MKWI 2018

Tagungsband

Multikonferenz Wirtschaftsinformatik 2018

Data driven X — Turning Data into Value

Band I

Paul Drews, Burkhardt Funk, Peter Niemeyer und Lin Xie (Hrsg.)

6. - 9. März 2018, Leuphana Universität Lüneburg

LEUPHANA
UNIVERSITÄT LÜNEBURG

Paul Drews, Burkhardt Funk, Peter Niemeyer und Lin Xie (Hrsg.)

Multikonferenz Wirtschaftsinformatik 2018

Data driven X — Turning Data into Value

Leuphana Universität Lüneburg
6. - 9. März 2018

Band I

LEUPHANA
UNIVERSITÄT LÜNEBURG

Impressum

Leuphana Universität Lüneburg
Institut für Wirtschaftsinformatik
Universitätsallee 1
21335 Lüneburg

ISBN 978-3-935786-72-0

mkwi2018.leuphana.de

Vorwort der Tagungsleitung

Die Multikonferenz Wirtschaftsinformatik (MKWI) ist eine etablierte im zweijährigen Turnus stattfindende Konferenz der Wirtschaftsinformatik-Community. In diesem Jahr findet die MKWI vom 6. – 9. März 2018 an der Leuphana Universität Lüneburg statt und bietet mit 29 Teilkonferenzen und Plenumsveranstaltungen ein Forum für den Austausch und die Diskussion innovativer Forschungsergebnisse.

2018 steht die MKWI unter dem Motto **„Data driven X – Turning Data into Value“** und adressiert damit die für Unternehmen hochaktuelle und potenzialträchtige Frage, wie die exponentiell wachsenden Datenbestände in Unternehmen optimal genutzt werden können. Das Konferenzmotto wird insbesondere in den Teilkonferenzen Analytics in Mobility, Business Analytics, cyber-physische Systeme, Data Analytical Processes in IS, digitale Ökosysteme und Plattformen, E-Commerce, E-Health, Internet of Value and Blockchain sowie Social Media Analytics wieder aufgegriffen. Darüber hinaus zeigt die Breite der Themen in den Teilkonferenzen, dass die Wirtschaftsinformatik in der Forschung Kernthemen der digitalen Transformation aufgreift und vorantreibt. Mit dem „Student Track“ zeigt die Wirtschaftsinformatik-Community erneut ihren Einsatz für den wissenschaftlichen Nachwuchs.

Insgesamt wurden 342 wissenschaftliche Beiträge eingereicht. Von 297 eingereichten Vollbeiträgen wurden 162 für die Präsentation auf der Konferenz akzeptiert (Annahmequote 55%). Der Begutachtungsprozess erfolgte in den einzelnen Teilkonferenzen und wurde von den jeweiligen Teilkonferenzleitungen organisiert. Dabei wurde jeder Beitrag in anonymisierter Form von mindestens zwei Personen begutachtet. Unser großer Dank geht an die Leitungen der Teilkonferenzen für die Organisation des Begutachtungsprozesses sowie an die Programmkomitee-Mitglieder und die Gutachter für ihre Unterstützung.

Eine wissenschaftliche Konferenz mit mehr als 700 Teilnehmerinnen und Teilnehmern ist ohne die Unterstützung vieler engagierter Personen undenkbar. Von den mehr als 50 beteiligten Personen gilt unserer besonderer Dank Stefanos Dimitriadis, der als Projektleiter mit großem Geschick die Fäden über die letzten eineinhalb Jahre zusammengehalten hat, sowie Sabine Prigge und Madlen Schmaltz, die die unzähligen kleinen und großen Herausforderungen einer erfolgreichen Konferenz mit beeindruckendem Engagement angenommen und gelöst haben. Unser Dank gilt auch Christoph Martin für die Begleitung des Begutachtungsprozesses und die Erstellung des vorliegenden Tagungsbands.

Last but not least, danken wir ausdrücklich unseren Sponsoren, die die Wirtschaftsinformatik-Community unterstützen und durch die diese Konferenz finanziell aber auch inhaltlich erst möglich wurde. Insbesondere danken wir der SENACOR Technologies und der AXA Versicherung als Hauptsponsoren sowie Dell EMC, Volkswagen Financial Services, Werum und 29forward als Sponsoren. Als Medienpartner unterstützen die Konferenz: BISE, GI, IHK Lüneburg-Wolfsburg, Landeszeitung und Springer Verlag.

Lüneburg, im März 2018

Paul Drews   Burkhardt Funk   Peter Niemeyer   Lin Xie

Organisation der MKWI 2018

Konferenzleitung

Paul Drews
Burkhardt Funk
Peter Niemeyer
Lin Xie

Konferenzorganisation

Martina Darda
Karin Dening-Bratfisch
Stefanos Dimitriadis
Philip Dunkhase
Andrea Japsen
Anna Lena Johannsen
Christoph Martin
Sabine Prigge
Marc Stephan Riedel
Madlen Schmaltz
Norbert Tschritter

Konferenzteam

Christian Albrecht | Erik Balzer | Dennis Becker | Daje Brahms | Timm Brümmendorf | Björn Brunke | Luca Cardinali | Kim Lasse Carlsen | Melissa Dogan | Guenay Dogan | Salina Dürrlich | Jonas Fährmann | Rigmor Franke | Fynn Freund | Janna Geesen | Sophie Wilhelmine Godding | Christoph Gooß | Tim Hadler | Michel Hansen | Florian Hermes | Lisa Heubner | Bianca Hoth | Marilen Jungandreas | Silas Fenimore Kißmer | Patrick Klose | Jonas Kößmeier | Fabian Kramer | Emir Kucevic | Michelle La | Klara Lentz | Aina Lindlahr | Fritz Lüders | Lisa Maslanka | Korab Mehmeti | Kevin Mischlich | Elena-Sophie Muhle | Yannic Neubauer | Mareike Neuser | Yannek Nyenhuis | Philip Ostermann | Anne-Sophie Ostmeyer | Amos Pietzcker | Lukas Reck | Denica Reimer | Felix Scheffler | Mathilde Schmidt-Rhen | Altje-Marie Schneider | Lena Schöning | Tim Schulze | Carlotta Schütte | Christoph Söhnlein | Anna-Prisca Sommerfeldt | Annika Steckelberg | Esther Stephan | Bastian Tamke | Nils Matti Thieme | Lukas Tippmann | Kim Christine Torster | Christopher Tschubel | Timon Tyzler | Dustin Uetrecht | Tina Wedmann | Leon Weis | Julia Woronko

Teilkonferenzleitungen

Frederik Ahlemann | Rainer Alt | Christine Bauer | Roman Beck | Benjamin Wagner Vom Berg | Daniel Beverungen | Eva Bittner | Michael Breitner | Benedikt Buchner | Catherine Cleophas | Marcus Dapp | Thomas Deelmann | Peter Dietrich | Paul Drews | Jan Fabian Ehmke | Carsten Felden | Peter Fettke | Kathrin Figl | Christoph Flath | Bogdan Franczyk | Albrecht Fritsche | Hansjoerg Fromm | Burkhardt Funk | Martin Gersch | Ralph Grothmann | Nadine Guhr | Jorge Marx Gómez | Axel Hahn | Andreas Harner | Bernd Heinrich | Sara Hofmann | Christian Huemer | Christian Janiesch | Harald Kienegger | Dennis-Kenji Kipker |

Michael Koch | Helmut Krcmar | Heidi Krömker | Christoph Lattemann | Ulrike Lechner | Christine Legner | Sebastian Lehnhoff | Jan Marco Leimeister | Christian Leyh | Thomas Lux | Martin Matzner | Philipp Melzer | Kathrin Möslein | Egon Müller | Peter Niemeyer | Boris Otto | Sven Overhage | Hans-Gert Penzel | Michael Prilla | Marco Pruckner | Alexander Richter | Susanne Robra-Bissantz | Hannes Rothe | Steffi Rudel | Thomas A. Runkler | Michael Räckers | Bernd Schenk | Sebastian Schlauderer | Hannes Schlieter | Eric Schoop | Mareike Schoop | Reinhard Schütte | Martin Sedlmayr | Thomas Setzer | Dirk Stelzer | Stefan Stieglitz | Susanne Strahringer | Marlin Ulmer | Nils Urbach | Edzard Weber | Peter Weber | Christof Weinhardt | Markus Weinmann | Lauri Wessel | Tobias Wienken | Christoph Wunck | Lin Xie | Gregor Zellner

Programmkomitee

Witold Abramowicz | Stephan Aier | Antonia Albani | Lars Algermissen | Rainer Alt | Jörn Altmann | Elske Ammenwerth | Nicolas Apostolopoulos | Christian Appel | Hans-Knud Arndt | Uwe Arnold | Sören Auer | Henning Baars | Constance Baban | Andrea Back | Mark Christopher Ballandies | Dian Balta | Ana Paula Bezerra Barquet | Christian Bauer | Sabine Baumann | Tobias Baumgaertner | Ulrike Baumöl | Felix Becker | Jörg Becker | Helmut Beckmann | Daniel Beimborn | Martin Benedict | Frank Bensberg | Jan Betzing | Jürgen Beyerer | Markus Bick | Michael Blaschke | Alexander von Bodisco | Franziska Boehm | Matthias Brand | Tobias Brandt | Michael Breidung | Berhard Breil | Rüdiger Breitschwerdt | Ulrich Bretschneider | Marius Brinkmann | Jan Vom Brocke | Tobias Brogt | Axel Bruns | Julian Bruns | Arne Buchwald | Christoph Buck | Hans Ulrich Buhl | Angelika Bullinger-Hoffmann | Jan Bumberger | Tilo Böhmann | Georg Carle | Oliver Christ | Hashim Chunpir | Catherine Cleophas | Nico Clever | Patrick Delfmann | Benjamin Dietrich | Barbara Dinter | Clemens van Dinther | Viktor Dmitriyev | Verena Dorner | Andreas Drechsler | Gabi Dreo | Sonja Dreyer | Martin Ebner | Andreas Eckhardt | David Eckhoff | Jan Ehmke | Jörn Eichler | Claus-Peter H. Ernst | Werner Esswein | Oliver Eulenstein | Torsten Eymann | Bruce Ferwerda | Peter Fettke | Stefan Feuerriegel | Hans-Georg Fill | Ralf Finger | Simon Fischer | Daniel Fischer | Christoph Flath | Carlos Rodriguez Flick | Bogdan Franczyk | Ulrich Frank | Lena Frankenhauser | Gilbert Fridgen | Thomas Friedrich | Christoph Fuchss | Burkhardt Funk | Jens Förderer | Daniel Fürstenau | Roland Gabriel | Nick Gehrke | Reinhard German | Martin Gersch | Heiko Gewald | Steffen Gilge | Henner Gimpel | Volkan Gizli | Matthias Goeken | Andreas Grahl | Matthias Griebel | Norbert Gronau | Tom Gross | Jürgen Großmann | Andreas Grzemba | Karin Gräslund | Alex Gröschel | Nadine Guhr | André Göbel | Jochen Gönsch | Janine Hacker | Marit Hansen | Basel Hasan | Nicolas Haubner | Reinhold Haux | Robert Heidekrüger | Irina Heimbach | Moreen Heine | Andreas Heinemann | Bernd Hellingrath | Patrick Helmholz | Thomas Herrmann | Georg Herzwurm | Thomas Hess | Markus Heuchert | Andreas Hilbert | Stefan Hirschmeier | Patrick Hoberg | Roland Holten | Bo Hu | Paul Huttner | Anke Häber | Ralf Isenmann | Christian Janssen | Lutz Johanning | Florian Johannsen | Jan Juerjens | Reinhard Jung | Anna Jung | Dimitris Karagiannis | Manfred Kirchgeorg | Kathrin Kirchner | Stefan Kirn | Stefan Klein | Mathias Klier | Julia Klier | Thomas Koehler | Lutz Kolbe | Henning Kontny | Jana Korunovska | Agnes Koschmider | Johann Kranz | Hanna Krasnova | Helmut Krcmar | Irene Krebs | Simone Kriglstein | Max Krüger | Thomas Kude | Niklas Kuehl | Dennis O. Kundisch | Hartmut König | Sabine Köszegi | Angelina Lange | Kevin Laubis | Ulrike Lechner | Christine Legner | Jan Marco Leimeister | Susanne Leist | Richard Lenz | Ulrike Leopold-Wildburger | Joerg Leukel | Karl Liebstueckel | Peter Loos | Helena Lovasz-Bukvova | Carsten Lucke | André Ludwig | Thomas Ludwig | Stephan Lukosch | Reinhard Madlener | Ronald Maier | Christian Maier |

Marian Margraf | Bernd Markscheffel | Christian Matt | Dirk Christian Mattfeld | Florian Matthes | Florian Maurer | Pascal Meier | Michael Meier | Jan Mendling | Christian Meske | Udo Milkau | Milad Mirbabaie | Christoph Mittendorf | Jürgen Moormann | Volkmar Mrass | Thomas Myrach | Alexander Mädche | Andreas Möller | Kathrin Möslein | Sven Müller | Benjamin Müller | Dirk Neumann | Peter Niemeyer | Astrid Niesse | Volker Nissen | Markus Nuettgens | Michael Nürnberg | Markus Nüttgens | Andreas Oberweis | Sebastian Olbrich | Boris Otto | Barbara Paech | Mario Parade | Christoph Peters | Stefan Pickl | Gunther Piller | Niels Pinkwart | Volkmar Pipek | Jens Poeppelbuss | Wolfgang Prinz | Jens Pöppelbuß | Gerald Quirchmayr | Kai Rannenberg | Holger Raphael | Barbara Rapp | Asarnusch Rashid | Matthias Rass | Michael Rebstock | Beke Redlich | Olaf Reinhold | Florian Remark | Wolfgang Renz | Christian Reuter | Sebastian Richter | Konrad Rieck | René Riedl | Stefanie Rinderle-Ma | Susanne Robra-Bissantz | Thomas Rose | Michael Rosemann | Kristina Rosenthal | Sven Rosinger | Peter Rossbach | Volker Roth | Martin Roth | Hannes Rothe | Nico Rödder | Alexander Sandau | Kurt Sandkuhl | Gerhard Satzger | Jürgen Sauer | Thomas Schaeffer | Carsten Schauer | Kevin Scheibe | Christian Scheiner | Birgit Schenk | Barbara Schiller | Thomas Schlegel | Hannes Schlieter | Jenny Schoch | Detlef Schoder | Hendrik Scholta | Peter Schuler | Carsten Schultz | Volker Schöber | Isabella Seeber | Martin Semmann | Thomas Setzer | Katerina Shapoval | Dominik Siemon | Markus Siepermann | Alexander Simons | Elmar Sinz | Stefan Smolnik | Matthias Soellner | Karsten Sohr | Andreas Solsbach | Carsten Sommer | Thorsten Staake | Katarina Stanoevska | Ulrike Steffens | Nikolai Stein | Gunnar Stevens | Volker Stich | Stefan Stieglitz | Steffen Stock | Alexander Stocker | Susanne Strahringer | Stefan Strecker | Nick Street | Matthias Stürmer | Jan Sudeikat | Safwan Sulaiman | Ali Sunyaev | Matthias Söllner | Fabian Taigel | Alfred Taudes | Frank Teuteberg | Oliver Thomas | Verena Tiefenbeck | Kowatsch Tobias | Manuel Trenz | Apostolos Tryfonidis | Martin Tröschel | Klaus Turowski | Rainer Unland | Daniel Veit | Rudolf Vetschera | Oliver Vins | Michael Voessing | Bernhard Volz | Gottfried Vossen | Heinz-Theo Wagner | Konrad Walser | Kristin Weber | Anke Weidlich | Stefan Weidner | Markus Weinmann | Dirk Werth | Hannes Werthner | Mathias Weske | Lauri Wessel | Markus Westner | Thomas Widjaja | Martin Wiener | Axel Winkelmann | Till Winkler | Robert Winter | Dominique Winter | Alfred Winter | Rikka Wittstock | Matthias Wittwer | Clemens Wolff | Andreas Wombacher | Volker Wulf | Stefan Wunderlich | Marius Wybrands | Alex Zarifis | Rüdiger Zarnekow | Torsten Zesch | Hans-Dieter Zimmermann | Frank Öchsle

Reviewer

Olga Abramova | Witold Abramowicz | Safa'A Abujarour | Frederik Ahlemann | Stephan Aier | Antonia Albani | Lars Algermissen | Rainer Alt | Jörn Altmann | Boris Amberg | Elske Ammenwerth | Lorenz Ammon | Gabriele Anderst-Kotsis | Nicolas Apostolopoulos | Christian Appel | Hans-Knud Arndt | Sören Auer | Henning Baars | Constance Baban | Andrea Back | Nikolina Bader | Lucas Baier | Mark Ballandies | Dian Balta | Victoria Banken | Ana Barquet | Ana Paula Bezerra Barquet | Marcus Basalla | Verena Batt | Christine Bauer | Christian Bauer | Sabine Baumann | Annika Baumann | Ruth Baumgart | Ulrike Baumöl | Roman Beck | Dennis Becker | Felix Becker | Jörg Becker | Helmut Beckmann | Kathrin Bednar | Jannis Beese | Daniel Beimborn | Martin Benedict | Frank Bensberg | Benjamin Wagner Vom Berg | Sabine Berghaus | Arne Bergmann | Jan Betzing | Daniel Beverungen | Manoj Bhat | Markus Bick | Christian Bierwirth | Matthias Billert | Eva Bittner | Michael Blaschke | Alexander Bock | Florian Bock | Alexander von Bodisco | Tilo Böhmann | Nicolai Bohn | Lars Bollweg | Dominik Bork | Hendrik Brakemeier | Matthias Brand | Tobias Brandt | Laura Braun | Daniel Braunnagel | Michael Breidung | Michael H. Breitner | Rüdiger Breitschwerdt | Vincent

Bremer | Ulrich Bretschneider | Marius Brinkmann | Jan Vom Brocke | Julian Bruns | Arne Buchwald | Christoph Buck | Hans Ulrich Buhl | Angelika Bullinger-Hoffmann | Jan Bumberger | Marc Burkhalter | Simone Burster | Georg Carle | Alexandra Cetto | Peter Chamoni | Oliver Christ | Hashim Chunpir | Catherine Cleophas | Nico Clever | Carl Corea | Marcus Dapp | Thomas Deelmann | Patrick Delfmann | Amra Delić | Benjamin Dietrich | Barbara Dinter | Clemens van Dinther | Sven Dittes | Viktor Dmitriyev | Katharina Drechsler | Andreas Drechsler | Paul Drews | Sonja Dreyer | Christoph Drodt | Niklas Ebell | Martin Ebner | Andreas Eckhardt | David Eckhoff | Jan Fabian Ehmke | Jörn Eichler | Adrian Engelbrecht | Elias Englmeier | Claus-Peter H. Ernst | Markus Esswein | Werner Esswein | Oliver Eulenstein | Torsten Eymann | Torsten Fahle | Carsten Felden | Bruce Ferwerda | Peter Fettke | Stefan Feuerriegel | Kathrin Figl | Hans-Georg Fill | Simon Fischer | Daniel Fischer | Thomas Fischer | Christoph Flath | Carlos Rodriguez Flick | Jens Förderer | Bogdan Franczyk | Ulrich Frank | Daniel Franzmann | Gilbert Fridgen | Thomas Friedrich | Albrecht Fritzsche | Christoph Fuchß | Burkhardt Funk | Daniel Fürstenau | Kai Gand | Nick Gehrke | Manfred Geiger | Ulf Gerhardt | Reinhard German | Martin Gersch | Heiko Gewald | Steffen Gilge | Henner Gimpel | Volkan Gizli | André Göbel | Jochen Gönsch | Andreas Grahl | Karin Gräslund | Matthias Griebel | Norbert Gronau | Tom Gross | Jürgen Großmann | Christian Grotherr | Andreas Grzemba | Nadine Guhr | Jens Gulden | Knut Haase | Anke Häber | Janine Hacker | Simon Hacks | Axel Hahn | André Halckenhäußer | Julian Hamann | Timo Hammerl | Richard F. Hartl | Patrick Hartmann | Martina Hartner-Tiefenthaler | Basel Hasan | Nicolas Haubner | Reinhold Haux | Irina Heimbach | Moreen Heine | Robert Heininger | Patrick Helmholz | Thomas Herrmann | Matthias Herterich | Georg Herzwurm | Thomas Hess | Markus Heuchert | Stefan Hirschmeier | Robin Hirt | Sven Hoffmann | Sara Hofmann | Roland Holten | Bettina Horlach | Thomas Leopold Huber | Christian Huemer | Marcus Hülsdau | Dominik Huth | Paul Huttner | Florian Imgrund | Ralf Dr. Isenmann | Ralf Isenmann | Roope Jaakonmäki | Christian Janiesch | Christian Janssen | Florian Johannsen | Marco Johns | Felix Jordan | Christopher Jud | Jan Juerjens | Anna Jung | Reinhard Jung | Ekaterina Jussupow | Lena Kaiser | André Kaiser | Piotr Kałużny | Dimitris Karagiannis | Robert Keller | Ateeq Khan | Harald Kienegger | Sybren de Kinderen | Dennis-Kenji Kipker | Stefan Kirn | Stefan Klein | Julia Klier | Mathias Klier | Christoph Klingenberg | Ralf Knackstedt | Michael Koch | Thomas Koehler | Hartmut Koenig | Henning Kontny | Jakob Johannes Korbel | Jana Korunovska | Agnes Koschmider | Claudia Koschtial | Antonia Köster | Johann Kranz | Hanna Krasnova | Helmut Krcmar | Simone Kriglstein | Max Krüger | Thomas Kude | Niklas Kuehl | Niklas Kühl | Patrick Kummler | Dennis O. Kundisch | Gerrit Küstermann | Olga Labazova | Angelina Lange | Christoph Lattemann | Kevin Laubis | Ralf Laue | Ulrike Lechner | No-San Lee | Christine Legner | Sebastian Lehnhoff | Christiane Lehrer | Jan Marco Leimeister | Susanne Leist | Richard Lenz | Daniel Leonhardt | Joerg Leukel | Christian Leyh | Karl Liebstueckel | Alexander Löffler | Peter Loos | Gustavo Lopez | Helena Lovasz-Bukvova | Carsten Lucke | André Ludwig | Thomas Ludwig | Stephan Lukosch | Alexander Mädche | Maria Madlberger | Christian Maier | Ronald Maier | Monika Malinova | Bernd Markscheffel | Christoph Martin | Christian Matt | Dirk Mattfeld | Dirk Christian Mattfeld | Florian Matthes | Martin Matzner | Florian Maurer | Michael Meier | Pascal Meier | Frank Meisel | Philipp Melzer | Jan Mendling | Alexander Merz | Alexander Benedikt Merz | Christian Meske | Udo Milkau | Milad Mirbabaie | Tobias Mirsch | Cristoph Mittendorf | Christoph Mittendorf | Andreas Mladenow | Frederik Möller | Andreas Möller | Jürgen Moormann | Stefan Morana | Marcel Morisse | Kathrin Möslein | Volkmar Mrass | Benjamin Müller | Christoph Müller-Bloch | Thomas Myrach | Esther Nagel | Oliver Neumann | Jürgen Neumann | Peter Niemeyer | Steffen Nienke | Astrid Niesse | Astrid Nieße | Volker Nissen | Markus Noeltner | Mario Nolte | Markus Nuettgens | Michael Nürnberg | Markus

Nüttgens | Andreas Oberweis | Sebastian Olbrich | Stefan Oppl | Hubert Österle | Uwe Ostermann | Boris Otto | Sven Overhage | Barbara Paech | Mario Parade | Ute Paukstadt | Christoph Peters | Florian Pethig | Stefan Pickl | Jan-Hendrik Piel | Gunther Piller | Fabiano Pinatti | Niels Pinkwart | Ante Plazibat | Jens Poeppelbuss | Jens Pöppelbuß | Loina Prifti | Michael Prilla | Wolfgang Prinz | Marco Pruckner | Jochen Pruemper | Thomas Quirchmar | Thomas Quirchmayr | Maximilian Raab | Michael Räckers | Kai Rannenberg | Holger Raphael | Barbara Rapp | Asarnusch Rashid | Matthias Raß | Michael Rebstock | Beke Redlich | Jana-Rebecca Rehse | Tim Reimers | Olaf Reinhold | Florian Remark | Christian Remfert | Wolfgang Renz | Christian Reuter | Sebastian Prof. Dr. Richter | Alexander Richter | Andreas Rieb | Konrad Rieck | René Riedl | Stefanie Rinderle-Ma | Daniel Ritter | Susanne Robra-Bissantz | Nico Rödder | Kristina Rosenthal | Sven Rosinger | Patrick Roßner | Martin Roth | Hannes Rothe | Steffi Rudel | Alexander Sandau | Kurt Sandkuhl | Gerhard Satzger | Jürgen Sauer | Silvia Schacht | Thomas Schaeffer | Thomas Schäffer | Carsten Schauer | Kevin Scheibe | Johannes Schlaich | Darius Schlangenotto | Sebastian Schlauderer | Thomas Schlegel | Hannes Schlieter | Christopher Schmitz | Michael Schneider | Sofia Schöbel | Volker Schöber | Detlef Schoder | Hendrik Scholta | Jörn Schönberger | Eric Schoop | Maximilian Schreieck | Melanie Schreiner | Carsten Schultz | Reinhard Schütte | Josef Schwaiger | Martin Sedlmayr | Isabella Seeber | Martin Semmann | Thomas Setzer | Dominik Siemon | Markus Siepermann | Alexander Simons | Elmar Sinz | Stefan Smolnik | Matthias Soellner | Karsten Sohr | Matthias Söllner | Andreas Solsbach | Carsten Sommer | Marc-Oliver Sonneberg | Andreas Sonntag | Jeannette Stark | Ulrike Steffens | Nikolai Stein | Dirk Stelzer | Tobias Steudner | Gunnar Stevens | Volker Stich | Stefan Stieglitz | Steffen Stock | Alexander Stocker | Susanne Strahringer | Stefan Strecker | Nick Street | Matthias Stürmer | Jan Sudeikat | Safwan Sulaiman | Ali Sunyaev | Fabian Taigel | Alfred Taudes | Navid Tavanapour | Oliver Terhaar | Frank Teuteberg | Oliver Thomas | Verena Tiefenbeck | Matthias Tietz | Kowatsch Tobias | Manuel Trenz | Martin Tröschel | Apostolos Tryfonidis | Klaus Turowski | Marlin Ulmer | Ömer Uludag | Rainer Unland | Nils Urbach | Svitlana Vakulenko | Daniel Veit | Eric Veith | Rudolf Vetschera | Oliver Vins | Michael Voessing | Patrick Vogel | Kai-Ingo Voigt | Matthias Volk | Bernhard Volz | Stefan Voss | Gottfried Vossen | Amina Wagner | Heinz-Theo Wagner | Michael Walch | Jonas Wanner | Peter Weber | Kristin Weber | Edzard Weber | Charlotte Wehking | Benjamin Wehner | Julian Weidinger | Anke Weidlich | Stefan Weidner | Christoph Weinert | Markus Weinmann | Michael Werner | Johannes Werner | Oliver Werth | Dirk Werth | Hannes Werthner | Mathias Weske | Lauri Wessel | Markus Westner | Arnold Wibmer | Thomas Widjaja | Martin Wiener | Axel Winkelmann | Till Winkler | Alfred Winter | Karolin Winter | Dominique Winter | Robert Winter | Anne-Katrin Witte | Rikka Wittstock | Matthias Wittwer | Isabell Wohlgenannt | Clemens Wolff | Anselma Wörner | Volker Wulf | Christoph Wunck | Stefan Wunderlich | Bastian Wurm | Marius Wybrands | Lin Xie | Alex Zarifis | Ruediger Zarnekow | Rüdiger Zarnekow | Novica Zarvic | Torsten Zesch | Hans-Dieter Zimmermann | Benedikt Zobel

Sponsoren der MKWI 2018

Inhaltsverzeichnis

Analytics in Mobility: Planning for Transportation and Logistics

Teilkonferenzleitung

Catherine Cleophas
Jan Ehmke
Marlin Ulmer

Truck Platooning: Towards Future Business Models

Nicolai Krüger[1] and Frank Teuteberg[1]

[1] University of Osnabrueck, Accounting and Information Systems, Osnabrueck, Germany
{nikrueger, frank.teuteberg}@uni-osnabrueck.de

Abstract. Automated driving trucks promise significant improvements with regards to traffic, logistic costs and emissions. Platooning can be seen as the next technological step in this direction. While the body of literature reflects the engineering perspective of platooning extensively, an Information Systems discussions about user-centered business models are rare. Our paper aims to strengthen the Information Systems perspective on platooning and its inherent business models. Up to the current state of our project, we found that intra-fleet platooning can be understood as an amortization issue, while inter-fleet platooning requires complex matching algorithms, motivational incentives for drivers and fleets and easy plus fast payment solutions. We built a Monte Carlo simulation for an inter-fleet platooning app and present preliminary results.

Keywords: Platooning, Autonomous Driving, Logistic Research, Logistics, Vehicle2Vehicle Communication

1 Introduction

1.1 Problem Relevance and Research Questions

Platooning can be defined as "(…) coupling two or more vehicles without a physical link to form a train" [14]. Public challenges with wireless connected trucks took already place in California, Japan and Europe; across national borders and under real traffic conditions [16]. The primary motivation for engineers to develop platooning was for sure the fuel-saving potential. It is a scalable factor for transportation companies even with little savings because fuel equals one-third of the operating costs of a heavy duty vehicles (HDV) [10]. Platooning of long-haul trucks is moving more and more from a science-fiction to a real-life technology. However, while technological confidence is evolving, research rarely covers the economic aspects of platooning, as pointed out in [3]: "One of the greatest neglects in the literature is how platooning will be developed and operated in practice." and further "It must also be understood that there will form a business ecosystem around platooning, with many different actors cooperating and competing with each other, and this raises important issues related to safety, due to interoperability requirements and shared responsibilities".

What makes platooning very interesting for researchers, is the multi-perspective challenge of it [15]. Thus, an Information System (IS) view can support the research initiatives in this field to set impulses into a user-centric added value of this new

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

technology and business opportunity. For the latter one, some basic ideas can be found in the literature, for instance [17] sees tremendous opportunities for truck fleets in the mining industry, but also opens up the wide field for the entire logistics branch. As the amount of potential savings will matter, linking as many trucks as possible might be a good choice. In [10], up to 200 trucks get connected, which brings up the important question, how platooning partners will find their matches; may it be within their own fleet and transport company or may it be in the open market. We summarize the described aspects with the following research questions: **RQ1.** Which basic business models can be described for intra-fleet and inter-fleet platooning? **RQ2.** How could a market for inter-fleet platooning look like, based on the current data available?

The paper is structured as follows: In the next section, we take a brief look at related articles and – in the context of this paper – relevant platooning technologies. Based on that, we create a concept of a business model for platooning in section 2. Then, a simulation of our business model will be conducted, using Monte Carlo. In section 3, we wrap up this research in progress paper, anticipating learnings of the simulation and sharing an outlook onto the future research and project scope.

1.2 Related Work

The literature research will form the baseline for section 2, where we aim to build a realistic market scenario for platooning. We conduct the review in common databases, including SCOPUS, ScienceDirect, AISEL and Google Scholar. To ensure proper quality, we limit the search to articles from peer-reviewed journals and conferences.

Synapsis. While automated driving in the passenger car segment came up already in the 1950s, experiments with autonomous trucks started only in the 1990s [16]. Practical tests proved already in 1995 the possibility of visual-based platooning. But even in such an early stage of research in the 90s, researchers were aware of the fact which impact platooning will have on the way to fully autonomous driving in technological and its' economic perspectives [7]. General benefits of platooning are fuel saving and a variety of other advantages: [1] analyzed, that 72% of greenhouse gas emissions has been caused by road transportation; in [16] safety and drivers' comfort are reflected, of course under the assumption, that the long-term strategy of automated trucks is driverless.

Fuel saving potential. It is common sense that shortening the distance between trucks reduces the air drag, which reduces the fuel consumption. A study with 1,800 heavy-duty trucks driving through Europe, conducted in [11], shows that spontaneous, *manual* platoons were formed in 1.2% of all routes by the drivers, simply by reducing the safety clearance to the max. However, such a non-optimized and non-coordinated approach only let to 0.07% fuel savings overall. Thus, a technological platooning solution is obviously desirable. Early research of truck platooning predicted potential savings in the range from 30 to 40 percent. In [2], a fuel reduction between 4.7 and 7.7% has been achieved, taking driving two identical loaded trucks at 70 km/h. But fuel savings vary based on the individual circumstances, for instance [12] was able to confirm a 7% fuel reduction for a 350km trip driving with 80km/h. In contrast to this, [4] achieved in their experiments a saving potential from 12.1 to 19.8%, driving at a speed of 60.1 km/h. Theoretical savings for the non-leading vehicle up to 20% are also described in [10], but the authors identified the speed of the platoon, and the distance between the trucks as critical factors. Thus, a realistic saving range from 9-10% could be achieved, if all trucks start at the same node of a route. [2] has also shown, that the load weight and the time gap between the trucks determine the fuel saving.

Economic scenarios and use cases. Scania and Volvo built industry-ready solutions for well-defined scenarios, like mining, logistics or industry. Thus, productivity could be increased, as well as safety for other vehicles and the truck drivers, especially in dangerous mining or tunnel situations. But logistics and industry solutions in general also describe a huge potential for platooning [17].

An unsolved question is an incentivization model. Why should a truck driver offer platooning (and thus, fuel saving potential) to following trucks, instead of joining an existing platoon chain? In [6] a game-theoretic approach focusing on road traffic can be found, where platoons get incentivized but face a trade-off concerning road traffic, time and so forth.

2 Towards a Business Model for a Truck Platooning Platform

2.1 Business Model Conception

Intra-fleet platooning. Seeking for future business models which can utilize this technology, our approach follows two steps: Based on the literature research presented, intra-fleet platooning and inter-fleet platooning can be clustered; both may inherent own business models. To answer the research question, we strive for mathematical or statistical evidence. We analyze intra-fleet platooning first by seeking for a mathematical expression. The economic potential can be seen as the geographical area between the original cost function and the new cost function with platooning:

$$p_{savings} \int_a^b [c(x) - c_p(x)]\, dx \tag{1}$$

where c(x) is the cost function without, cp(x) with platooning and they simply differentiate in the fact that platooning vehicles will have a reduced diesel consumption rate (which will not be further elaborated here). Thus, the area between the original and the new cost graph geometrically symbolize the saving potential of platooning within, for instance, a mining fleet. Typically, this potential will follow the bandwidth of 7-9%, as discussed in section 1.2. As [16] has shown, the cost for platooning technology will consist of the cost of the Adaptive Cruise Control (ACC) and the V2V communication system. But even taking labor costs into account, it remains an amortization issue. In other words, to give the first answer to our research question:

Intra-fleet platooning is a cost-saving business opportunity, but not a new business model on its' own.

Further, an intra-fleet platoon might be organized by the fleet company itself, without the need of further IT-infrastructure: Drivers or dispatchers can match schedules and routes to achieve a maximal platooning-based efficiency inside the company. Out of scope for this paper but a crucial factor in the daily routine of logistics companies might become the driver rating of the digital tachograph or fleet optimization software like Fleetboard. While the driver rating of all following drivers will increase, based on the reduced diesel consumption, any solution and benefit for the leading driver must be created to prevent penalties.

Inter-fleet platooning. Therefore, we continue to a second perspective, the inter-fleet platooning to evaluate this business model in terms of cost and revenue streams. As nothing comparable is existing in the presented body of literature we have to come up with a platooning ecosystem, following the analogy of car sharing providers [8]: We build the hypothesis, that there will be a business model for the platform which brings the different platooning partners together. This asset should solve the problem of needed starting nodes in a highway nets, as discussed in [10]. In the following, we are going to describe the players in that market more in detail.

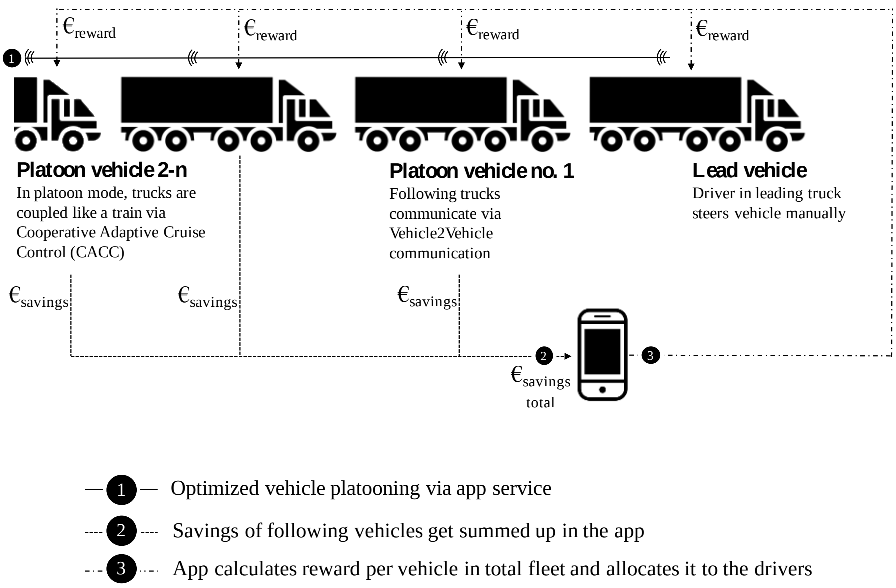


Figure 1: Scheme of a platooning app business model, matching trucks in a platoon virtually to share the diesel savings belong all participants in the platoon

Fig. 1 describes our business model, where we see one central app provider as matchmaker between potential platooning partners. The app would detect route parameters, geo-positions and weight/speed indicators from the truck onboard system, which allows a precise and vehicle-individual computation of fuel savings. In some cases, this might happen via the Fleet Management Systems Interface (FMS) or an existing OEM telematics unit with a third-party access API, like Fleetboard (Daimler) or Rio (MAN). When the app computes a positive platooning matching chance, both trucks have to accept. An important assumption here is, that a truck can also decline a platooning offer and can't be forced to platoon, for instance by law, OEM technology or any other reason. Assuming that a platoon activation follows a positive match, the app extracts the savings of all following vehicles from the trucks' telematics system, as described above. All needed parameters to do so are already part of today's FMS interface,[1] therefor the app would always be delivered together with a small hardware dongle which the trucker can attach easily to his/her truck. For our simulation, we iterate different business logics first, to find a good way to sum up and allocate savings to the group of a platoon. One idea could be to pay directly via e-wallets or blockchain technology to the truck in front. We skip that idea because it is not route-based and makes it hard to include the entire platoon with different weights (and therefore fuel savings, as earlier discussed based on the results of [12]). To tackle this problem and to grant maximum (financial) motivation to the truckers – to include the learnings of [6] – we decide to calculate the truck-individual diesel savings on the fly. The virtual saving amount will be transmitted to the app's backend, where all savings of the entire platoon for the time it exists in a specific constellation will be captured. When a platoon formation ends, the rewards will be calculated and virtually transferred to the in-app account of all trucks. For the reward calculation, we take the weighted portion of savings generated by a truck, relative to the entire platoon (important: including the lead vehicle, which won't generate savings but should participate in the total savings). By doing so, we motivate any potential leading truck starting a platoon.

2.2 Simulation

Simulations are a valid source of statistical evidence to cover those aspects of platooning, which can't be evaluated in real-life situation, such as the optimal distance for the case of emergency breaks of the leading platoon vehicle [13] or business opportunities and models for a market to come, as proposed in this paper.

Market parameters. The basic input value for our simulation is the amount of new truck registrations per year, which is around 200,000 trucks in Germany [9]. We assume 10% market share of platooning hardware within the newly sold trucks for our scenarios, which equals in total 19,961 platooning-ready trucks. This is based on the assumption, that platooning equipment will become a cost-intensive feature, and only early adopters will purchase the technology after the market launch. Next, we calculate the expected maximum distance of platoons for all platooning-ready trucks in Germany. Given a yearly mileage between 100,000 and 120,000 km for a regular fleet truck [5], the total potential distance for truck platooning would end up (taking the average of

[1] See also the official FMS web page http://www.fms-standard.com/ (accessed: 9/26/17)

110,000 km mileage per year) at 2,195,743,000 km. Reducing this number to a more realistic level, we considered the expected overlap of trucking routes with respect to time schedules, starting points, destinations and delivery stops. Thus, we see 5% (109,787,150 km) as the total platooning potential per year (for all trucks combined, having platooning equipment installed).

Finally, the business model of our platooning app is based on a small fee, which all truckers have to pay, who benefit from the platoon. This fee should be relative to the total platoon savings; we chose a range from 0.5 to 5% to have an appropriate and – from a user point of view – acceptable price and good statistical data to find to conduct further analytics on the price impact. At this point of research, we focus on fuel savings purely (and not driver efficiency etc.), which can be understood as a limitation in our paper. However, fuel savings can be calculated and simulated already today, without field data. Future research projects might allow the research community to enhance this.

Computation. The entire simulation algorithm, including every single truck, all routes and market conditions is implemented in Python 3.0.[2] Running the entire simulation took 15:36:25 hours on a MacBook Air with an Intel Core i7 processor running at 2.20 GHz, in total, more than 250 GB of platooning data has been created.

3 Conclusion and Outlook

In total, we ran the simulation 1,000 times with an average app fee of 2.69€ (standard deviation: 1.29), which led to an average app revenue per simulation of 5,654,936.01€ (standard deviation: 4,462,360.50€). The broad spread in the total app revenue is based on the mileage bandwidth, which was 54,778,745.14km in average (standard deviation: 31,708,886.03km). These first results underline the potential of a Monte Carlo simulation in our context. As the input and output parameters face such a huge bandwidth, we can start now selecting worst-, best- and most-likely cases out of the generated data and build a sensitivity analysis. We expect to get a clear understanding of minimum revenue, such an app would achieve, even under worst-case conditions. With that, a holistic concept for platooning business solutions, including the utilization of blockchain and smart contract technology, can follow. With this research in progress paper, we aim to communicate the current state of our project and to gather feedback from the research community.

[2] We also considered using VISSIM for our simulation, as well as UNITY (where also a platooning package exists), MATLAB and Python. Experimenting with these tools showed us, that we could neglect visual output for our simulation (which reduced our tests to MATLAB and Python). Due to personal preference of the authors, we finally chose Python 3.0 for the simulation itself and SPSS for the interpretation of the data.

References

1. Alam, A. et al.: Heavy-duty vehicle platooning for sustainable freight transportation: A cooperative method to enhance safety and efficiency. IEEE Control Syst. Mag. 35, 6, 34–56 (2015)
2. Alam, A. Al et al.: An Experimental Study on the Fuel Reduction Potential of Heavy Duty Vehicle Platooning. In: IEEE Annual Conference on Intelligent Transportation Systems. , Madeira Island (2010)
3. Axelsson, J.: Safety in Vehicle Platooning: A Systematic Literature Review. IEEE Trans. Intell. Transp. Syst. 18, 5, (2017)
4. Caltagirone, L. et al.: Truck Platooning Based on Lead Vehicle Speed Profile Optimization and Artificial Physics. In: IEEE Conference on Intelligent Transportation Systems. , Las Palmas (2015)
5. Daimler: Kraftstoffvergleichstests in Europa, https://www.mercedes-benz.com/de/mercedes-benz/fahrzeuge/lkw/kraftstoffvergleichstests-in-europa/
6. Farokhi, F., Johansson, K.H.: A game-theoretic framework for studying truck platooning incentives. In: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. pp. 1253–1260 , The Hague (2013)
7. Franke, U. et al.: Truck Platooning in Mixed Traffic Abstract. In: Intelligent Vehicles '95 Symposium. , Detroit (1995)
8. Hanelt, A. et al.: A Taxonomy of Carsharing Business Models. Int. Conf. Inf. Syst. September, 1–19 (2016)
9. Kraftfahrtbundesamt: Neuzulassungen von Lkw im Jahr 2016 nach ausgewählten Herstellern, https://www.kba.de/DE/Statistik/Fahrzeuge/Neuzulassungen/MarkenHersteller/2016_n_herst_lkw_dusl.html;jsessionid=BB0B9426B9A3D7A10DA43F0938B0821F.live21301?nn=658682
10. Larsson, E. et al.: The vehicle platooning problem: Computational complexity and heuristics. Transp. Res. Part C Emerg. Technol. 60, 258–277 (2015)
11. Liang, K., Jonas, M.: Fuel-Saving Potentials of Platooning Evaluated through Sparse Heavy-Duty Vehicle Position Data. Iv, 1061–1068 (2014)
12. Liang, K.Y. et al.: When is it fuel eficient for a heavy duty vehicle to catch up with a platoon? IFAC Proc. Vol. 7, PART 1, 738–743 (2013)
13. Nunen, E. Van et al.: Sensor Safety for the European Truck Platooning Challenge. In: ITS World Congress. , Melbourne (2016)
14. Royer, E. et al.: A FRAMEWORK FOR VEHICLE PLATOONING BASED ON MONOCULAR VISION. IFAC Proc. Vol. 40, 15, 294–299 (2007)
15. Segata, M. et al.: Toward communication strategies for platooning: Simulative and experimental evaluation. IEEE Trans. Veh. Technol. 64, 12, 5411–5423 (2015)
16. Tsugawa, S. et al.: A Review of Truck Platooning Projects for Energy Savings. IEEE Trans. Intell. Veh. 1, 1, 68–77 (2016)
17. Uhlemann, E.: Platooning: Connected Vehicles for Safety and Efficiency [Connected Vehicles]. IEEE Veh. Technol. Mag. 11, 3, 13–18 (2016)

Examining Delay Propagation Mechanisms for Aircraft Rotations

Lucian Ionescu, Natalia Kliewer

Department of Information Systems, Freie Universität Berlin,
Garystr. 21, 14195 Berlin, Germany
{lucian.ionescu, natalia.kliewer}@fu-berlin.de

Abstract.
State-of-the-art optimization software for robust airline resource scheduling are often based on stochastic approaches for which understanding and modeling realistic delay occurrence and propagation mechanisms are essential. In this context, this paper deals with the evaluation of a theoretical delay propagation model with real-world data of a major European carrier. The results show that especially in the event of delayed departure, the turnaround process can be accelerated and thus the operational robustness is underestimated. The findings can be used to estimate delay propagation in robust aircraft scheduling and simulation closer to operational reality.

Keywords: Robust Efficiency, Airline Resource Scheduling, Propagation, Simulation

1 Introduction

In the globalized world, the rising demand for air transportation leads to increased frequencies on existing flight routes and new destination offerings by airlines. This inevitably goes along with additional complexity of resource schedules since more crews, aircraft but also airport equipment for ground operations and scares airspace must be efficiently scheduled. A wide range of sophisticated commercial optimization suites and IT solutions has been developed in recent decades from specialized providers such as Sabre, SITA, Lufthansa Systems, Jeppesen or EDS, see [1] for details.

The traditional objective of cost minimization preferably leads to the efficient usage of resources at preferably high utilization levels. Since idle times are costly, ground times between flights tend towards a minimum in cost-optimized schedules. We refer to the costs for a schedule that can be operated as planned as nominal costs. However, airlines frequently have to deal with exogenous delays during operations, caused by bad weather conditions, technical failure or congestion, to name just a few. In case of insufficient buffer times, delays may propagate further on consecutive flights, inducing additional reactionary costs for schedule recovery, resource

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

reassignment or passenger rerouting. The real costs that actually emerge for airline operations can therefore substantially differ from nominal costs.

This issue is addressed by the concept of *robust efficiency* which aims at the minimization of real costs by already taking into account reactionary costs during the scheduling stage. Besides cost-efficiency, the additional objective of robustness is considered, aiming at the reduction of delay propagation effects. A stronger coupling between scheduling and operations can be achieved.

In recent decades, stochastic optimization approaches have been developed to improve either the stability [2]-[4] or flexibility [5] of resource schedules, mostly focusing on crews and aircraft. The benefit of these approaches greatly depends on realistic assumptions on occurrence probabilities of exogenous delays. Related studies can be found in [6] and [7]. In addition, the examination of realistic delay propagation mechanisms is crucial for the anticipation of potential propagated delays, allowing the realistic assessment of schedule robustness.

In this work, we address this topic by evaluating the prediction accuracy of the theoretical delay propagation model of [2] based on real-world data for four years from a major European carrier. The data set comprises 2,197,406 flight delay records for a hub-and-spoke network from March 2003 to February 2007. Findings can be used in a prototypical scheduling and simulation framework in which scheduling strategies are evaluated for practical application.

The remainder of the paper is organized as follows. In Section 2, we present the formalization of a generalized turnaround process for aircraft and the evaluated theoretical propagation model. Section 3 deals with determining essential minimum ground time values for the aircraft turnaround. Based on this, propagation mechanisms for aircraft rotations are analyzed in Section 4. Conclusions and an outlook are given in Section 5.

2 Fundamental Assumptions on the Turnaround Process and Propagation Effects

The generalized aircraft turnaround process between two consecutive flights within a rotation R is illustrated in Figure 1. Case (1) represents scheduled times in regular operations. Every flight $f \epsilon R$ has a scheduled departure time STD_f and a scheduled arrival time STA_f.

The turnaround process starts with the aircraft arriving at the gate. The scheduled ground time $sgt_{a(f)f}$ between flight f and its aircraft predecessor $a(f)$ consists of the minimum ground time $mgt^A_{a(f)f}$ needed for the turnaround and a potential buffer time $b^A_{a(f)f} \geq 0$.

Cases (2) and (3) depict cases in which flight $a(f)$ arrives late. ATD_f and ATA_f describe the actual times of departure and arrival of flight f, respectively. d^D_f denotes the departure delay $ATD_f - STD_f$ and d^A_f the arrival delay $ATA_f - STA_f$. $agt_{a(f)f}$ is the actual ground time $ATD_f - ATA_{a(f)}$.

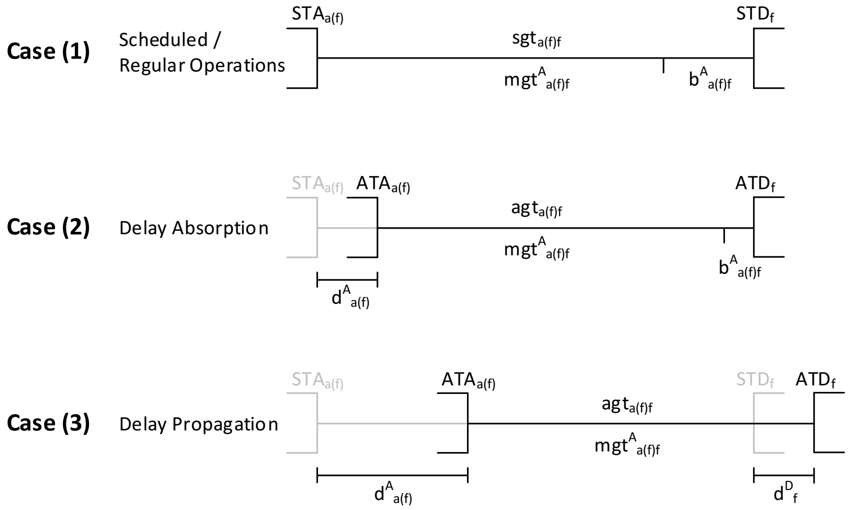


Figure 1. Delay Propagation in during the Aircraft Turnaround Process

In Case (2), the delay can be absorbed by the buffer time and it holds

$$agt_{a(f)f} \geq mgt^{A}_{a(f)f}, \tag{1}$$

$$b_{a(f)f} \geq 0. \tag{2}$$

In contrast, Case (3) implies a delay propagation to flight f with

$$agt_{a(f)f} = mgt^{A}_{a(f)f}, \tag{3}$$

$$b_{a(f)f} = 0. \tag{4}$$

Delay propagation by crew itineraries follow the same mechanisms with the respective minimum ground time $mgt^{C}_{c(f)f}$ between flight f and the crews' predecessor $c(f)$. Based on these assumptions, a basic propagation model for crews and aircraft is formulated in (5)-(9), following [2]:

$$ATA_f = \max\{STA_f, ATD_f + t_f\}, \ \forall f \in F, \tag{5}$$

$$ATD_f = max\left\{STD_f, max\begin{Bmatrix} ATA_{a(f)} \ + mgt^{A}_{a(f)f} \\ ATA_{c(f)} + mgt^{C}_{c(f)f} \end{Bmatrix}\right\} + X_f, \forall f \in F. \tag{6}$$

$$d^D_f = ATD_f - STD_f, \forall f \in F. \tag{7}$$

$$d^A_f = ATA_f - STA_f, \forall f \in F. \tag{8}$$

$$s_f = d_f - X_f, \forall f \in F. \tag{9}$$

X_f represents a stochastic variable for exogenous delays in ground processes that can be modeled based on findings from related studies such as [6] and [7]. A block time t_f is associated to each flight f, determining the time from gate to gate, including taxi-out, flying time and taxi-in.

A flight can depart only if both crew and aircraft are available (6). Equalities (7) and (8) explicitly describe departure and arrival delays, respectively. Note that one main assumption of the model is that negative delays do not propagate, i.e. early arrivals do not imply early departures of subsequent flights. s_f depicts the amount of propagated departure delay for every flight f and therefore is the common target value concerning the evaluation of schedule robustness. Since the duration of the minimum ground time plays a central role in terms of delay propagation, its specification is discussed in further detail in the following.

3 Minimum, Scheduled and Actual Ground Times

Necessarily, minimum ground times *mgt* are determined prior to the construction of aircraft rotation based on generalized rules. It is defined as the shortest time span in which the turnaround process can be performed. Several – partly interconnected – tasks have to be carried out, leading to a fairly complex system of interactive processes. Several studies on a high granular level have been carried out on this topic, see e.g. [8] and [9]. According to [9, p.81], the *mgt* for continental fleets is 45 minutes at the main hub. It can be reduced to 40 minutes when both the inbound and outbound flights are domestic. Since *mgt* values are only available for the main hub, we derive a sufficient estimator for *mgt* values at all other airports of the flight network based on scheduled ground times *sgt*.

At first, possible influential factors that may affect the length of turnaround times are checked in a CART analysis[1]. The results confirm the expectation that fleets are the most substantial influential factor for the *sgt*, followed by O&Ds[2]. As an additional parameter we use a binary variable indicating whether the inbound and outbound flights are both domestic or not. Significant differences by seasonal attributes cannot be observed. We check the following candidates as *mgt* estimators:

- **Minimum Value** as theoretical threshold for data without outliers,
- **.01, .05, .075, .10, and .15 Quantiles** as estimators that are robust against outliers,
- **Modal Value**, based on the assumption that most turnarounds are performed in minimum time available.

[1] See https://cran.r-project.org/package=rpart for details (last visited November 9th, 2017).
[2] The Origin & Destination (O&D) of a flight is determined by its departure and arrival airport.

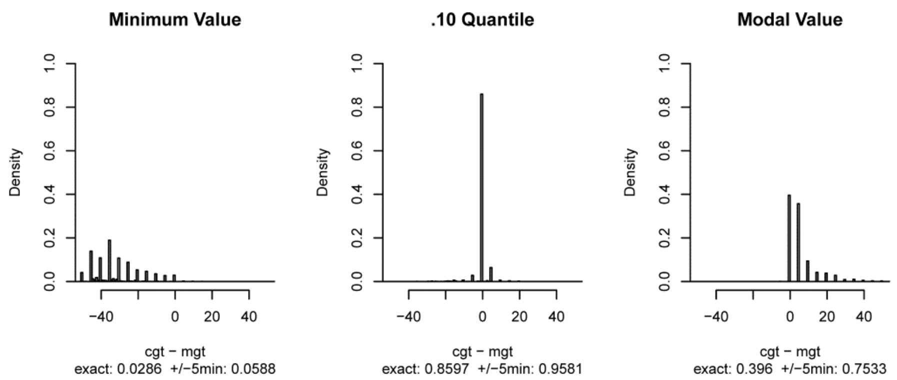


Figure 2. Selected estimators for scheduled *mgt* values

The results for the main hub are presented in Table 1. The second and third column depict the mean and standard deviation of $(cgt - mgt)$ where *cgt* is the minimum ground time computed using the respective estimator and *mgt* is the minimum ground time value given in the data set. The last two columns are the relative amount of flights for which the minimum ground time is derived exactly or within a tolerance range of ±5. It turns out that the .10-quantile suits best for retrieving mgt from sgt, leading to an exact derivation for 86% of all flights. Within the tolerance range of ±5 minutes the value increases to 95.9%.

For further illustration, Figure 2 shows the histogram for the .10-quantile in comparison to the minimum and modal values, indicating the extremes concerning under- and overestimation. In combination with figures of Table 1, we can assume that the .10-quantile is the best estimator of *mgt* values for all airports in the flight network for the scope of this analysis.

Table 1. Prediction accuracy for *mgt* value estimators

			Accuracy	
Estimator	*Mean*	*SD*	*exact*	*±5 minutes*
min	33.705	13.339	0.029	0.059
.01 quantile	20.152	11.129	0.043	0.115
.05 quantile	3.514	6.916	0.587	0.911
.075 quantile	1.734	5.765	0.832	0.951
.10 quantile	1.326	5.132	0.860	0.958
.15 quantile	1.901	4.475	0.690	0.944
mode	6.615	10.820	0.396	0.753
mean	24.436	7.565	0.000	0.001

4 Examination of Propagation Mechanisms on real-world Data

Prior to the assessment of the prediction accuracy of the propagation model, we give an introducing example of a daily rotation based at Hub H_1 that is affected by exogenous delays. Table 2 represents the daily rotation of an Airbus A321 at one representative day in summer. ΔGT indicates the difference between scheduled and actual ground time of flight f and its predecessor $a(f)$:

$$\Delta GT = (STD_f - STA_{a(f)}) - (ATD_f - ATA_{a(f)}). \tag{10}$$

Negative ΔGT values are commonly induced by the usage of buffers for the absorption of incoming delays. In case of a delayed turnaround, the actual ground time exceeds the scheduled ground time, leading to positive ΔGT values. Analogously, ΔBT describes the block time difference

$$\Delta BT = (ATA_f - ATD_f) - (STA_f - STD_f). \tag{11}$$

Concerning O&Ds, H_1 stands for a hub airport while $\{S_1, \ldots, S_4\}$ depict four spoke airports that are served by the out-and-back principle. Column *in* shows the incoming arrival delay of the predecessor flight, columns *rot* and *prim* stand for actual rotation and primary delays. The first flight has a primary departure delay of 69 minutes due to weather conditions at the destination airport (IATA Delay Code 84). Additionally, the scheduled block time is exceeded by 23 minutes, resulting in an arrival delay of 92 minutes. 24 minutes can be absorbed during the following ground time. The turnaround is performed within 46 minutes, close to the minimum ground time of 45 minutes. Nevertheless, the second flight departs 68 minutes late and again experiences an increase of the scheduled block time duration (16 minutes).

The ground time between flight 2 and 3 does not contain any buffer time and takes 7 minutes more than scheduled. The initial delay can finally be absorbed prior to departure of flight 4. The aircraft spends 165 minutes on ground rather than the originally scheduled 255 minutes. The extended ground time prior to flight 5 is a result from the early arrival of flight 4. Eventually, flight 6 is severely delayed, awaiting a crew interchange (IATA Delay Code 95, *flight deck or entire crew*). Overall, a primary delay of 69 minutes in combination with additional 39 minutes of block time delay entails accumulated propagated delays of 159 minutes.

Table 2. Exemplary one-day rotation of an Airbus A321

O&D	*STD*	*ATD*	*STA*	*ATA*	*ΔGT*	*ΔBT*	*in*	*rot*	*prim*
S_1-H_1	06:00	07:09	07:10	08:42	-	23	14	0	69
H_1-S_2	08:20	09:28	09:15	10:39	-24	16	92	68	0
S_2-H_1	09:55	11:26	11:10	12:41	7	0	84	91	0
H_1-S_3	15:25	15:25	16:35	16:25	-91	-9	91	0	0
S_3-H_1	17:25	17:22	18:55	18:42	6	-10	-9	0	0
H_1-S_4	19:55	21:17	20:40	22:01	95	-1	-13	0	0

In the following, we examine the prediction accuracy of the delay propagation and absorption estimation by the presented propagation model. The key question is if operational propagation effects are sufficiently represented or if severe over- or underestimations become apparent. The estimated rotation delay $\hat{d}_f^r$ of flight f is computed as the difference between the arrival delay of the preceding flight (denoted as $inDLY$) and the scheduled buffer time:

$$\hat{d}_f^r = \max(\max(inDLY, 0) - (sgt - mgt), 0). \quad (12)$$

We validate the accuracy of the rotation delay estimation of the model for the 28,502 daily rotations of the A321 fleet in the data set. Each rotation consists of 4.86 daily flights on average. For 58.09% of the underlying 138,662 flights, the predecessor flight in the rotation arrives late so that a rotation delay may emerge. An average of 7.30 minutes of incoming arrival delay leads to 3.30 minutes of outgoing rotation delay which is an absorption rate of 45.24%.

Table 3 shows certain figures concerning the accuracy of rotation delay estimations for flights when the predecessor flight of the aircraft arrives at least one minute late. To compare, Table 4 shows the same values for flights whose aircraft predecessor flight arrives on-time. H1 and H2 stand for the two main hub airports, respectively, while S depicts all spoke airports. Furthermore, we differentiate between domestic (index *d*) and continental flights (index *c*).

Table 3. Estimated rotation delay when the preceding flight of the aircraft arrives late

	all	all_d	all_c	$H1_d$	$H1_c$	$H2_d$	$H2_c$	S_d	S_c
avg	-0.54	-0.49	-0.57	-0.42	-0.20	-0.07	-0.15	-0.78	-0.91
SD	3.51	2.87	3.81	2.98	3.00	2.47	2.27	2.93	4.43
±0 (%)	70.80	74.91	68.64	79.03	76.54	79.21	86.30	68.96	60.47
±5 (%)	90.30	92.67	89.05	92.49	92.45	94.88	97.15	91.71	85.49
$r\Delta_{ta}$ *(%)*	13.33	13.62	13.18	7.99	6.65	8.36	3.81	21.42	19.43
$a\Delta_{ta}$	-5.18	-4.58	-5.51	-4.65	-3.91	-3.69	-3.27	-4.73	-6.00

Table 4. Estimated rotation delay when the preceding flight of the aircraft arrives on-time

	all	all_d	all_c	$H1_d$	$H1_c$	$H2_d$	$H2_c$	S_d	S_c
avg	0.02	0.02	0.02	0.02	0.01	0.06	0.04	0.00	0.02
SD	0.44	0.45	0.43	0.41	0.38	0.76	0.77	0.17	0.41
±0 (%)	99.72	99.68	99.75	99.78	99.81	99.14	99.50	99.91	99.72
±5 (%)	99.87	99.86	99.87	99.85	99.88	99.66	99.77	99.98	99.88
$r\Delta_{ta}$ *(%)*	1.02	1.41	0.78	0.35	0.24	0.61	0.14	2.78	1.39
$a\Delta_{ta}$	-2.58	-2.42	-2.78	-2.07	-2.58	-2.16	-2.00	-2.49	-2.82

In a first step, we discuss the overall prediction accuracy. The first row depicts the average deviation between estimated rotation delay $\hat{d}_f^r$ and actual rotation delay d_f^r which is given in the data set. Throughout all categories, the value is negative, indicating a general overestimation of rotation delays by the model. For continental flights departing at spokes (column S_c), rotation delays are overestimated the most, while values are closest to zero for Hub 2. The rotation delay is estimated correctly on the precise minute for 70.8% of all flights. Overestimation emerges in 19.87%, underestimation in 9.33% of all cases. Best results can be obtained at Hub 2, followed by Hub 1. Obtaining a ±5 minute threshold, results are significantly better but still in the same order for all categories. This is a direct consequence of the fact that departure and arrival times are commonly scheduled in 5 minute intervals. In addition, delay recording at many spoke airports is performed within 5 minute steps.

Table 4 shows respective values for all turnarounds without an incoming arrival delay of the preceding flight. It allows counter-checking if rotation delays occur even when they are theoretically impossible. Slightly positive values in the first row are implied by 160 flights containing rotation delay records without a preceding late aircraft arrival. Apart from these obviously inconsistent records, no systematic falsifying effects become apparent.

Concerning the rotation delay overestimation of the propagation model for turnarounds with incoming arrival delays (Table 3), there are differences between actual operational turnaround durations and generalized *mgt* values. Target *mgt* times do not necessarily determine the operational turnaround duration and may already contain certain slack times as it has already become apparent in Section 3. Referring to this, the two bottom rows of Tables 3 and 4 depict

- $\boldsymbol{r}\Delta_{\boldsymbol{ta}}$ as the relative share of turnarounds that are performed in less than the minimum ground time $(agt < mgt)$ if a late aircraft arrival is likely to postpone the following flight departure, and
- $\boldsymbol{a}\Delta_{\boldsymbol{ta}}$ as the average of the absolute difference $(agt - mgt)$ in minutes.

The actual ground time falls below the minimum ground time at spoke airports significantly more often. Assumingly, this is due to conservatively scheduled *mgt* values at spoke airports since the lack of strong presence makes it harder for an airline to perform recovery actions in case of unscheduled events.

On the one hand, spare resources like reserve crews are mostly available at hubs rather than spokes, on the other hand less possibilities for aircraft swaps exist at spokes. The importance of this effect becomes apparent when comparing the results to corresponding values in Table 4. If an aircraft arrives late for the turnaround, it is thirteen times more likely that the turnaround is performed faster than scheduled.

Going into more detail, we check this mechanism by specifically considering data for the main Hub 1. Depending on the exact arrival delay of the aircraft predecessor flight, we examine if length and frequency of turnaround speed-ups change accordingly. Finally, we check the overall operational importance in terms of the total amount of delay savings.

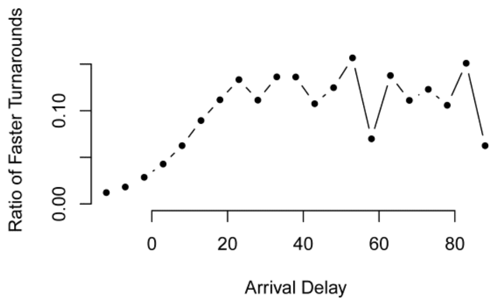


Figure 3. Ratio of turnarounds that are performed in less than minimum ground time

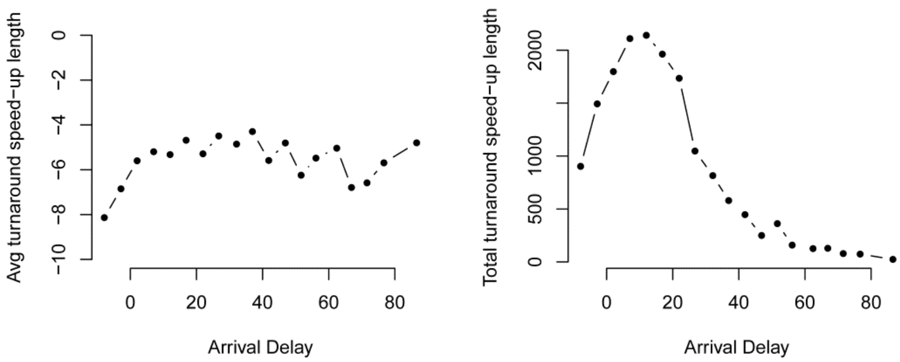


Figure 4. Operational emergence of turnarounds in less than the minimum ground time

For a start, the left panel of Figure 3 shows the ratio of turnarounds with $(agt < mgt)$ depending on the arrival delay of the aircraft predecessor flight. We use bins of 5 minutes for arrival delays up to 60 minutes. There is a substantial and steady increase of faster turnarounds until an arrival delay of 25 minutes. For larger arrival delays, inconclusive fluctuations at high levels can be observed.

In addition, the left panel of Figure 4 shows the average absolute difference $(agt - mgt)$ in minutes depending on the arrival delay. Values are in a constant interval, however, with considerable fluctuations. No relevant pattern can be observed. Finally, the right panel of Figure 4 is intended to indicate operational relevance of turnaround speed-ups. The y-axis shows the sum of absorbed arrival delay by turnaround speed-ups in minutes. It thus depends on both absolute turnaround speed-up and the occurrence frequency of certain arrival delay values. Most substantial savings are achieved for arrival delays between 0 and 25 minutes.

The dominant factor for the curve progression is that significantly large arrival delays are extremely rare and although turnarounds are performed faster in these cases, too, the operational relevance is minor in terms of a holistic assessment of rotation delays.

5 Conclusion

In the presented analysis, the assessment of the delay propagation model of [2] has been provided as a first step concerning the refinement of delay propagation mechanisms for a prototypical robust aircraft scheduling and simulation framework.

It has turned out that 70.8% of rotational delay propagation can be precisely estimated. Within a ±5 minute tolerance threshold, the estimation is correct for about 90.3% of all flights. The prediction provides best results for large hub airports with values up to 97.15% while at continental spokes only 85.49% can be reached. On average, rotation delays are overestimated by 0.54 minutes per flight. This value ranges from nearly 0 at Hub 2 to around 0.91 minutes at continental spoke airports.

A substantial responsibility for the overestimation of rotation delays lies in the fact that actual ground times sometimes fall below scheduled target minimum ground. It happens in 13.33% of all turnarounds when the aircraft arrives late. On average, 5.18 minutes of propagated delay can be saved in these cases. Deeper insight in the dependency between arrival delays and turnaround speed-ups shows that operational relevance is given for arrival delays between 1 and 25 minutes. In this range, delays are often overestimated by the propagation model in comparison to actual operations.

Future work has to deal with the incorporation of specific correction terms into the propagation model. Subsequently, the scheduling framework can be used to examine in how far refined delay propagation actually leads to an improved assessment of schedule robustness. Furthermore, it has to be evaluated if findings are comparable when crew-related propagation effects are considered additionally. Final results form the basis for the synchronization of assumptions on delay propagation in scheduling with the operational reality in practical robust resource scheduling.

References

1. Bazargan, M. (2016). Airline operations and scheduling. Routledge.
2. Dück, V., L. Ionescu, N. Kliewer, and L. Suhl (2012). Increasing stability of crew and aircraft schedules. Transportation research part C: emerging technologies 20 (1), 47–61.
3. Yen, J. W. and J. R. Birge (2006). A stochastic programming approach to the airline crew scheduling problem. Transportation Science 40 (1), 3–14.
4. Dunbar, M., G. Froyland, and C.-L. Wu (2012). Robust airline schedule planning: Minimizing propagated delay in an integrated routing and crewing framework. Transportation Science 46 (2), 204–216.
5. Ionescu, L. and N. Kliewer (2011). Increasing flexibility of airline crew schedules. Procedia-Social and Behavioral Sciences 20, 1019–1028.
6. Tam, M. B. B. (2011). Optimisation approaches for robust airline crew scheduling. Ph. D. thesis, University of Auckland.
7. Ionescu, L., Gwiggner, C., & Kliewer, N. (2016). Data analysis of delays in airline networks. Business & Information Systems Engineering, 58(2), 119-133.
8. Fricke, H., & Schultz, M. (2009, June). Delay impacts onto turnaround performance. In ATM Seminar.
9. Schlegel, A. (2010). Bodenabfertigungsprozesse im Luftverkehr: Eine statistische Analyse am Beispiel der Deutschen Lufthansa AG am Flughafen Frankfurt/Main. Springer-Verlag.

Decision Support Rules for Flexible Time Window Management of Attended Home Deliveries

Charlotte Köhler[1], Ann Melissa Campbell[2]

[1] Management Science Group, Otto von Guericke University, Magdeburg, Germany
charlotte.koehler@ovgu.de
[2] Department of Management Sciences, University of Iowa, Iowa City, United States
ann-campbell@iowa.edu

Abstract. In the competitive world of online retail, customers choose from a selection of delivery time windows on retailers' websites. Creating a set of convenient and cost-efficient delivery time windows is challenging, since customers want short time windows, but short time windows can increase the delivery costs of a retailer significantly. Furthermore, the acceptance of a particular request can restrict the ability of accommodating future requests significantly. In this paper, we present decision support rules that enable flexible time window management in the booking of time window based deliveries. We build tentative delivery routes and check if a new customer request can be accommodated feasibly with the remaining delivery capacity. We maintain routing flexibility through offering short or long time windows based on customer characteristics as well as characteristics of the evolving route plan. We investigate the presented approaches with a case study in the area of online supermarkets.

Keywords: Time Window Management, Customer Acceptance, Attended Home Deliveries, Vehicle Routing with Time Windows, Online Supermarkets

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Ein Entscheidungsunterstützungssystem zur Tourenplanung am Beispiel eines innovativen Lebensmittel-Lieferkonzeptes

Sebastian Adam Gebski, Patrick Czerwinski, Max Leyerer, Marc-Oliver Sonneberg und Michael H. Breitner

Leibniz Universität Hannover, Institut für Wirtschaftsinformatik, Hannover, Deutschland
{gebski,czerwinski,leyerer,sonneberg,breitner}@iwi.uni-hannover.de

Abstract. Die weltweite Urbanisierung führt zu einer erhöhten Nachfrage von innerstädtischen Gütertransporten, wodurch das Verkehrsnetz belastet wird. Moderne Technologie sowie die Nutzung verfügbarer GIS Daten bieten den Anreiz zur Entstehung alternativer und innovativer Lieferkonzepte. Davon betroffen sind ebenfalls urbane Lebensmittel-Lieferdienste, die durch Verwendung verfügbarer mobiler Daten Lieferungen effizienter gestalten können. Dieser Beitrag beschäftigt sich mit einem innovativen Lebensmittel-Lieferkonzept. Hierzu wird ein gemischt ganzzahliges Optimierungsmodell auf Basis eines Vehicle Routing Problems (VRP) mit dem Ziel der dynamischen Zuordnung und Reihenfolgeplanung formuliert. Charakteristisch für die vorliegende Thematik ist die Existenz mehrerer Fahrer in Kombination mit offenen Touren. Die Zuordnung und Planung der Touren wird durch die Implementierung des Optimierungsmodells in ein Entscheidungsunterstützungssystem (EUS) unter Verwendung von Echtzeit Geoinformationsdaten (GIS-Daten) automatisiert und optimiert.

Keywords: City Logistik, Routenplanung, Optimierung, Entscheidungsunterstützungssystem.

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

1 Einleitung und Motivation

Der Begriff Urbanisierung beschreibt ein bilaterales Wachstum von Städten hinsichtlich Größe und Einwohnerzahl [1]. Dieses Wachstum geht dabei mit einem erhöhten Aufkommen an logistischen Aktivitäten innerhalb urbaner Räume einher, was auf eine erhöhte Nachfrage nach Produkten und Dienstleistungen zurückzuführen ist. Die Konsequenzen sind dabei erhöhtes Verkehrsaufkommen sowie erhöhte Schadstoffemissionen. Verstärkt wird dieser Effekt durch zunehmende Anforderungen von Kunden an den Servicegrad der Logistikdienstleister, was an zusätzlichen Dienstleistungen, wie z. B. der „same day delivery" zu erkennen ist. Das resultierende erhöhte Aufkommen an logistischen Aktivitäten innerhalb von (Groß-) Städten ist gleichzeitigt ein Treiber für die Entwicklung neuer, innovativer und disruptiver Geschäftsmodelle und Lieferkonzepte. Hierbei werden im Zuge der digitalen Transformation mobile Endgeräte, die sich besonders durch ihre Ubiquität sowie diverse Schnittstellen zu mobilen Datendiensten kennzeichnen, genutzt, um unter anderem Lieferprozesse zu optimieren. Dieser Trend ist beispielsweise an der steigenden Anzahl von Applikationen, wie etwa Uber und Lyft zu erkennen, die sogenannte on-demand Transportdienste anbieten [2]. Neben Personentransporten können Lebensmittel-Lieferdienste ebenfalls durch die Verwendung innovativer mobiler Technologien optimiert werden. Diese Möglichkeiten bestehen sowohl für die Auslieferung von Einkäufen und gesamten Warenkörben von Supermärkten als auch für Lieferungen von Mahlzeiten, auf welche sich dieser Ansatz fokussiert. Lieferdienste werden daher als Unternehmen verstanden, die auf Bestellungen von Kunden zubereitete Mahlzeiten an die Kunden ausliefern. Da Lieferdienste zur Auslieferung ihrer Produkte das städtische Verkehrsnetz nutzen, können diese den Akteuren der urbanen Logistik zugeordnet werden [3]. Aus Sicht der Lieferdienste zeichnet sich die Optimalität einer Route durch mehrere Aspekte aus. Hierzu zählt, dass innerhalb der Tour möglichst viele Kunden beliefert und dabei möglichst kurze Distanzen zurückgelegt werden. Dadurch werden verfügbare Ressourcen (z. B. Arbeitszeit der Fahrer, Kraftstoff der Fahrzeuge) effizienter genutzt und die Kundenzufriedenheit durch kurze Lieferzeiten gewährleistet.
Beim klassischen Konzept eines Lieferdienstes kehren die Fahrer nach jeder erfolgreichen Auslieferung wieder zur Filiale zurück. Das innovative Lieferkonzept besteht darin, den Servicegrad durch verkürzte Lieferzeiten zu erhöhen, da die auszuliefernden Produkte bereits vorgefertigt und im Lieferfahrzeug verladen sind. Diese Fahrzeuge sind im Stadtgebiet verteilt und bedienen Bestellungen von ihrem jeweils aktuellen Standpunkt. Somit ermöglicht dieses Konzept on-demand Lieferungen mit minimalen Wartezeiten für die Kunden. Die Forschungsfrage des Beitrags lautet daher: *Wie können on-demand Lieferdienste von vorab zubereiteten Mahlzeiten bei der operativen Tourenplanung effizient unterstützt werden?*

Im folgenden Kapitel wird das vorliegende Geschäftsmodell präzisiert und in die Literatur eingeordnet. Das entwickelte Optimierungsmodell ist Gegenstand von Kapitel 3 und dient als Grundlage für das Entscheidungsunterstützungssystem (EUS) in Kapitel 4, bevor anschließend ein Anwendungsbeispiel vorgestellt wird. Kapitel 6 umfasst Diskussionen und Limitationen des Ansatzes, ein Fazit rundet den Beitrag ab.

2 Geschäftsmodell und Einordnung in die Literatur

Im klassischen Lieferdienst-Geschäftsmodell wird der Großteil der Mahlzeiten nach Bestellungseingang individuell zubereitet. Anschließend wird die Bestellung, bestehend aus einer oder mehreren zubereiteten Mahlzeiten, einem Fahrer zugewiesen, der diese von der Filiale zum Kunden transportiert. Teilweise werden dem Fahrer auch Bestellungen von verschiedenen Kunden zur Auslieferung gleichzeitig ausgehändigt. Nach erfolgreicher Auslieferung der Bestellungen kehrt der Fahrer zur Filiale bzw. Zentrale zurück, sodass die Lieferungen in sogenannten Rundreisen bzw. Touren mit einem oder eben mehreren Kunden stattfinden. Bei diesem klassischen Lieferdienst-Geschäftsmodell kann es so zu ineffizienten Routen der Fahrer kommen, falls nur wenige Kunden in einer Tour beliefert werden. Dies ist meist durch zeitverzögerte Bestellungen, der frischen Zubereitung der Mahlzeiten nach Bestellungseingang (first-in first-out-Prinzip) sowie große Distanzen zwischen Kundenorten begründet. Diese Nachteile klassischer Lieferdienste werden durch ein alternatives Lieferkonzept umgangen, welches ein Berliner Start-Up (*GreenGurus)* bereits anwendet. Dabei werden sämtliche Produkte morgens für den gesamten Arbeitstag vorgefertigt, wobei das Produktangebot auf eine geringere Anzahl begrenzt ist. Die vorgefertigten Produkte, bei denen es sich ausschließlich um haltbare Mahlzeiten (z. B. Salate) handelt, werden daraufhin in Lastenfahrräder und Elektro-Roller vorgeladen, die mit Kühlboxen ausgestattet sind und anschließend in einem zuvor definierten Liefergebiet verteilt. Dadurch gelingt es, die Distanzen zu potentiellen Kunden zu minimieren und Rückfahrten zur Zentrale zu vermeiden. Die individuellen Positionen und Produktbestände aller Lieferfahrzeuge sind dabei einer Zentrale durch Verwendung von GPS Daten und einer mobilen App jederzeit bekannt. Erfolgt eine Bestellung durch einen Kunden, so wird dem Kunden durch einen Mitarbeiter in der Zentrale derjenige Fahrer zugeordnet, der die bestellten Mengen an Produkten vorrätig hat und sich am nächsten zu dem entsprechenden Kunden befindet. Dem Fahrer wird bei erfolgreicher Zuordnung per mobiler App eine Benachrichtigung über die Menge der auszuliefernden Produkte, die entsprechenden Kundendaten sowie die zu fahrende Route gegeben.

Im Vergleich zu klassischen Lieferdiensten bestehen demnach mehrere Unterschiede. Zunächst beginnen die Fahrer den Lieferungsprozess von keinem zentralen Standpunkt, sondern von ihren jeweils aktuellen Positionen aus, da sie die nachgefragten Produkte bereits geladen haben. Zusätzlich beenden die Fahrer ihre Touren, beim jeweils zuletzt belieferten Kunden, da keine Rückkehr zu einem zentralen Standpunkt notwendig ist und demnach die nächste Lieferung von der aktuellen Position beginnen kann. Die Verkehrsinfrastruktur wird somit durch zwei Faktoren entlastet: Es können Lastenfahrräder und Elektroroller anstatt Autos eingesetzt werden und es sind keine Rückfahrten zur Zentrale erforderlich. Besonders Lastenräder ermöglichen eine Entlastung des Straßenverkehrsnetzes, da diese zusätzlich Radwege der Stadt nutzen können und, ebenso wie Roller, weniger Parkraum beanspruchen.

Da bei einer zunehmenden Anzahl an Kunden und Fahrern eine manuelle Zuweisung dieser zueinander nicht mehr optimal möglich ist, erscheint es sinnhaft, diesen Prozessschritt zu optimieren und zu automatisieren. Dabei kann das zugrunde liegende Optimierungsproblem zunächst formalmathematisch modelliert und anschließend mittels EUS softwaregestützt gelöst werden. Während die Distributionsprozesse im klassischen Lieferdienst mit dem bekannten Vehicle Routing Problem (VRP) modelliert werden können, ist dies beim beschriebenen Lieferkonzept nicht eindeutig. Einerseits existieren aufgrund mehrerer, individueller Startpunkte mehrere Depots, die den Startpunkten jeder Tour entsprechen, was der Ausgangssituation im sogenannten Multiple Depot Vehicle Routing Problem (MDVRP) entspricht. Andererseits werden die Depots, anders als im MDVRP angenommen, nach Beendigung der letzten Lieferung nicht mehr angefahren, sodass die Touren bei den zuletzt belieferten Kunden enden, was wiederum der Ausgangssituation im sogenannten Open Vehicle Routing Problem (OVRP) entspricht. Aus diesem Grund kann die Tourenplanungssituation im innovativen Lieferkonzept mit einer Kombination aus OVRP und MDVRP modelliert werden. Für beide Probleme existieren in der Literatur mehrere Modellieransätze, auf deren Basis ein entsprechend angepasstes Optimierungsmodell formuliert werden kann. Surekha und Sumathi [4] entwickeln ein gemischt-ganzzahliges Optimierungsmodell (engl. mixed-integer programming (MIP)) für das MDVRP und entwickeln zudem ein Lösungsverfahren mittels genetischer Algorithmen. Hierbei verwenden die Autoren unterschiedliche Teilmengen zur Abbildung der Depots (Startpunkte der Fahrer) und der Kundenorte. Crevier et al. [5] erweitern das MDVRP um inter-Depot-Routen, wobei zusätzlich ein heuristisches Lösungsverfahren präsentiert wird. MIP-Formulierungen und Lösungsverfahren für das OVRP werden von Brandao [6] und Aksen et al. [7] präsentiert. Letchford et al. [8] verweisen zudem darauf, dass es zur Bildung offener Routen bei Lösung des Optimierungsmodells genügt, die Rückwege zum Depot bzw. Startpunkt der Routen mit „0“ zu gewichten. Da das zu formulierende Optimierungsmodell in einem EUS implementiert wird, kann diese Eigenschaft berücksichtigt werden. Das Problem kann laut der Klassifizierung nach Pillac et al. [9] als dynamisches VRP mit deterministischem Input eingeordnet werden.

Das in Kapitel 3 vorgestellte Optimierungsmodell wurde auf Basis des *Design Science Research* Forschungsansatzes nach Hevner et al. [10] entwickelt. Die Forschungsmethodik basiert darauf, das Forschungsobjekt an einem Relevanz- und Stringenz-Zyklus auszurichten. Durch den Relevanz-Zyklus wird die Aktualität des Forschungsprojekts durch existierende Probleme und Herausforderungen sichergestellt. Der Stringenz-Zyklus sorgt für eine Analyse der bestehenden Wissensbasis auf dem jeweiligen Betrachtungsgebiet und soll die Innovation des Forschungsprojektes gewähren. Beide Zyklen wurden in diesem und im voran gestellten Kapitel adressiert. Schließlich wird das Forschungsobjekt im sogenannten Design-Zyklus entwickelt, bewertet und iterativ erweitert um eine kontinuierliche Verbesserung zu erreichen [10]. Als finales Artefakt dieses Prozesses resultiert ein EUS mit eingebettetem Optimierungsmodell, welches Anbietern von vorab zubereiteten Mahlzeiten in der Tourenplanung und kurzfristigen Tourenoptimierung in der täglichen Auslieferung unterstützt. Eingehende Kundenbestellungen werden bestehenden Touren in Echtzeit zugeordnet und diese optimal angepasst. Mit diesem Artefakt zu dem beschriebenen emissionsarmen Lieferkonzept adressieren wir die Aufforderung von Gholami et al. [11], welche den überdurchschnittlichen Anteil von Konzeptualisierungen und theoretischen Analysen kritisieren und damit mehr anwendungsorientierte Forschung auf dem Gebiet der Wirtschaftsinformatik fordern.

3 Mathematisches Modell

Im Folgenden wird das formulierte gemischt-ganzzahlige Optimierungsmodell zur Tourenplanung sowie die zugrundliegende Notation vorgestellt und erläutert. Das Modell unterliegt dabei folgenden Annahmen:

- Zielsetzung: Minimierung der gesamten Lieferzeit
- Die Lieferzeit setzt sich aus der Reisezeit zwischen zwei Knoten und einer Servicezeit beim Kunden zusammen.
- Die Servicezeit pro Kunde beträgt 5 Minuten.
- Es werden verschiedene Arten von haltbaren und vorgefertigten Produkten ausgeliefert, welche in homogene Lieferfahrzeuge geladen werden.
- Alle Kunden müssen bedient werden.
- Teillieferungen und Nachladeprozesse sind nicht erlaubt. Ein zugeordneter Fahrer muss alle bestellten Produkte transportieren.

Tabelle 1. Notation

Mengen	
H	$\{h = 1, \dots, K\}$ *Fahrer*
I^A	$\{i, j = 1, \dots, K\}$ *Fahrerstandorte*
I^B	$\{i, j = K+1, \dots, L\}$ *Kundenorte*
I	$\{I^A\} \cup \{I^B\}$ *Menge aller Orte*
P	$\{p = 1, \dots, N\}$ *Produkte*
Parameter	
a_{ip}	*Angebot von Produkt* p *am Fahrerstandort* $i \in I^A$
n_{ip}	*Nachfrage nach Produkt* p *am Kundenort* $i \in I^B$
sz	*Servicezeit beim Kunden*
tt_{ij}	*Fahrzeit zwischen den Orten* $i \in I^A$ *und* $j \in I^B$ *bzw.* i *und* $j \in I^B$
Variablen	
x_{ijh}	$\in \{0,1\}$ *Reihenfolge in der Fahrer* h *Orte* i *und* j *besucht*
y_{ih}	$\in \{0,1\}$ *Zuordnung von Fahrer* h *zu Ort* i
z_i	$\in \mathbb{R}$ *Hilfsvariable*

$$Min\ Z = \sum_{i \in I} \sum_{j \in I} \sum_{h \in H} tt_{ij} \cdot x_{ijh} + \sum_{i \in I} \sum_{h \in H} y_{ih} \cdot sz \qquad (1)$$

u. B. d. R.

$$\sum_{i \in I^B} n_{ip} \cdot y_{ih} \leq \sum_{i \in I^A} a_{ip} \cdot y_{ih} \quad \forall\, h, p \qquad (2)$$

$$\sum_{j \in I} x_{ijh} = y_{ih} \quad \forall\, i \in I, h \qquad (3)$$

$$\sum_{i \in I} x_{ijh} = y_{jh} \quad \forall j \in I, h \tag{4}$$

$$\sum_{h \in H} y_{ih} = 1 \quad \forall i \in I^B \tag{5}$$

$$\sum_{h \in H} y_{ih} \leq 1 \quad \forall i \in I^A \tag{6}$$

$$y_{ih} = 0 \quad \forall i \in I^A, h; h \neq i \tag{7}$$

$$x_{iih} = 0 \quad \forall i \in I, h \tag{8}$$

$$z_i - z_j + L \cdot \sum_{h \in H} x_{ijh} \leq L - 1 \quad \forall i \in I^B; j \in I^A; i \neq j \tag{9}$$

$$x_{ijh} \in \{0,1\} \quad \forall i, j \in I, h \tag{10}$$

$$y_{ih} \in \{0,1\} \quad \forall i \in I, h \tag{11}$$

In (1) wird die gesamte Lieferzeit aller Touren, bestehend aus Reise- sowie und Servicezeiten minimiert. Gleichung (2) ist eine Kapazitätsrestriktion für alle Produkte, die ausgeliefert werden, wobei das Optimierungsmodell um beliebig viele Produkte erweitert werden kann. Jeder Knoten der angefahren wird, muss auch wieder verlassen werden (3). Zudem muss jeder Knoten von einem anderen Knoten aus erreicht werden (4). Durch Gleichung (5) wird sichergestellt, dass jeder Kunde genau einem Fahrer zugeordnet wird. Gleichung (6) besagt, dass die Standorte der Fahrer maximal einer Tour zugeordnet werden können. Gleichung (7) stellt sicher, dass die Fahrerstandorte $i \in I^A$ nur den entsprechenden Fahrern $h \in H$ zugeordnet werden, wobei für die Fahrerstandorte und die Fahrer dieselbe Notation verwendet wird. Gleichung (8) stellt sicher, dass ein Fahrer einen Ort innerhalb einer Tour nicht doppelt anfahren kann. Gleichung (9) verhindert die Bildung von Kurzzyklen, die den jeweiligen Fahrerstandort nicht beinhalten. (10) und (11) geben die Definitionsbereiche der Entscheidungsvariablen an.

4 Entscheidungsunterstützungssystem

Das EUS besteht in Form einer Web-Applikation (App), die mittels Ruby on Rails entwickelt wurde. Die App dient dabei erstens als Benutzeroberfläche zur Eingabe und Darstellung von Daten und zweitens zur automatisierten Steuerung der Hintergrundaktivitäten zwischen den Schnittstellen. Das EUS interagiert dabei mit einer SQLite Datenbank (DB), dem im General Algebraic Modeling System (GAMS) implementierten Optimierungsmodell, dem IBM ILOG CPLEX Solver und der Google Maps API (engl., application programming interface).

Die Hintergrundaktivitäten umfassen:

1. die Berechnung von Distanzen zwischen den Standorten mithilfe der Google Maps API,
2. die Datenübertragung der eingegebenen Parameterwerte in die DB,
3. die Erzeugung einer Input-Datei, welche die Daten der DB enthält,
4. den Start des Optimierungsvorgangs in GAMS,
5. die Lösung der Probleminstanz mithilfe des CPLEX-Solvers und
6. die Erzeugung einer Lösungsdatei.

Mithilfe der App können Kunden Bestellungen aufgeben, Fahrer ihre Route einsehen und abgeschlossene Lieferungen dokumentieren sowie Administratoren Kunden, Bestellungen, Fahrer und alle dazugehörigen Daten verwalten. Dazu gehört neben der Einsicht der Daten auch die Darstellung von Routen sowie Positionen der Fahrer und Kunden in Google Maps Karten. Zum Start der Hintergrundaktivitäten, und somit zur Lösung des Optimierungsmodells, führen mehrere Ereignisse. Diese sind der Bestellungseingang, Änderungen von Fahrern, sowie abgeschlossene Lieferungen.

Um die Funktionsweise konkretisiert darzustellen wird das Ereignis eines Bestellungseingangs betrachtet. Der Kunde gibt seine Bestellung in einer Maske ein, in der dieser seinen Namen, seine Adresse und die gewünschte Anzahl an Produkten angibt. Noch vor Beginn der Hintergrundaktivitäten wird anhand der DB geprüft, ob es verfügbare Fahrer mit ausreichendem Produktbestand gibt, welche die Bestellung erfüllen können. Ist dies der Fall, werden die Bestellungsdaten in die DB übertragen und ein neuer Kunde angelegt. Mithilfe der Google Maps API werden anhand der Adressen der Kunden und Fahrer die Fahrtzeiten zwischen allen Kunden sowie zwischen Kunden und Fahrern berechnet. Diese Werte werden in die DB übertragen. Somit sind alle, für das Lösen des Optimierungsmodells benötigten Daten vorhanden, welche darauf in eine entsprechend für GAMS lesbare Input-Datei geschrieben werden. In diese werden die Daten für die Mengen (Sets) i,j und h sowie die Parameter n_{ip}, a_{ip} und tt_{ij} eingetragen. Außerdem werden die ersten beiden Kunden einer jeden Tour eines Fahrers fixiert, um den Workflow des Fahrers zu berücksichtigen. Die Daten werden dabei automatisch generiert und anschließend in die Input-Datei geschrieben. Nach der Erzeugung der Input-Datei wird die Lösung des in GAMS implementierten Optimierungsmodells gestartet, wobei die Input-Datei eingelesen und anschließend mithilfe von CPLEX eine optimale Lösung generiert wird. Die Lösung wird in eine Textdatei geschrieben, welche von der App automatisch gelesen wird. Die Ergebnisse werden dabei in die DB übertragen, welche der Zuordnung der Bestellung zu einem Fahrer sowie die angepassten Routen und Fahrtzeiten der Fahrer entsprechen.

Im Falle der anderen Ereignisse zum Start der Hintergrundaktivitäten unterscheiden sich diese lediglich in der Anpassung der Daten in der DB. Bei einer Bestellung werden Kunden-Daten hinzugefügt. Beim Hinzufügen und Löschen von Fahrern werden ggf. Routen gelöscht, welche neu geplant werden müssen. Bei der Fertigstellung von Lieferungen werden die Standorte aktualisiert, was jedoch bei jedem der Ereignisse geschieht. Abbildung 1 fasst die beschriebene Methodik zusammen.

1. Auslöser-Ereignisse und Datenerzeugung
    a. Fall 1: Bestellungseingang
        i. Schreiben der Bestellungsdaten in die DB
    Fall 2: Löschen oder Hinzufügen von Fahrern
        i. Aktualisierung der Fahrerdaten (Orte und Routen) in der DB
    Fall 3: Abschließen einer Lieferung
        i. Löschung der abgeschlossenen Bestellung in der DB
        ii. Aktualisierung von Fahrerdaten (Ort und Route) in der DB
    b. Prüfung der Verfügbarkeit an bestellten Produkten
    c. Berechnung der neuen Fahrtzeiten zwischen allen Orten durch die Google Maps API
2. Aktualisierung der Input-Daten für das Optimierungsmodell in der Input-Datei Festlegung der Teilmengen I^A und I^B durch Prüfung der Anzahl von Fahrern
    a. Setzen der Fahrer und Kunden mit individueller Kennzeichnung in die Sets i,j und h
    b. Setzen der Nachfrage- und Angebots-Parameter n_{ip} und a_{ip}
    c. Setzen der Fahrtzeiten tt_{ij} zwischen allen Orten i und j
    d. Fixierung vorhandener Routen (ersten beiden Kunden jeder Tour)
3. Mathematisches Modell
    a. Erzeugung der Gleichungen
    b. Lesen der Input-Daten aus der Input-Datei
        i. Daten der Paramater n_{ip} und a_{ip}
        ii. Daten der Fahrtzeiten tt_{ij}
        iii. Setzen der pauschalen Servicezeit sz auf 5 Minuten
    c. Lösung des Modells und Berechnung der Variablen mit CPLEX
4. Transfer der Lösungsdaten
    a. Erzeugung einer Lösungsdatei mit den Ergebnissen der Optimierung
    b. Übertragung der Ergebnisse in die DB
        i. Aktualisierung der zugeordneten Fahrer zu Kunden (y_{ih})
        ii. Aktualisierung der Routen der Fahrer (x_{ijh})
        iii. Aktualisierung der Gesamtfahrtzeiten der Fahrer

Abbildung 1. Zusammenfassung der Methodik

Die vorliegende Problematik ähnelt der Situation von sogenannten Online VRPs. Hierbei gehen Kundenbestellungen inkrementell in Echtzeit ein. Diese werden ebenfalls in Echtzeit Fahrern zugeordnet, deren bereits bestehende Touren ggf. um die neuen Kunden erweitert werden [12]. Jaillet und Wagner [13] bieten zusätzlich einen Überblick über derartige Optimierungsprobleme und algorithmische Lösungsmethoden. Im Rahmen des EUS wird zur Lösung des Optimierungsmodells mit CPLEX ein exaktes Verfahren genutzt, wobei die Rechenzeit aufgrund der beschriebenen Datenstruktur gering gehalten und die optimale Lösung gewährleistet wird. Durch das betrachtete Geschäftsmodell wird so die Gesamtfahrtzeit aufgrund der vermiedenen Rundreisen gegenüber herkömmlichen Geschäftsmodellen reduziert.

5 Anwendungsfall

Die Funktionsweise des EUS wird folgend anhand eines Anwendungsbeispiels vorgestellt. Wie beim Lieferdienst *GreenGurus* wird in diesem Anwendungsfall auch von einem begrenzten Liefergebiet (z. B. nach Postleitzahlen) ausgegangen, wodurch sowohl die Lieferzeiten, als auch die zurückzulegenden Strecken der Fahrer innerhalb des Gebiets begrenzt werden. Die Fahrer werden dabei auf Basis von Erfahrungswerten im Liefergebiet verteilt. Im Anwendungsfall wird von einer Situation im laufenden Betrieb mit fünf Fahrern, zwei Produkten und zwölf zu beliefernden Kunden ausgegangen. Die bestehenden Kunden sind dabei bereits zu den Touren der Fahrer zugeordnet (siehe Tabelle 2). Im Beispiel geht nun eine Bestellung des neuen Kunden K13 ein, welcher über die Eingabemaske in der App den Auftrag abgesandt hat. Mit der erfolgreichen Bestellung erfolgt die Registrierung des Kunden. Dabei werden die vorgestellten Hintergrundaktivitäten bis hin zur Lösung des Optimierungsmodells mit neuen Daten ausgelöst. Die durch das Modell berechneten Variablenwerte werden in die DB übertragen und in der App visualisiert. Das Optimierungsergebnis ist in Tabelle 2 zu sehen und zeigt, dass der Kunde K13 dem Fahrer F2 zugeordnet wurde. Die Route ist in dieser Darstellung von oben nach unten zu lesen.

Tabelle 2. Anwendungsfall

	Startsituation		**+ K13**		**- K1**		**+ K14**	
Fahrer	P1/P2	Route	P1/P2	Route	P1/P2	Route	P1/P2	Route
F1	15/6	K7 K9	**15/6**	K7 K9	15/6	K7 K9	15/6	K7 K9
F2	26/22	K1 K8	26/22	K1 K8 **K13**	**24/19**	**K8** **K13**	24/19	K8 K13
F3	30/20	K4 K5 K11	30/20	K4 K5 K11	30/20	K4 K5 K11	30/20	K4 K5 K11
F4	15/15	K3 K6 K12	15/15	K3 K6 K12	15/15	K3 K6 K12	15/15	K3 K6 K12
F5	25/25	K2 K10	25/25	K2 K10	25/25	K2 K10	25/25	K2 K10 **K14**

Legende: F – Fahrer; K – Kunde; P – Verfügbare Produktmengen

Im weiteren Verlauf des Szenarios stellt der Fahrer F2 die Belieferung des Kunden K1 fertig und gibt dies in der App an. Damit wird die Route des Fahrers (K1, K8, K13) um K1 verkürzt. Ebenfalls zu sehen ist, dass sich die Anzahl der verfügbaren Produkte um die ausgelieferte Menge verringert hat (in diesem Fall zwei Einheiten des Produktes P1 und drei Einheiten des Produktes P2). Das Optimierungsmodell wurde in diesem Schritt erneut gelöst, um mögliche neue optimale Routen auf Basis der neuen Daten zu finden, was in diesem Fall jedoch zu keinen Änderungen führt. Anschließend trifft durch K14 eine weitere Bestellung ein, die nach bereits erläutertem Prinzip dem Fahrer F5 zugeordnet wurde.

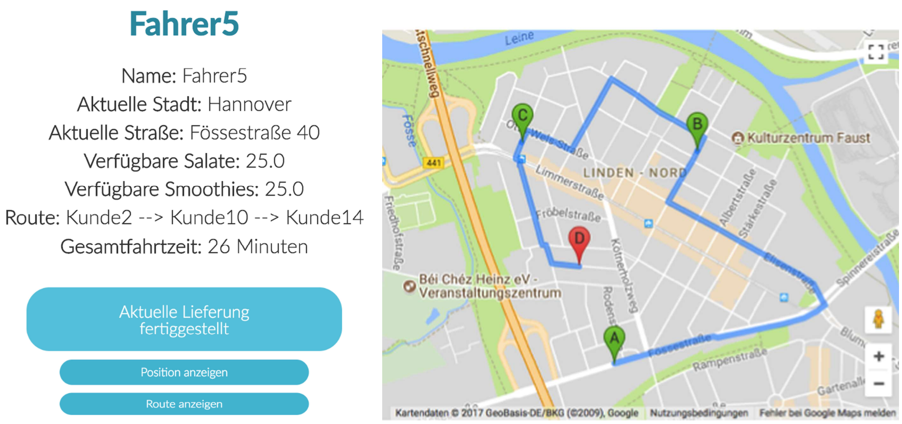


Abbildung 2. Daten eines Fahrers nach Ereignissen im Szenario

Abbildung 2 zeigt beispielhaft, wie die Daten des Fahrers F5 nach den Ereignissen im Szenario in der App dargestellt werden. Dabei sind der Fahrername, dessen Position, die Art (in diesem Fall Salate und Smoothies) und Anzahl der Produkte, sowie die Gesamtfahrtzeit und die in Google Maps dargestellte Route des Fahrers zu sehen. Das Lösen des Optimierungsmodells in der vorgestellten Probleminstanz erfolgt auf einem Intel Core i5 Prozessor mit 1,6 GHz und 8 GB DDR3 unter macOS 10.12.4 in durchschnittlich 0,242 Sekunden je Bestellung. Im Vergleich dazu beträgt die Rechenzeit bei einer größeren Instanz von 30 Fahrern, 80 Kunden (davon 50 fixiert) und 3 aufeinander eingehenden Bestellungen durchschnittlich 17,47 Sekunden je Bestellung.

6 Diskussion und Limitationen

Das entwickelte EUS kann durch automatisierte Lösung des formulierten Optimierungsmodells die Fahrer den Kunden optimal zuordnen und gibt zudem die optimal zu fahrenden Routen inklusive der Gesamtfahrzeiten an. Dies entspricht einem Workaround zur Simulation der realen Situation des innovativen Lieferdienstes durch Verwendung von Echtzeitdaten, da innerhalb des entwickelten EUS nicht mit GPS-Schnittstellen gearbeitet wird. Dennoch muss eine dauerhafte und störungsfreie Verbindung zwischen Zentrale und Fahrern gewährleistet sein. Wie bei der durchschnittlichen Rechenzeit der größeren Beispielinstanz ersichtlich, steigt die Rechenzeit proportional zur Problemgröße, wie üblich für gemischt-ganzzahlige Optimierungsprobleme. Die Probleminstanz und die entsprechende Rechenzeit werden jedoch minimiert, indem immer auf Basis der aktuellen Zuordnungen, iterativ pro Bestellung optimiert wird und bestehende Zuordnungen teilweise fixiert werden. Durch dieses Vorgehen muss nicht der gesamte Datensatz eingelesen werden und der Lösungsbereich wird stark reduziert. Zudem werden Kundendaten nach erfolgreicher Belieferung aus dem Datensatz gelöscht, wodurch dieser möglichst gering gehalten wird. Eine weitere Limitation besteht darin, dass keine Teillieferungen möglich sind. Durch deren Einführung könnten alle verfügbaren Produkte gemeinsam betrachtet werden und nicht nur der Bestand, welchen ein einzelner Fahrer im Moment der Bestellung transportiert. Auf diese Weise könnten Bestellungen, die zuvor abgelehnt wurden ggf. ausgeliefert und der Servicegrad gesteigert werden. Eine zusätzliche Erweiterungsmöglichkeit des EUS könnte darin bestehen, historische Daten zu nutzen, um Fahrer am Tagesbeginn sowie in Zeiten ohne zugewiesene Kunden optimal im Liefergebiet zu platzieren. Dadurch könnten die Fahrzeiten der Fahrer, sowie die Wartezeiten der Kunden weiter verkürzt und hohe Nachfragen in Stoßzeiten effizienter bedient werden. Weiterhin gilt es zu beachten, dass sowohl das vorgestellte Lieferkonzept als auch das entwickelte Optimierungsmodell ausschließlich für vorab zubereitete, haltbare Mahlzeiten geeignet sind (siehe Beispiel mit Salaten und Smoothies, welche durch Kühlboxen hinreichend frisch gehalten werden können). Dadurch ist es nur für eine begrenzte Art an Lieferdiensten verwendbar. Insgesamt können on-demand Lieferdienste von vorab zubereiteten Mahlzeiten mit Hilfe eines EUS unterstützt werden, indem die Zuordnungen zwischen Fahrern zu Kunde, sowie die Routenbildung auf Basis eines EUS automatisiert wird. Insgesamt wird ein Beitrag zur Verkehrsentlastung geleistet, da im Falle des betrachteten Geschäftsmodells Elektroroller und Lastenräder zur Belieferung der Kunden verwendet werden können, welche weniger Verkehrsraum benötigen und durch ihre Fahrzeugeigenschaften agiler sind.

7 Fazit und Ausblick

In diesem Beitrag wird ein innovatives Geschäftsmodell für einen on-demand Lebensmittel-Lieferdienst vorgestellt, welches durch Verwendung mobiler Technologien und Datendienste kürzere Lieferzeiten und eine größere Effizienz gegenüber klassischen Lieferdiensten aufweist. Hierfür wird zur automatischen Zuordnung zwischen Fahrern und Kunden sowie zur Routenplanung ein gemischt-ganzzahliges Optimierungsmodell formuliert und in ein intuitiv nutzbares EUS integriert. Mit dem EUS ist es möglich, das operative Tagesgeschäft eines Lebensmittel-Lieferdienstes unter Verwendung von Echtzeit Daten zu simulieren. Das Geschäftsmodell und das entwickelte EUS als finales Artefakt liefern einen Beitrag zur anwendungsorientierten Forschung auf dem Gebiet der Wirtschaftsinformatik, indem die Nutzung des beschriebenen emissionsarmen Lieferkonzeptes gefördert wird.

Für das Optimierungsmodell sowie das EUS wurden zudem Erweiterungsmöglichkeiten vorgestellt. Die eingangs formulierte Forschungsfrage wird maßgeblich durch das entwickelte EUS beantwortet, welches für Anbieter vorab zubereiteter Mahlzeiten geeignet ist. Zukünftige Forschung könnte darauf fokussiert werden, das bestehende EUS und Optimierungsmodell hinsichtlich der genannten Nachteile zu verbessern. Zusätzlich könnten weitere Geschäftsmodelle untersucht bzw. neue Geschäftsmodelle mit ähnlicher Ausgangssituation (mehrere Startpunkte und offene Touren) analysiert werden, für welche das Optimierungsmodell angepasst werden könnte.

8 Referenzen

1. Jonas, A.E.G., McCann, E., Thomas, M.: Urban Geography. A critical introduction. Wiley Blackwell, Chichester, West Sussex (2015)
2. Watanabe, C., Naveed, K., Neittaanmäki, P., Fox, B.: Consolidated challenge to social demand for resilient platforms - Lessons from Uber's global expansion. Technology in Society 48, 33–53 (2017)
3. van Audenhove, F.J., Durance, M., Jongh, S. de: Urban Logistics-How to unlock value from last mile delivery for cities, transporters and retailers. Arthur D. Little Future of Urban Mobile Lab (2015)
4. Surekha, P., Sumathi, S.: Solution To Multi-Depot Vehicle Routing Problem Using Genetic Algorithms. World Applied Programming 1 (1), 118–131 (2011)
5. Crevier, B., Cordeau, J.-F., Laporte, G.: The multi-depot vehicle routing problem with inter-depot routes. European Journal of Operational Research 176, 756–773 (2007)
6. Brandão, J.: A tabu search algorithm for the open vehicle routing problem. European Journal of Operational Research 157, 552–564 (2004)
7. Aksen, D., Özyurt, Z., Aras, N.: Open vehicle routing problem with driver nodes and time deadlines. Journal of Operational Research Society 58 (9), 1223–1234 (2007)
8. Letchford, A.N., Lysgaard, J., Eglese, R.W.: A branch-and-cut algorithm for the capacitated open vehicle routing problem. Journal of Operational Research Society 58 (12), 1642–1651 (2007)
9. Pillac, V., Gendreau, M., Guéret, C., Medaglia, A.L.: A review of dynamic vehicle routing problems. European Journal of Operational Research 225, 1–11 (2013)
10. Hevner, A. R., March, S. T., Park, J., and Ram, S.: Design Science in Information Systems Research. Management Information System Quarterly 28 (1), 75–105 (2004)
11. Gholami, R., Watson, R. T., Molla, A., Hasan, H., and Bjørn-Andersen, N.: Information Systems Solutions for Environmental Sustainability: How Can We Do More? Journal of the Association for Information Systems (JAIS) (17:8), 521–536 (2016)
12. Hungerländer, P., Maier, K., Pöcher, J., Rendl, A., and Truden, C.: Solving an On-line Capacitated Vehicle Routing Problem with Structured Time Windows. In: Fink A., Fügenschuh A., Geiger M. (eds.) Operations Research Proceedings, 127–132, Springer, Cham (2016)
13. Jaillet, P., Wagner, M. R.: Online Routing Problems: Value of Advanced Information as Improved Competitive Ratios. Transportation Science 40 (2), 200–210 (2006)

Business Analytics - Nutzen und Grenzen analytischer Methoden zur Datenanalyse

Teilkonferenzleitung

Carsten Felden

Business Analytics in the Finance Department – A Literature Review

Markus Esswein[1], Peter Chamoni[1]

[1] University of Duisburg-Essen, Mercator School of Management, Duisburg, Germany
{markus.esswein,peter.chamoni}@uni-due.de

Abstract. Business analytics has become more and more relevant to practitioners and academics over the past few years. With vast amounts of data from in- and outside an organization and the availability of fast enough hardware to process it, all functions in an organization aim to benefit from analytics. We conduct a literature review looking at the finance department, i.e. financial and management accounting, and identify hotspots of current and potential future research. We summarize our findings in a framework for literature classification with the two dimensions accounting activity and rationale for business analytics. On this grid, better organization performance and better decision outcomes for most management accounting activities have been covered the most, while support of strategic and tactical goals as well as obtaining value from data should be considered more in detail by future research.

Keywords: Business analytics, predictive analytics, financial and management accounting, literature review.

1 Introduction

Business analytics is one of the main organizational levers to benefit from digitization [1]. However, the use of business analytics varies substantially among the different functions of a company. According to a 2014 study, 64% of respondents said they already use predictive analytics in marketing, with an additional 24% saying they will use it within the next three years [2]. Finance, on the other hand, was only mentioned by 39% and 26%, respectively. While this is still considerably more than, e.g., human resources (17% and 22%), it is noteworthy that the number-crunching finance department is not the first stop for advanced statistical methods. The demand, however, for business analytics adoption is clear when looking at practitioner literature. For example, in its 2017 CFO Agenda, the Hackett Group states that the finance department needs to step in and support the company strategy facing more constraints on funding and headcount and, secondly, provide the organization with more and better information [3].

The finance or accounting department has as a long tradition as a supporting function for corporate management [4]. It spans two areas: (1) *Financial accounting* addressing external stakeholders and covering bookkeeping, statutory reporting, and consolidation on the one hand [5], and (2) *management accounting* addressing internal stakeholders

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

and covering strategic cost management, planning and decision making, performance measurement as well as financial statement preparation [6] on the other hand. Thus, there are many opportunities to apply analytics in both financial and management accounting. Although, there is a wide range of applications for business analytics in the treasury and corporate finance function, too, the focus of this paper is on accounting.

Analytical (i.e. advanced statistical) techniques to gain insights from data are and have always been one of the main concerns in the field of statistics. Only today, the fast pace at which transactions are moving online allows for the collection of vast amounts of data [7]. Thus, analytics is becoming more relevant to practitioners (e.g., [8]) and scholars (e.g., [9]) alike. Building on the omnipresence of data arising from all kinds of sources such as enterprise systems, social networks, mobile devices, public data, and the internet-of-things, analytics goes beyond traditional business intelligence to generate better insights. Linking the data to a set of explanatory variables in order to determine causal inferences or in a predictive sense [10] enables a shift from a reactive towards a proactive, forward-looking management of the organization [11].

Although, there has been an increasing interest in the benefits of business analytics for the finance department, only certain aspects have been considered so far. Existing research mostly covers the impact of digitization on management accounting from a practitioner's perspective or the benefits of business analytics in general with no particular focus on the finance department.

Addressing the divide between potential use cases and actual application of business analytics, this paper will answer two research questions:

1. How can the application of business analytics to accounting be structured?
2. What has been the main focus of research for business analytics in accounting and where is potential for future research?

In order to answer these questions, we conducted a literature review following vom Brocke et al. [12] and propose a framework for business analytics in the finance department. After laying out our *research methodology* and giving an overview of the results in chapter 2, we will focus on the *framework for literature classification* in chapter 3 followed by an *agenda for future research* based on the identified research gaps in chapter 4 and a *conclusion* and outlook in chapter 5.

2 Research Methodology

Literature reviews are a widely accepted methodology not only as a first step for any research project, but also as a means to categorize existing research, present avenues for future research, and facilitate theoretical progress [12–14].

2.1 Search Strategy

We started our literature review with a *journal search* in leading journals followed by a backward and forward search to look for articles cited in the identified papers (backward) and newer articles citing the identified papers (forward) [14]. Since the focus of this research project is at the intersection of statistics and operations research on the one hand and accounting on the other hand, literature was searched "from both ends".

Regarding accounting, we chose the top ten accounting journals[1] in line with Nitzl [15], complemented them with the top information systems (IS) journals from the Senior Scolar's Basket of Journals[2] and with AIS conferences[3]. We then used the search terms business analytics and predictive analytics.

With respect to statistics and operations research, we chose five journals[4] from the list of top journals based on their scope and used the search terms management (or managerial) accounting and financial accounting.

As a second step of our literature search we broadened our scope and started a *comprehensive database search* in ScienceDirect, EBSCOhost, and Google Scholar combining different search terms according to the *citation pearl growing* approach [16]. We started with "finance" and "business analytics" and then widened our search to include different accounting, information systems, and planning terms in the finance context and predictive modeling and various forecast terms in the analytics context. For a list of search terms used in this second step, see **Figure 1**.

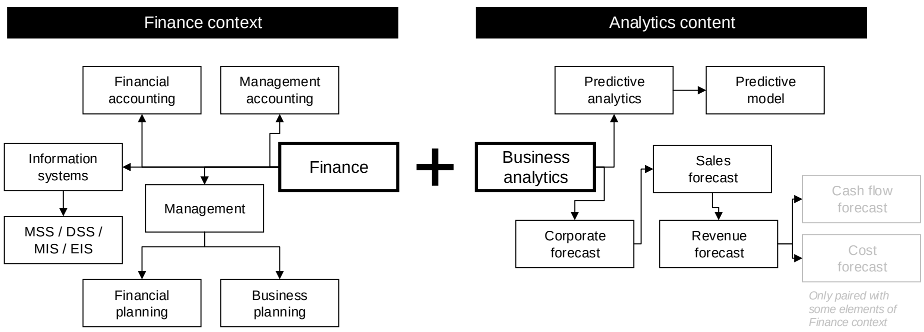


Figure 1. Citation pearl growing search terms for the database search

Due to the importance of the field to practitioners, a number of accounting organizations and consulting agencies have published surveys and point-of-view reports. In

[1] Journal of Accounting and Economics; Journal of Accounting Research; The Accounting Review; Management Accounting Research; Journal of Management Accounting Research; Contemporary Accounting Research; Behavioral Research in Accounting; Accounting, Auditing & Accountability Journal; Accounting and Business Research; Accounting, Organizations and Society

[2] European Journal of Information Systems; Information Systems Research; Information Systems Journal; Journal of the Association for Information Systems; Journal of Information Technology; Journal of Management Information Systems; Journal of Strategic Information Systems; MIS Quarterly

[3] Americas Conference on Information Systems; European Conference on Information Systems; International Conference on Information Systems; Pacific and Asia Conference on Information Systems

[4] International Journal of Forecasting; Journal of Forecasting; Operations Research; European Journal of Operational Research; International Journal of Production Economics

this work, however, the focus is on academic, peer reviewed, literature. While this may omit a number of recent developments, we consider it justified for a literature review. For an overview of practitioner statements regarding management accounting see, for example, [17].

2.2 Overview of Results

The journal search lead to a large number of initial hits (see **Table 1**), for which we checked title, abstract, and keywords. We ended up with an initial 22 relevant results from this first step[5].

Table 1. Search results

Journals	Search term	Total results	Relevant results
Accounting	Predictive analytics	13	**1**
	Business analytics	33	**2**
IS	Predictive analytics	26	**2**
	Business analytics	88	**4**
IS conferences	Predictive analytics	15	**0**
	Business analytics	52	6
Stats / OR	Financial accounting	84	**1**
	Management accounting	101	**6**
		Sum	**22**
Additional results from backward, forward, and database search			**47**
Total number of relevant results			**69**

From our consecutive backward and forward search and the broader database search, we found an additional 47 relevant results, which lead to a *total of 69 relevant results*.

Looking at the results on a timeline, different phases can be noted (**Figure 2**). (1) Owing to the roots of analytics in mathematical models that were developed already in the 1970s, e.g. the seminal work on time series analysis by Box-Jenkins [18], a small number of studies in the 1980s can be considered as relevant. They focus on the transfer of methods from operations research to management accounting [19] as well as the state of adoption of these methods [20, 21]. (2) By the end of the 1990s and early 2000s, sales forecasting was a common in practice and (fuzzy) neural networks were increasingly used [22, 23]. Additionally, there were further studies looking at the adoption of forecasting methods [24]. (3) The third phase is characterized by an increasing availability of data and a more comprehensive application of advanced statistical methods – by then called analytics. With a number of articles covering the added value of business or predictive analytics [25, 26], analytics adoption is no longer only a question of ability. It is a question of organizational transformation [10] and a new way of working with information in a digitally enabled business [27].

[5] Note that no sum for total results is given in **Table 1** because some results were found with both search terms. Relevant results, on the other hand, are all unique.

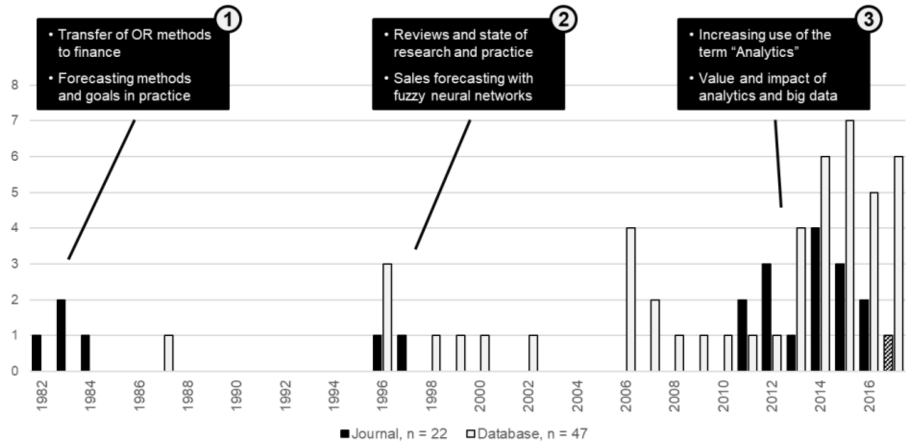


Figure 2. Search results on a timeline

3 Framework for Literature Classification

Based on the results of our literature search, we propose a framework to classify the existing applications of business analytics in financial and management accounting. With the help of this framework, we then identify what we call "hotspots" of current interest and potential hotspots of future interest.

3.1 Dimensions

Our framework has two dimensions: first, the accounting activities and second, the rationale for using business analytics with respect to a specific accounting activity.

Accounting Activities are the tasks that an accountant in financial or management accounting performs on a regular basis. Although, the scope of financial accounting is not the same for all companies, there is some common denominator in companies of a certain size. We follow a list of three activities in financial accounting presented by Horngren [5]: (1) *Bookkeeping* (incl. accounts payables, receivables, and credit management), (2) *Statutory reporting*, and (3) *Consolidation*.

Equally, management accounting can be set up differently in an organization, but four core tasks are common as well, as described by Blocher et al. [28] and Brands and Holtzblatt [6]: (1) Strategic (cost) management, (2) *Performance measurement*, (3) *Planning and decision making*, and (4) *Support in financial statement preparation*.

Rationale for Business Analytics is the reason why business analytics are applied in this specific situation. Generally, there are numerous possible nuances, however, we follow a list of six endogenous elements summarized by Holsapple et al. [29]: (1) *Achieving a competitive advantage*, (2) *Support of strategic and tactical goals*, (3) *Better organizational performance*, (4) *Better decision outcomes*, (5) *Knowledge production*, and (5) *Obtaining value from data*

3.2 Classification of the literature

Comparing the analytics coverage in financial accounting and management accounting, it is clear that the latter has attracted more attention. While this is partly due to the type of work in each of the two domains, it should not lead to an exclusion of financial accounting from consideration. In the following, we will propose three categories – less relevant (white shading), relevant (light grey shading), and highly relevant (middle grey shading) – of increasing interest and highlight some of the applications of business analytics in each of these categories. Our categorization is based on the nature of the activity and the general potential for statistical methods as well as current literature coverage. **Figure 3** shows the results for financial accounting with two highly relevant areas in bookkeeping, a couple of relevant and some less relevant areas.

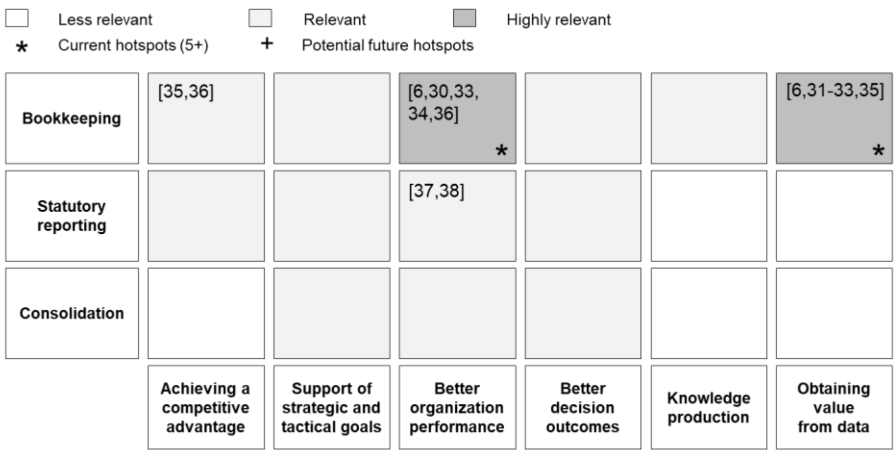


Figure 3. Classification for financial accounting

Brands and Holtzblatt [6] address better organization performance in bookkeeping and state that accounts payable and payment monitoring can greatly benefit from an analytics integration. Analytics can also help in choosing and contacting the right customers in order to improve collections cash flows [30]. Achieving a competitive advantage and obtaining value from data are generally not directly associated with bookkeeping, but can become relevant goals when it comes to fraud detection, bankruptcy prediction, or credit default prediction [31–36]. Dybvig [37] propose an optimized income statement improving organization performance in statutory reporting by including more accurate forecasts and Schneider et al. [38] see potential for predictive analytics in an early identification of financial accounting discrepancies.

Currently, there is only one hotspot of research with the integration of external data in order to improve credit default and bankruptcy prediction (bookkeeping – obtaining value from data). We do not see a real future hotspot, but consider better organization performance in bookkeeping very relevant. Amani and Fadlalla [39] found 11% of data mining applications in financial accounting, 25% in managerial accounting, and 64% in assurance and compliance. The papers cited for financial accounting apply neural networks or other data mining techniques to predict, e.g., quarterly cash flows, risk factors in financial statements or sentiments between different public statements. Yet, most of them take an external perspective, which is not the focus of our study.

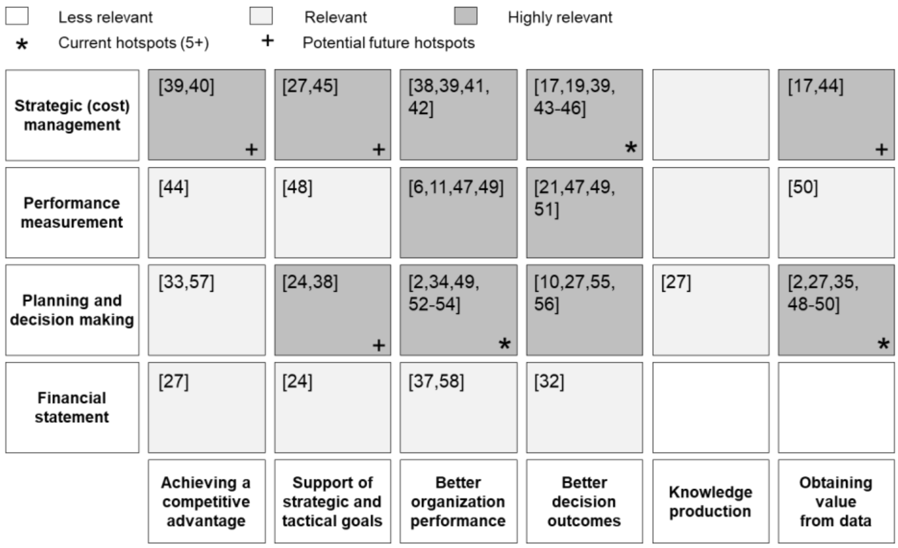


Figure 4. Classification for management accounting

With respect to management accounting, the overall picture is different. We see a number of highly relevant and only two less relevant areas in our grid. More in detail, researchers mention almost all rationales for business analytics in connection with strategic (cost) management. Marchant [40] states that management accountants are perfectly prepared to help management find ways to use data for a competitive advantage. Bhimani and Willcocks [27] consider the impact of novel forms of information on corporate strategy and goals and even organizational structures. Better organization performance, for example through the creation and revision of business rules with the help of business analytics, is addressed by many authors [38, 39, 41, 42]. Likewise, better decision outcomes, for instance using the analytical hierarchy process for cost driver selection [43] or through a holistic view and integrated thinking [17], are covered sufficiently [19, 39, 44–46]. Looking at performance measurement, there is less literature coverage of business analytics. Schläfke et al. [47] provide a framework that consists of the four layers capture (performance drivers in inputs, processes, and outputs), couple (performance drivers), control (knowing cause-effect relationships and crucial levers), and communicate (internally and externally). Recent conference proceedings look at critical success factors for business analytics in performance management to support strategic goals [48] or the mechanisms through which business analytics supports strategic decision making [49]. Further research emphasizes better decision outcomes or identifies ways to obtain value from data [11, 21, 50, 51].

Planning and decision making is another area of high interest. However, at the current point it is focused mainly on better organization performance [2, 34, 49, 52–54] and better decision outcomes [10, 27, 55, 56] due to more accurate and fact-based data, even for small and medium-sized companies. Business analytics is also applied to planning and decision making to achieve a competitive advantage, for instance with the help of a generalized advanced analytics competency in the finance department [33, 57], or to support strategic goals with improved forecasting [24]. Obtaining value from data [2, 27, 35, 48–50] was also covered from various angles like looking at what possible actions customers might take. Finally, financial statement preparation was covered only occasionally with articles focusing on the impact of selecting different accounting methods [32] or better organization performance in preparing the statements [37, 58].

Finally, there were a number of articles in top accounting or information systems journals, which did not address accounting-specific benefits of business analytics. Some of them highlighted a better organization performance from a general business perspective [59, 60] or obtaining value from data in the business functions [61, 62].

4 Discussion and agenda for further research

The combination of financial accounting data in the narrower sense and business analytics was mostly used from an external perspective on the company. This changes a bit when looking at credit default prediction, bankruptcy prediction or fraud detection. For example, as a predecessor of fraud detection, data quality issues and irregularities can be addressed with analytics. Consolidation is quite a complex activity; however, potential use cases for business analytics still need to be investigated. Knowledge production, on the other hand, seems less relevant in financial accounting. Bookkeeping in combination with better organization performance or obtaining value from data are currently the only two areas addressed by a number of researchers. Support of strategic and tactical goals as well as better decision outcomes come to mind as potential areas for further research.

Management accounting has clearly received more attention by researchers. Current hotspots with more than five articles are better decision outcomes in strategic (cost) management and two combinations with planning and decision making. Looking ahead, *strategic (cost) management* with business analytics to *achieve a competitive advantage, support strategic or tactical goals or to obtain value from data* should be addressed more in detail. Moreover, the use of business analytics in order to *support strategic or tactical goals in planning and decision making* should be addressed by future research. Despite these promising content areas, more research should focus on the application of *prescriptive analytics*. Currently, the focus is on descriptive and predictive analytics, as has been elaborated by [39]. Prescriptive analytics, on the other hand, goes one step further and combines analytics with intelligent automation.

5 Conclusion, limitations, and outlook

Motivated by the increasing interest in analytics by practitioners, we surveyed literature at the intersection of business analytics and the finance department. Looking at financial and management accounting separately we proposed a framework to get a comprehensive overview of motivations for business analytics with respect to different accounting activities. Identifying current hotspots like better decision outcomes for strategic (cost) management as well as planning and decision making, we also highlighted potential future hotspots like achieving a competitive advantage or obtaining value from data in strategic (cost) management. For research purposes, this paper contributes to a more comprehensive coverage of an emerging field of interest. For practice, it contributes to a more relevant and directed research, exploring possibilities in combining accounting activities and motivations for using business analytics. A next study should broaden the scope and include grey reports published by accounting organizations or consulting agencies. Besides, a closer look at the identified hotspots should be beneficial to researchers and practitioners alike.

References

1. Gartner: Gartner Survey Finds That Two-Fifths of IT Professionals Consider Their IT Organisation Ready for Digital Business (2016)
2. Halper, F.: Predictive Analytics for Business Advantage (2014)
3. Essaides, N., Willman, T., O'Connor, J.: The CFO Agenda: Finance's Top Four Strategic Priorities in 2017 (2017)
4. Strahlem, R.E.: Introduction to accounting. Dickenson, Encino (1972)
5. Horngren, C.T.: Introduction to financial accounting. Prentice-Hall, Englewood Cliffs (1981)
6. Brands, K., Holtzblatt, M.: Business Analytics. Transforming the Role of Management Accountants. Management Accounting Quarterly 16 (2015)
7. Agarwal, R., Dhar, V.: Editorial—Big Data, Data Science, and Analytics. The Opportunity and Challenge for IS Research. Information Systems Research 25, 443–448 (2014)
8. Lavalle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Special Report. Analytics and the New Path to Value. MIT Sloan Management Review 52, 22–32 (2011)
9. Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics. From big data to big impact. MIS Quarterly 36, 1165–1188 (2012)
10. Baesens, B., Bapna, R., Marsden, J.R., Vanthienen, J., Zhao, J.L.: Transformational Issues of Big Data and Analytics in Networked Business. MIS Quarterly 40, 807–818 (2016)
11. Cokins, G.: Enterprise Performance Management (EPM) and the Digital Revolution. Performance Improvement 56, 14–19 (2017)
12. Vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., Cleven, A.: Reconstructing the giant. On the importance of rigour in documenting the literature search process. In: Newell, S., Whitley, E.A., Pouloudi, N., Wareham, J., Mathiassen, L. (eds.) Proceedings of the 17th European Conference on Information Systems, Verona, Italy 2009 (2009)

13. Cooper, H.M.: Synthesizing research. A guide for literature reviews. Sage Publications, Thousand Oaks, Calif (2010)
14. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future. Writing a literature review. MIS Quarterly 26, xiii–xxiii (2002)
15. Nitzl, Christian: The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research. Directions for future theory development. Journal of Accounting Literature 37, 19–35 (2016)
16. Rowley, J., Slack, F.: Conducting a literature review. Management research news 27, 31–39 (2004)
17. Nielsen, S.: New and Interesting Perspectives for the Management Accountant in a World of Data
18. Box, G.E.P., Jenkins, G.M.: Time series analysis. Forecasting and control. Holden-Day, San Francisco (1970)
19. Kriens, J., van Lieshout, J., Roemen, J., Verheyen, P.: Management accounting and operational research. European Journal of Operational Research 13, 339–352 (1983)
20. Brown, L.D., Griffin, P.A.: Perspectives on forecasting research in accounting and finance. Journal of Forecasting 2, 325–330 (1983)
21. Dino, R.N., Riley, D.E., Yatrakis, P.G.: The role of forecasting in corporate strategy. The Xerox experience. Journal of Forecasting 1, 335–348 (1982)
22. Kuo, R.J., Wu, P., Wang, C.P.: An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination. Neural Networks 15, 909–925 (2002)
23. Kuo, R.J., Xue, K.C.: A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights. Decision Support Systems 24, 105–126 (1998)
24. Winklhofer, H., Diamantopoulos, A., Witt, S.F.: Forecasting practice. A review of the empirical literature and an agenda for future research. International Journal of Forecasting 12, 193–221 (1996)
25. Shanks, G., Bekmamedova, N.: Achieving benefits with business analytics systems. An evolutionary process perspective. Journal of Decision Systems 21, 231–244 (2012)
26. Seddon, P.B., Constantinidis, D., Tamm, T., Dod, H.: How does business analytics contribute to business value? Info Systems J 27, 237–269 (2017)
27. Bhimani, A., Willcocks, L.: Digitisation, 'Big Data' and the transformation of accounting information. Accounting and Business Research 44, 469–490 (2014)
28. Blocher, E.J., Stout, D.E., Cokins, G.: Cost management. A strategic emphasis. McGraw-Hill/Irwin, New York (2010)
29. Holsapple, C., Lee-Post, A., Pakath, R.: A unified foundation for business analytics. Decision Support Systems 64, 130–141 (2014)
30. Buckham, C.: Using Predictive Modelling to Improve Collections. Credit Control 27, 54–59 (2006)
31. Acito, F., Khatri, V.: Business analytics. Why now and what next? Business Horizons 57, 565–570 (2014)
32. Paliwal, M., Kumar, U.A.: Neural networks and statistical techniques. A review of applications. Expert Systems with Applications 36, 2–17 (2009)
33. Davenport, T., Tay, A.: Ramping Up the Role Of Analytics Leader. CFO 32, 14–16 (2016)

34. Ramrathan, D., Sibanda, M.: Impact of Analytics in Financial Decision Making. Evidence from a Case Study Approach. Acta Universitatis Danubius: Oeconomica 10, 124–135 (2014)
35. Tangsucheeva, R., Prabhu, V.: Stochastic financial analytics for cash flow forecasting. International Journal of Production Economics 158, 65–76 (2014)
36. Luber, M.: Employ Statistics and Analytics to Up Your Collections Game. Controller's Report 2013, 8–9 (2013)
37. Dybvig, A.: A New Application of TRUE Math Programming Optimization for Finance. The Optimized Income Statement (OIS). Journal of Corporate Accounting & Finance (Wiley) 27, 11–16 (2016)
38. Schneider, G.P., Dai, J., Janvrin, D.J., Ajayi, K., Raschke, R.L.: Infer, Predict, and Assure. Accounting Opportunities in Data Analytics. Accounting Horizons 29, 719–742 (2015)
39. Amani, F.A., Fadlalla, A.M.: Data mining applications in accounting. A review of the literature and organizing framework. International Journal of Accounting Information Systems 24, 32–58 (2017)
40. Marchant, G.: Management Accounting in the 21st Century. A Profession for which the Time Has Come. Journal of Applied Management Accounting Research 11, 1–4 (2013)
41. Quaadgras, A., Weill, P., Ross, J.W.: Management commitments that maximize business impact from IT. Journal of Information Technology 29, 114–127 (2014)
42. Nielsen, E.H., Jacobsen, A., Pedersen, L.B.: Management Accounting and Business Analytics. Dan. J. Manag. Bus. 78, 31–44 (2014)
43. Schniederjans, M.J., Garvin, T.: Using the analytic hierarchy process and multi-objective programming for the selection of cost drivers in activity-based costing. European Journal of Operational Research 100, 72–80 (1997)
44. Appelbaum, D., Kogan, A., Vasarhelyi, M., Yan, Z.: Impact of business analytics and enterprise systems on managerial accounting. International Journal of Accounting Information Systems 25, 29–44 (2017)
45. Cokins, G.: Top 7 trends in management accounting. Strategic Finance 95, 21–30 (2013)
46. Nielsen, S.: The Impact of Business Analytics on Management Accounting, Available at SSRN: https://ssrn.com/abstract=2616363
47. Schläfke, M., Silvi, R., Möller, K.: A framework for business analytics in performance management. International Journal of Productivity and Performance Management 62, 110–122 (2012)
48. Shanks, G., Bekmamedova, N., Willcocks, L.: Business Analytics. Enabling Strategic Alignment and organisational Transformation. In: Pries-Heje, J., Chiasson, M., Wareham, J., Busquets, X., Valor, J., Seiber, S. (eds.) Proceedings of the 20th European Conference on Information Systems, Barcelona, Spain 2012 (2012)
49. Cao, G., Duan, Y.: The Affordances of Business Analytics for Strategic Decision-Making and Their Impact on Organisational Performance. In: Pan, S.L., Cao, T.H. (eds.) Proceedings of the 16th Pacific Asia Conference on Information Systems, Ho Chi Minh City, Vietnam 2012 (2012)
50. Sun, Z., Strang, K., Firmin, S.: Business Analytics-Based Enterprise Information Systems. Journal of Computer Information Systems 57, 169–178 (2017)
51. Guo, L., Shi, F., Tu, J.: Textual analysis and machine leaning. Crack unstructured data in finance and accounting. The Journal of Finance and Data Science 2, 153–170 (2016)

52. Lu, C.-J., Kao, L.-J.: A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server. Engineering Applications of Artificial Intelligence 55, 231–238 (2016)
53. Duan, L., Xiong, Y.: Big data analytics and business analytics. Journal of Management Analytics 2, 1–21 (2015)
54. Zotteri, G., Kalchschmidt, M.: Forecasting practices. Empirical evidence and a framework for research. International Journal of Production Economics 108, 84–99 (2007)
55. Quattrone, P.: Management accounting goes digital. Will the move make it wiser? Management Accounting Research 31, 118–122 (2016)
56. Dinan, T.P.: Predictive Analytics Can Move You from Scorekeeper to Proactive Manager. Pennsylvania CPA Journal 86, 9–10 (2015)
57. Payne, R.: Discussion of 'Digitisation, 'Big Data' and the transformation of accounting information' by Alnoor Bhimani and Leslie Willcocks (2014). Accounting and Business Research 44, 491–495 (2014)
58. Bräuning, M., Hüllermeier, E., Keller, T., Glaum, M.: Lexicographic preferences for predictive modeling of human decision making. A new machine learning method with an application in accounting. European Journal of Operational Research 258, 295–306 (2017)
59. Asadi Someh, I., Shanks, G.: How Business Analytics Systems Provide Benefits and Contribute to Firm Performance? In: Becker, J., Vom Brocke, J., Marco, M. de (eds.) Proceedings of the 23th European Conference on Information Systems, Münster, Germany 2015 (2015)
60. Wixom, B.H., Yen, B., Relich, M.: Maximizing Value from Business Analytics. MIS Quarterly Executive 12 (2013)
61. Lacity, M.C., Solomon, S., Yan, A., Willcocks, L.P.: Business process outsourcing studies. A critical review and research directions. Journal of Information Technology 26, 221–258 (2011)
62. Sharma, R., Mithas, S., Kankanhalli, A.: Transforming decision-making processes. A research agenda for understanding the impact of business analytics on organisations. European Journal of Information Systems 23, 433–441 (2014)

Success or Failure of Big Data: Insights of Managerial Challenges from a Technology Assimilation Perspective

Johannes Weibl[1] and Thomas Hess[1]

[1] LMU Munich, Institute for Information Systems and New Media
{weibl,thess}@bwl.lmu.de

Abstract. The popularity of big data analytics (BDA) as one of the leading digital technologies in organizations has increased tremendously. However, organizations, currently encounter various barriers to the successful adoption of BDA. Grounded in technology assimilation theory, this study goes beyond a binary view of BDA adoption and describes crucial managerial challenges along the more complex process of assimilation by conducting an exploratory and international Delphi study in collaboration with 21 experts. To exploit the full potential of BDA our results reveal that the assimilation process requires 1) appropriate organizational structures and 2) well defined business cases and business value from the very beginning (initiation phase) and on a continuous basis. From an academic perspective, the study sheds light on crucial challenges that influence the assimilation of BDA on its pathway to maturity. In addition, this study contributes to assimilation theory by providing a new perspective for BDA.

Keywords: Big Data Analytics, Assimilation, Challenges

1 Introduction

The increasing evidence for the potential benefits of big data analytics (BDA) has led the majority of all Fortune 500 companies (85%) in the US, to invest in BDA [1]. These organizations expect to create new business models, as well as products and services that support, optimize and automate organizational decisions and processes by analyzing large amounts of data [2]. BDA is increasingly incorporated in organizations in order to establish a more agile and efficient decision-making process and is often seen, as "one of the major innovations in the last decade" [3, p.1]. It enables organizations to transform towards a data-driven company and is therefore one of the leading digital technologies to support the digital transformation of companies [4].

However, as for many new technologies or information systems (IS), the process of implementing BDA in organizations poses challenges. According to Gartner [5], 60 percent of BDA projects fail because they are not completed within budget or on schedule, or fail to deliver certain features. Other practitioners see organizational change issues and insufficient access to data as relevant barriers [6]. Further reports claim that due to these various challenges more than half of BDA projects fail to go

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

farther than piloting and experimentation with Big Data [7]. Prior research has also started to observe some of the challenges of introducing BDA. For example, studies have identified technical issues, such as data security and privacy issues, and data integration complexities [8]. Other authors emphasize structural challenges, such as insufficient resources in general or a lack of skilled personal [9]. Overall, research and practical publications both highlight a diverse spectrum of managerial challenges that may be relevant for successful BDA projects.

However, it still remains unclear why BDA projects fail during different stages of introduction. Typically, the stated barriers are mainly examined with the assumption that there is just a single point of introduction. But the introduction of any technological innovation is usually not binary, being instead multi-staged and often a highly complex phenomenon [10]. Various issues which occur before, during and after implementation can put the success of BDA at risk. Furthermore, current research lacks adequate analysis of the prioritization of potential barriers, which need to be addressed primarily to leverage the full potential of BDA initiatives. Therefore, our study goes beyond viewing the implementation and use of BDA in a binary manner and instead analyzes the more complex phenomenon of BDA assimilation by drawing on the technology assimilation theory [e.g. 11, 12] to identify the most crucial barriers along the multi-staged process. Thus, this paper addresses the following research question: *Which crucial managerial challenges need to be addressed for a successful assimilation of Big Data Analytics in organizations?*

In order to answer this question, we build on the literature of BDA and the theory of technology assimilation as a theoretical foundation. We decided to perform an exploratory Delphi study in the field of BDA, because we recognized there was a lack of prior knowledge to build upon. Our study provides insights into the introduction of BDA in organizations and its related managerial challenges. It highlights crucial barriers of a digital technology along the assimilation process in organizations.

The paper is organized as follows: First (section 2), we present the background for our study by describing BDA and the technology assimilation theory. Next (section 3), we explain our methodological approach and describe the four-phase Delphi study applied to our context. In the results and discussion section (section 4 and 5), we present and discuss the managerial challenges identified, prioritized and highlighted along the assimilation stages by our experts. Finally, we suggest implications for research and practice and state the limitations of our research (section 6).

2 Background

This section briefly discusses the literature of BDA and related managerial challenges as well as the technology assimilation theory that is used to structure our Delphi study and interpret its results.

2.1 Big data analytics and related managerial challenges

Big Data as a term refers to the process of managing large amounts of data that come from several, heterogeneous data sources (e.g. internal and external, structured and unstructured) that can be used for collecting and analyzing enterprise's data [13]. In this paper we adopt a definition of BDA as techniques (e.g., analytical methods) and technologies (e.g., databases and data mining tools) that a company can employ to manage and analyze large-scale, complex data for various applications intended to augment a business's performance [14]. This definition comprises high-tech data storage, management, analysis capability, and visual technologies as essential parts of BDA [1]. BDA is about the extraction of unknown patterns, correlations and information across multiple sources of data to enrich the information depths and produce new insights for decision makers [15]. Subsequently, the implementation and use of BDA can enable both efficient managerial decisions as well as leverage process improvements [16].

However, the introduction of BDA in organizations poses challenges that need to be addressed to fully leverage its potential. In previous IS-related studies on the adoption of BDA several generic factors affecting the adoption process [e.g. 17], as well as detailed determinants (e.g. data privacy issues [8]) have been discussed. Typically the identified barriers are related to people, technology, strategic or organizational domains of a company. We follow Alharti et al. [18], who reported that the adoption barriers of BDA fall into three categories: technological, human and organizational.

Organizational barriers are typically those factors in the organization's structure and culture that are not compatible with the new technology. These can include, among others, communication, authority flows, and show how the company has traditionally been working. Typical symptoms are a failure to perceive the strategic benefits of investment and a lack of co-ordination and co-operation due to organizational fragmentation [19]. Human barriers can be related both to employees lacking appropriate data analysis skills, and to challenges related to privacy and protection of personal information within and outside of organizations [18]. Technical barriers are factors in the information technology itself, such as finding a solution that integrates the new big data technology with existing legacy IT systems or solving the integration complexity of heterogeneous data sets [18].

2.2 Theory of technology assimilation

The theory of technology assimilation can be understand as an organizational process "unfolding in a series of decisions to evaluate, adopt, and implement new technologies" [20, p. 897]. It is defined as "the extent to which the use of the technology diffuses across the organizational projects or work processes and becomes routinized in the activities of those projects and processes" [21, p. 121]. Since the early applications of the technology assimilation theory to IS research, it has been applied and adapted in many ways. The range of application fields reaches from material requirements planning [11] to e-business [22] and cloud computing [10].

The theory of technology assimilation suggests that there is not just a single point of introduction of a new innovation in an organization. It is rather a multi-staged, sequential process, which is often far from simple and rarely unfolds in a linear and smooth manner [10]. There are often several different points in time where the assimilation may intensify or deteriorate and various obstacles and challenges need to be addressed [12]. Potential users, for example, face difficulties in learning about a new technology and in understanding how they must adapt their work-process activities to be able to use the new technology in an efficient way [23, 24]. In this context, potential users are more likely to use technologies that are perceived as user-friendly, having a clear benefit over existing ways of doing their job, being rather less complicated to use, and being suitable with the existing work processes and work domain [25].

In the IS literature several researchers have proposed various process and stage models describing the technology implementation process in organizations. These models are valuable in analyzing the context in which events (such as barriers) occur and show the causal linkages and temporal relationships between them [12]. Amongst the most cited is that of Gallivan [12], who created a six-staged assimilation model, based on the work of Zmud and colleagues [11], shown in table 1.

Table 1. Innovation assimilation stages [12]

Stages	*Description*
Initiation	A match is identified between an innovation and its intended application in the organization.
Adoption	The decision is made to invest resources to accommodate the implementation effort.
Adaption	The innovation is developed, installed and maintained, and organizational members are trained both in the new procedures and in the innovation.
Acceptance	Organizational members are committed to using the innovation.
Routinization	Usage of the innovation is encouraged as a normal activity in the organization.
Infusion	The innovation is used in a comprehensive and sophisticated manner which leads to increased organizational effectiveness.

3 Method: A Delphi-study design

The Delphi method is based on expert knowledge and aims to reach consensus on a specific research question via a structured process of iterative questionnaires with controlled feedback [26-28]. In IS research, the Delphi method has a long history of applications [29] and has been applied in the context of BDA recently [30]. This methodological approach enables an effective communication process through iterative rounds of feedback and prevents the direct confrontation of participants that could lead to potential biases [26, 29]. The expert panel size typically ranges between 10 to 30 experts [29].

The Delphi method is a useful explorative approach to gain insights from the collective experience of practitioners, in particular when the literature lacks sufficient empirical studies [27, 31]. Therefore, the Delphi method appeared to be the most useful approach in order to address our research question. By following the procedures and quality criteria stated by several authors [e.g. 26, 27, 28], we aimed to meet the call for higher methodological rigor in Delphi studies as well as this method's soundness [27].

3.1 Panel selection

Since the results of Delphi studies are entirely based on the panelists' statements and indications, the selection of suitable experts is a critical factor. Okoli and Pawlowski [38] suggested a five-step approach for the selection procedure in Delphi studies. We followed this process by identifying experts with differing professional backgrounds comprising consulting, research, IT and information management, technical sales, as well as IT focused project management. Subsequently, we compared the qualifications of the potential participants to create a classification and ultimately a prioritization methodology for the experts by a category system in order to ensure profound expertise among the panelists. For instance, a mandatory factor for researchers was to be currently involved in big data research. For consultants, corporation employees and other types of occupations, we decided to only choose persons in senior positions with long work experience in the field of data analytics. For example, practitioners had to have contributed to the set-up or usage of BDA systems on a practical level. On a strategy level, they had to be involved in decisions of BDA initiatives.

Based on the selection criteria, the first potential panelists were contacted through the research teams' professional networks. In addition, the authors searched in business- and employment-oriented social networking services (e.g. LinkedIn) for practitioners with expertise in the field of BDA. We identified 284 potential practitioners and 12 potential academic experts, which were invited to take part in the study. We aimed for a panel size of 20 experts to account for possible drop-outs during the study period.

Altogether, 21 BDA experts committed to participating and completed the first round of the Delphi study (corresponding to a response rate of 7.4%). The panelists' primary functional affiliation ranged from big data initiatives in organizations (30% of panelists) to academic research (20% of panelists) to consulting (50% of panelists) in various industries (e.g. aviation, banking, energy, IT). All experts had been extensively involved in big data projects in the United States, Europe, Brazil or Australia. The average number of years working in managing information technologies (IT) within our panel was 20 years, with 5 years working specifically in the field of BDA. The panelists held various senior positions with responsibilities for BDA initiatives, including positions such as CEO, head of BI applications, or managing director at management consultancies. All academic experts held chairs at universities, except for one associate professor.

3.2 Data collection and analysis

The entire Delphi study was conducted with an online survey platform. We conducted a total of four rounds over the period from May-June 2017, allowing one week for the experts to respond to each round. Extending the Delphi procedure of Schmidt [28], this study segmented data collection and analysis in four distinct phases: Brainstorming (1), Selection (2), Ranking (3) and Assignment (4), see table 2.

Phase 1 (Brainstorming) serves to brainstorm and identify a broad range of relevant managerial challenges. To do this, we provided the panelists with our definition and understanding of BDA as outlined in the introduction and theoretical section of this study to ensure a common understanding. We asked the experts to name and briefly explain at least 5 managerial challenges associated with BDA in their organizations. In the interest of not limiting the diversity of the initial value set, we did not bound the number of possible responses [28]. After obtaining 107 suggested challenges, the research team consolidated the list by deleting duplicates and sharpening the descriptions. The compiled results were subsequently discussed collectively by the research team in order to create one reliable and consolidated list of managerial challenges. It was then sent to all experts for the purpose of validation. After incorporating the experts' feedback, the final list resulted in 25 BDA-related managerial challenges. Following Alharti et al. [18], the research team classified the list of challenges into organizational, human and technological barriers according to their interrelation.

Table 2. Overview of the data collection and analysis process

Phase	*Objective and panelists*	*Panelists*
Brainstorming (1)	Collection and validation of initial managerial challenges	21
Selection (2)	Selection of most crucial challenges	13
Ranking (3)	Ranking of most crucial challenges	13
Assignment (4)	Highlighting of most crucial challenges along the assimilation process	13

The second phase (Selection) served to identify the most important managerial challenges. For this purpose each participant was requested to state, in his or her judgement, the 10 most crucial challenges for a successful assimilation of BDA on a randomized list of all challenges. Subsequently, the research team further consolidated the list on the basis of the experts' selection. Following Piccinini et al. [32], we aimed to receive 12-15 crucial managerial challenges and therefore cut-off each challenge on the list that had been selected by less than 30% of the panelists. This procedure led to 13 crucial managerial challenges at the end of the selection phase.

In phase 3 (Ranking), the experts were asked to rank the shortened list of challenges in order of their importance for a successful assimilation of BDA. To do this, the shortened list of challenges was sent to the panelists in a randomized order. In order to investigate whether a consensus among the experts had been reached, the Kendalls' coefficient of concordance (W) was calculated [28]. According to Schmidt [28], consensus levels are appraised as weak at W = 0.3, moderate at W = 0.5, and strong at W = 0.7. The ranked results in phase 3 implicate a Kendall's W consensus of 0.14.

In phase 4 (Assignment), the experts were asked to highlight each of the 13 managerial challenges of the shortened list along the six stages of the process of technology assimilation (multiple answers allowed). In keeping with other Delphi studies [e.g. 33], the research team identified an assimilation stage as relevant for a managerial challenge, if the majority of panelists (> 50%) assigned a specific challenge to one assimilation stage.

4 Results

In this section, we present the final results of our Delphi study. During the brainstorming and selection phases, crucial managerial challenges for the assimilation of BDA were identified, consolidated and condensed to the most important items. Table 3 shows the results of the selection phase of our study.

Table 3. Results of the selection phase of our Delphi study

Category	*Managerial challenges of BDA in organizations*	*Phase 2 – Selection**
Organizational Challenge (OC)	Lack of long-term view of data as an asset and its potential future business value	69
	Lack of clear use cases to motivate big data projects	69
	Lack of an enterprise wide data strategy	62
	Lack of top management commitment to big data projects	62
	Insufficient commitment or resistance to change of non-IT project stakeholders	62
	Gaps in alignment of goals and expectations between IT and business units	54
	Risk averse culture	31
Human Challenge (HC)	Lack of BDA-related skills of employees	69
	Lack of knowledge of end users in how to interpret and transform data analytic results	62
Technological Challenge (TC)	Data integration complexity	77
	Data quality issues	69
	Insufficient data governance practices	46
	Selection of an appropriate BDA technology	38

Note: * indicates the % of panelists who selected this challenge as very important

Subsequent to selecting the most crucial managerial challenges, the experts were asked to rank the challenges according to their importance (phase 3) and to highlight the ranked results along the six-stages process of technology assimilation (phase 4), as shown in figure 1.

Especially noticeable is the aspect that the top ranked managerial challenges of assimilating BDA already occur in the first two stages of the assimilation process and belong to the organizational domain of barriers. Human challenges, such as lack of skills and knowledge, are not seen as that crucial and can be addressed during the adaption, acceptance and routinization stage of BDA assimilation. Technological challenges are rather subordinated.

Phase 3		Crucial managerial challenges of BDA along the assimilation process	Phase 4 - Assignment					
Rank (mean rank)	Cat.		1. Initiation	2. Adoption	3. Adaption	4. Acceptance	5. Routinization	6. Infusion
1 (4)	OC	Lack of top management commitment to big data projects	■	■				
2 (4,5)	OC	Lack of long-term view of data as an asset and ist potential future business value	■	■				
3 (5,5)	OC	Insufficient commitment or resistance to change of non-IT project stakeholders					■	
4 (6)	OC	Lack of clear use cases to motivate big data projects	■	■	■			
4 (6)	TC	Data integration complexity			■			
6 (6,5)	OC	Lack of an enterprise wide data strategy		■	■			
6 (6,5)	TC	Data quality issues	■		■	■	■	■
8 (7)	HC	Lack of BDA-related skills of employees			■	■		
9 (7,5)	TC	Insufficient data governance practices			■	■	■	
10 (8)	HC	Lack of knowledge of end users how to interpret and transform data analytic results				■	■	
11 (10)	OC	Gaps in alignment of goals and expectations between IT and business units			■		■	
12 (11)	OC	Risk averse culture	■	■				
13 (11)	TC	Selection of an appropriate big data technology	■	■	■			

Kendall's coefficient (W): 0.14

Figure 1. Crucial managerial challenges of BDA along the assimilation process

5 Discussion

The findings of our study indicate that the managerial challenges of assimilating BDA are influenced by various factors, such as structural, technical, staffing and strategy-related issues which can be categorized according to the domains of technology, people and organization. In line with previous work [3] is the finding that introducing BDA in organizations is not a simple technical issue per se, but rather an organizational transformational challenge. Our panelists ranked most of the organizational issues, such as lack of top management commitment, as the most crucial challenges to be addressed for a successful assimilation of BDA in organization. Technical issues and human barriers are subordinated in the ranking.

Furthermore the result of the Delphi study indicate that lack of senior management support or insufficient commitment of the top management team prevents the successful introduction of BDA directly at the beginning of the assimilation (initiation phase). Analogous to that is the implication of having no or insufficient use cases for BDA. Our experts perceive the lack of clear use cases as crucial. This often directly hinders any BDA initiatives in the initiation and adoption phase. According to Rogers [25] an assimilation of a new technology will be more successful when it has a clear relative advantage over the existing business system. Clear use cases could demonstrate their potential future business value and would lead to sufficient commitment from project stakeholders.

Additionally, Figure 1 shows that the majority of managerial challenges of BDA occur in particular during the first three assimilation stages, namely initiation, adoption and adaption phases. This basically leads to two conclusions: First, several serious potential issues need be addressed even before the actual implementation process of a BDA initiative begins, otherwise every assimilation effort is strongly endangered. Second, the result may be indicative of the evolutionary stage companies are at on the BDA implementation journey. The majority of the crucial challenges are highlighted in particular within stages one to three of the assimilation process. Our expert panel is apparently mostly involved in managing those barriers in the first stages and perceives these as very important, which might suggest that many organizations are still at a pre-implementation or implementation stage of BDA initiatives. This finding expands Kiron and Shockley's [34] statement, that many organizations are still at a reactive stage, in which they deal with various issues of managing data itself while not necessarily adopting and using BDA in a comprehensive and sophisticated manner.

Our Delphi study reveals two additional results which have not, to the best of our knowledge, previously received appropriate attention in the current literature. Firstly, our study shows that the necessity of ridding a company of a risk averse culture is crucial right from, or even before, the beginning of a BDA initiative. Enabling a trial and error mentality, as well as an experimental setting to test various analytic approaches and derive potential use cases is necessary to be developed before the actual BDA implementation process starts. Secondly, in contrast to previous work which has often stated psychological challenges, resistance to loss of power or status while implementing complex information technologies [e.g. 19] does not play a crucial role in BDA initiatives in the view of our panelists.

6 Contribution, Limitations and future research

This study has several implications for research and practice. We make two specific contributions to the literature of innovation assimilation. First, our study extends the literature [e.g. 12] by emphasizing the context in which assimilation events, namely crucial managerial challenges, occur, and shows the sequential relationship between them. Second, although some overlaps to other studies analyzing the barriers of technology assimilation do exist, such as the fact that successful assimilation requires senior management support [22], our study reveals that organizational challenges are the most important barriers to be addressed. This contradicts Conboy and Morgan [10], for instance, who state that the biggest barrier to the technology assimilation is people-related. We also contribute to the contemporary BDA research by providing an overview of crucial BDA barriers in a sequential, multi-staged way while analyzed along the assimilation process.

For practitioners, our study sheds light on the question of why the majority of BDA projects fail [5] and which obstacles need to be addressed primarily. Our results suggest that companies often lack relevant organizational and structural aspects during the initiation and adaption stage. To exploit the full potential of a BDA project in the long-term, an organization has to successfully address the following challenges directly at the start of any BDA initiative: 1) Demand top management commitment from the very beginning and on a continuous basis 2) Create clear use cases upfront to motivate every project stakeholder and to demonstrate its future business value. People related challenges, such as training efforts and knowledge management are subordinated and can be addressed during the adaption, acceptance and routinization stage. The priority ranking of the selected challenges forms the starting point to set up BDA projects accordingly, bring relevant stakeholders to the table and direct their resources toward addressing the most important challenges in each assimilation stage.

Furthermore, we acknowledge that our Delphi study has some limitations. As with any Delphi-type study, the results rely on the experience and opinions of a limited number of individuals, in our case 21 experts. Although the panel size is comparable to other IS-related Delphi studies [31, 35] and is not required to be statistically representative [26], we must be reserved when generalizing the results [35]. A further limitation of the study is related to the relatively low level of consensus regarding the ranking phase of our study, indicated by a Kendall's W of 0.14. It might be explained by the broad diversity of our panelist's experiences concerning BDA initiatives in organizations as shown in other contemporary IS-related Delphi studies [e.g. 32].

In spite of the limitations, we are confident that our results serve as a fruitful starting point for empirically analyzing the challenging assimilation process of BDA in organizations. Our study lays the groundwork for research projects on how to overcome the identified challenges in each assimilation stage and find potential solutions. In a next step the panel could be extended to reach a broader group of experts, including different stakeholders such as potential recipients of BDA projects. Furthermore, the approach could change to one that is more in-depth, such as case studies to understand how BDA is being used in different industry settings at a more granular level, such as in the energy sector, or for specific divisions in organizations.

References

1. Kwon, O., Lee, N., Shin, B.: Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management 34, 387-394 (2014)
2. Bughin, J., LaBerge, L., Mellbye, A.: The case for digital reinvention. McKinsey Quarterly, January 1-15 (2017)
3. Kung, L., Kung, H.-J., Jones-Farmer, A., Wang, Y.: Managing Big Data for Firm Performance: A configurational approach. Proceedings of the 21st AMCIS, pp. 1-9, Puerto Rico (2015)
4. Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M.: Embracing digital technology: A new strategic imperative. MIT Sloan Management Review 55, 1-15 (2014)
5. Gartner: Gartner Predicts Business Intelligence and Analytics Will Remain Top Focus for CIOs Through 2017. In: Meulen, R.v.d. (ed.), http://www.gartner.com/newsroom/id/2637615 (2013)
6. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H.: Big data: The next frontier for innovation, competition, and productivity. (2011)
7. Gartner: Gartner Says Business Intelligence and Analytics Leaders Must Focus on Mindsets and Culture to Kick Start Advanced Analytics. In: Goasduff, L. (ed.), http://www.gartner.com/newsroom/id/3130017 (2015)
8. Gandomi, A., Haider, M.: Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management 35, 137-144 (2015)
9. Kim, G.-H., Trimi, S., Chung, J.-H.: Big-data applications in the government sector. Communications of the ACM 57, 78-85 (2014)
10. Conboy, K., Morgan, L.: Assimilation of the Cloud: Challenges to Acceptance, Routinisation and Infusion of Cloud Computing Proceedings of the 33d ICIS, pp. 1-8 (2012)
11. Cooper, R.B., Zmud, R.W.: Information technology implementation research: a technological diffusion approach. Management Science 36, 123-139 (1990)
12. Gallivan, M.J.: Organizational adoption and assimilation of complex technological innovations: development and application of a new framework. ACM Sigmis Database 32, 51-85 (2001)
13. Cárdenas, A.A., Manadhata, P.K., Rajan, S.P.: Big Data Analytics for Security. IEEE Security & Privacy 11, 74-76 (2013)
14. Chen, H., Chiang, R.H., Storey, V.C.: Business intelligence and analytics: From big data to big impact. MIS Quarterly 36, 1165-1188 (2012)
15. Duan, Y., Cao, G.: Understanding the Impact of Business Analytics on Innovation. Proceedings of the 23rd ECIS, pp. 1-16 (2015)
16. Elgendy, N., Elragal, A.: Big data analytics: a literature review paper. In: Industrial Conference on Data Mining, pp. 214-227. Springer, (2014)
17. Chen, D.Q., Preston, D.S., Swink, M.: How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. Journal of Management Information Systems 32, 4-39 (2015)
18. Alharthi, A., Krotov, V., Bowman, M.: Addressing barriers to big data. Business Horizons 60, 285-292 (2017)
19. Beatty, C.A., Gordon, J.R.: Barriers to the implementation of CAD/CAM systems. MIT Sloan Management Review 29, 1-25 (1988)

20. Meyer, A.D., Goes, J.B.: Organizational assimilation of innovations: A multilevel contextual analysis. Academy of management journal 31, 897-923 (1988)
21. Purvis, R.L., Sambamurthy, V., Zmud, R.W.: The assimilation of knowledge platforms in organizations: An empirical investigation. Organization science 12, 117-135 (2001)
22. Zhu, K., Kraemer, K.L., Xu, S.: The process of innovation assimilation by firms in different countries: a technology diffusion perspective on e-business. Management science 52, 1557-1576 (2006)
23. Attewell, P.: Technology diffusion and organizational learning: The case of business computing. Organization Science 3, 1-19 (1992)
24. Saga, V.L., Zmud, R.W.: The nature and determinants of IT acceptance, routinization, and infusion. Levine, L. (ed.) Diffusion, Transfer and Implementation of Information Technology, pp. 67-86. Elsevier Science Inc., North-Holland, Amsterdam (1993)
25. Rogers, E.: Diffusion of innovations. Free Press, New York (1983)
26. Okoli, C., Pawlowski, S.D.: The Delphi method as a research tool: an example, design considerations and applications. Information & Management 42, 15-29 (2004)
27. Paré, G., Cameron, A.-F., Poba-Nzaou, P., Templier, M.: A systematic assessment of rigor in information systems ranking-type Delphi studies. Information & Management 50, 207-217 (2013)
28. Schmidt, R.C.: Managing Delphi Surveys Using Nonparametric Statistical Techniques. Decision Sciences 28, 763-774 (1997)
29. Skinner, R., Nelson, R.R., Chin, W.W., Land, L.: The Delphi method research strategy in studies of information systems. Communications of the Association for Information Systems 37, 31-63 (2015)
30. Brinch, M., Stentoft, J., Jensen, J.K.: Big Data and its Applications in Supply Chain Management: Findings from a Delphi Study. In: Proceedings of the 50th HICSS, pp. 1-10. (2017)
31. Singh, R., KeiI, M., Kasi, V.: Identifying and overcoming the challenges of implementing a project management office. European Journal of Information Systems 18, 409-427 (2009)
32. Piccinini, E., Hanelt, A., Gregory, R., Kolbe, L.: Transforming Industrial Business: The Impact of Digital Transformation on Automotive Organizations. In: Proceedings of the 36th ICIS, pp. 1-20. (2015)
33. Loughlin, K.G., Moore, L.F.: Using Delphi to achieve congruent objectives and activities in a pediatrics department. Academic Medicine 54, 101-106 (1979)
34. Kiron, D., Shockley, R.: Creating business value with analytics. MIT Sloan Management Review 53, 57 (2011)
35. Keil, M., Lee, H.K., Deng, T.: Understanding the most critical skills for managing IT projects: A Delphi study of IT project managers. Information & Management 50, 398-414 (2013)

Deskriptive Analyse von Kennzahlenrelationen

Daniel M. A. Burdensky[1,2], Benny Kneissl[1], Rainer Alt[2]

1 BMW Group, Leipzig München, Deutschland
{daniel.burdensky, benny.kneissl}@bmw.de
2 Universität Leipzig, Professur für Anwendungssysteme, Leipzig, Deutschland
{db59vaqy, rainer.alt}@uni-leipzig.de

Abstract. Die Steuerung großer Produktionsunternehmen stellt aufgrund einer Vielzahl interdependenter Geschäftsprozesse eine Herausforderung für das Management dar. Es wird erwartet, dass Entscheidungen entsprechend der übergeordneten strategischen Zielstellungen der Unternehmen getroffen werden. Jedoch macht die Komplexität der Ursache-Wirkbeziehungen zwischen den Kennzahlen, die zur Prozesssteuerung dienen, eine Abschätzung der Auswirkungen von Maßnahmen für den Menschen nahezu unmöglich. Vor diesem Hintergrund wird ein Ansatz vorgestellt, der auf Basis von Korrelationen zwischen Kennzahlen sowie deren leicht verständlicher Visualisierung die datenbasierte und systemgestützte deskriptive Analyse von Kennzahlenrelationen ermöglicht. Dabei liegt ein besonderer Fokus auf der Objektivität und Praktikabilität des Modells. Dieses wird anhand des Beispiels eines deutschen Automobilwerkes vorgestellt und validiert.

Keywords: Ursache-Wirkbeziehung, KPI, Kennzahl, Data Mining, Korrelation.

1 Einleitung

Zahlreiche Produktionsunternehmen nutzen Business Intelligence (BI)-Systeme zur Entscheidungsunterstützung für das Management. Diese Anwendungen bilden Kennzahlendaten, welche die jeweiligen Geschäftsprozesse repräsentieren, bspw. durch Ampeldarstellungen ab und ermöglichen Analysen und die Überwachung der Zielerreichung. Die Bedeutung dessen liegt darin begründet, dass viele Unternehmen unter Wissensdefiziten der Entscheidungsträger, welche die strategische Willensbildung prägen [1], leiden. Diese Defizite sind eine Konsequenz der hohen Steuerungskomplexität großer Produktionsunternehmen. Komplexität wird als die „…Eigenschaft, viele Zustände oder Verhaltensweisen annehmen zu können" definiert [2, S. 6]. Zudem ist der Mensch in der Lage, ein System aus bis zu vier Variablen kognitiv zu verarbeiten und zu verstehen [3]. Angesichts der Vielzahl an Variablen (Kennzahlen), die beispielsweise ein Automobilwerk besitzt und die zahlreiche Zustände einnehmen können, lässt sich schlussfolgern, dass dessen Verhalten aufgrund seiner Komplexität intuitiv weder nachvollzogen noch antizipiert werden kann [4]. Somit werden die kognitiven Fähigkeiten der Führungskräfte überfordert und das Abschätzen der Konsequenzen von Entscheidungen massiv erschwert [2], [5].

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Üblicherweise treffen Manager auf Grundlage von Fachwissen und Intuition Annahmen bezüglich einzelner Bereiche zur Abschätzung der Auswirkungen von Maßnahmen [4]. Dabei helfen Kennzahlen, die Steuerungskomplexität zu reduzieren. Dennoch ist es sehr schwierig, die Ursachen für bestimmte Wirkungen im Kennzahlennetzwerk zu ermitteln und zielorientiert zu nutzen. Daher besteht eine große Nachfrage nach systembasierter Entscheidungsunterstützung zur Analyse der Kennzahlenrelationen [6], [7].

Vor diesem Hintergrund eignet sich das Automobilwerk Leipzig der BMW Group[1] aufgrund der Komplexität und Interdependenzen der zur Automobilfertigung notwendigen Prozesse, der Steuerung mit Hilfe von Kennzahlen sowie der hierarchischen Zielstruktur für die Durchführung einer Fallstudie. Anhand dessen Geschäftsprozessen und Daten wird im realen Umfeld ein quantitatives Modell[2] zur Analyse von Kennzahlenrelationen entwickelt, angewendet und validiert (Abb. 1).

Abbildung 1. Modell zur datenbasierten Analyse von Kennzahlenrelationen

Das Modell besteht aus fünf Kernelementen: Die Basis bildet die Analyse des Nutzens der Kenntnis von Ursache-Wirkbeziehungen (UWB). Dem folgt auf Grundlage von Kennzahlendaten eine deskriptive Analyse der UWB mittels einer Korrelationsanalyse und einer geeigneten Visualisierung. Mit Hilfe einer multiplen Regression wird im nächsten Schritt eine prädiktive Analyse durchgeführt, um detaillierte Untersuchungen der Wirkungen von Einfluss- auf Zielgrößen zu ermöglichen sowie „What-if"-Betrachtungen durchzuführen. Zur Überprüfung möglicher Kausalitäten der gefundenen Beziehungen folgt die Analyse von Moderations- und Mediationseffekten zwischen den Kennzahlen. Schließlich werden für jeden dieser Schritte die Potentiale hinsichtlich der datenbasierten Erzeugung von Transparenz der Kennzahlenrelationen untersucht.

Der vorliegende Beitrag fokussiert die Nutzenanalyse der Kenntnis von UWB zwischen Kennzahlen, die Analyse von Kennzahlenrelationen mittels einer Korrelationsanalyse sowie dessen Potentiale hinsichtlich der Schaffung von Transparenz.

---

[1] Im BMW Werk Leipzig werden die Fahrzeuge der BMW 1er und 2er Reihe sowie Autos mit Elektroantrieb gefertigt. Rund 5.300 Mitarbeiter produzieren täglich über 860 Fahrzeuge [8]

[2] Ein Modell wird im Entscheidungsunterstützungskontext als eine Verknüpfung mehrerer Methoden verstanden [9]. „Als implementierte Methoden kommen hierbei vor allem heuristische, statistische, finanzmathematische und prognostische Verfahren zum Einsatz" [9, S. 111]

2 Kennzahlenbasierte Unternehmenssteuerung

Die angesprochenen Geschäftsprozesse werden als "Menge von Aufgaben, die in einer vorgegebenen Ablauffolge zu erledigen sind und durch Applikationen der Informationstechnik unterstützt werden," [10, S. 62] verstanden. Eine der Kernaufgaben des Managements ist deren Gestaltung, was „sowohl den (strategischen) Entwurf bzw. die Neugestaltung als auch die ständige (operative) Weiterentwicklung in Form von Prozessverbesserungen, kurz alle Maßnahmen zur Planung, Steuerung und Kontrolle von Geschäftsprozessen" umfasst [11, S. 55]. Dabei ist es von Bedeutung, dass die Bestandteile des Produktionssystems in ihrer Vernetztheit untersucht und gesteuert werden, da komplexe, dynamische, wechselseitige Interdependenzen zwischen den Prozessen bestehen, welche das Verhalten des Ganzen bestimmen [12]. Folglich liefert die isolierte Betrachtung einzelner Prozesse einen unzureichenden Grad an Erklärbarkeit dessen Verhaltens und schöpft Verbesserungspotentiale für die Steuerung nicht aus. Eine Form, die Vernetztheit von Geschäftsprozessen zu betrachten, ist die Anwendung (und Aggregation) von Kennzahlen zu deren Steuerung.

2.1 Steuerung mittels Kennzahlen

Kennzahlen sind „…quantitative Informationen, die für die spezifischen Bedürfnisse der Unternehmensanalyse und –steuerung aufbereitet worden sind" [1, S. 9]. Deren Verwendung ermöglicht es, Zusammenhänge in hoch verdichteter Form transparent zu machen, die Steuerungskomplexität zu reduzieren und Ursachen bestimmter Ereignisse zu analysieren [1], [13]. Als Grundlage dient die Übersetzung der strategischen Unternehmensziele in Kennzahlen und die Definition von Zielwerten. Anschließend werden diese top-down für die Hierarchieebenen heruntergebrochen [14], [15], sodass letztere damit sowohl steuern als auch gesteuert werden. Berichtet werden die Kennzahlen üblicherweise bottom-up.

Indes muss den Verantwortlichen bewusst sein, dass das Kennzahlensystem einer Organisation das Verhalten von Managern und Angestellten maßgeblich beeinflusst [14]: Es wird nach der Erreichung der Zielvorgaben gestrebt, da dadurch der Erfolg der eigenen Arbeit gemessen wird und bspw. Bonuszahlungen daran geknüpft sein können. Daher müssen die Ziele jeder Hierarchieebene auf die Strategie ausgerichtet sein, um eine Fehlsteuerung zu vermeiden. Jedoch erfolgen die Erzeugung und Anpassung von Kennzahlen häufig in einem evolutionären Prozess induktiv aus Erfahrungswissen oder dem Bauchgefühl des Managements und nicht anhand von sachgerechten Kriterien [1], [16]. Zudem fügen wechselnde Manager in der betrieblichen Praxis neue Kennzahlen zu der bestehenden Kennzahlenlandschaft hinzu, sodass deren Menge ansteigt und folglich die Übersichtlichkeit des ohnehin komplexen Systems abnimmt. Die entstehende Intransparenz und die Vernetztheit der Geschäftsprozesse führen durch unvollständige Informationen, begrenzte Ressourcen und Kommunikationsbarrieren zu erhöhtem Aufwand. Daher ist es für eine effiziente Unternehmenssteuerung notwendig, ein bestmögliches Verständnis für die Kennzahlenrelationen zu entwickeln.

2.2 Ursache-Wirkbeziehungen zwischen Kennzahlen

Im Rahmen einer Studie zeigt [17] auf, dass Personen mit einem besseren Verständnis für die kausalen Beziehungen zwischen ihren Handlungen und den jeweiligen Ergebnissen bessere strategische Entscheidungen treffen als andere. Ein Grund dafür ist, dass die Kenntnis von UWB zwischen Kennzahlen das Aufbrechen von Silodenken fördert: Sind die Auswirkungen von Maßnahmen nicht nur im eigenen Verantwortungsbereich sondern im gesamten Kennzahlennetzwerk abschätzbar, können effiziente und am Gesamtziel des Unternehmens ausgerichtete Entscheidungen getroffen werden; dies gilt ebenfalls bereits bei der Herleitung von Kennzahlen und deren Zieldefinition. Somit werden Unsicherheiten und Schäden reduziert [1], [18] und die isolierte Betrachtung von Kennzahlen überwunden. Ferner ermöglicht die übergreifende Kenntnis der UWB von Kennzahlen durch gesteigerte Transparenz bezüglich interdependenter Bereiche des Unternehmens eine effektivere Kommunikation, Vorhersagen und kontinuierliches Lernen über das Systemverhalten [19]. Folglich kann geschlossen werden, ob die Resultate von Maßnahmen mit hoher Wahrscheinlichkeit, notwendigerweise oder nicht eintreten werden [19].

Status quo. Trotz der genannten Vorteile sind „…die Bemühungen, Ursache-Wirkbeziehungen empirisch-theoretisch zu fundieren, […] eher sporadisch und vielfach unzureichend“ [1, S. 404]. [16] zeigen auf, dass empirische Ansätze zur Identifikation, Quantifizierung und Projektion existierender UWB zwischen Kennzahlen fehlen. Gleiches stellt [18] im Kontext der Balanced Scorecard (BSC) fest. [20] legt zudem dar, dass die betriebliche Wirklichkeit und damit Entscheidungsprozesse inklusive ihrer UWB kaum umfassend statistisch beschreibbar sind, da die Realität zu komplex ist. Dem gegenüber konstatiert [21], dass UWB Teil der empirischen Welt und daher empirisch abbildbar sind. Dieser Diskurs stützt die Einschätzung von [19], dass im Bereich der empirischen quantitativen Analyse von UWB Forschungsbedarf besteht.

Verwandte Arbeiten. Durch zahlreiche Interviews mit Experten unterschiedlicher hierarchischer Ebenen (Fachbereichsmitarbeiter bis hin zu Führungskräften der Werkleitung) sowie diverser Fachbereiche (u.a. Qualität, IT, Logistik) und die Mitarbeit in einem strategischen Projekt zur Schaffung von Transparenz, Messbarkeit und Wirksamkeit der Kennzahlenlandschaft des BMW Werkes Leipzig über einen Zeitraum von mehr als einem Jahr konnten zahlreiche Erkenntnisse gewonnen werden: Beispielsweise kumulieren und aggregieren die meisten betrachteten BI-Tools Kennzahlen, ermöglichen aber keine detaillierten Analysen der UWB. Darauf aufbauend wurden zentrale Anforderungen an die Analyse von UWB (siehe Kopfzeile Tab. 1) abgeleitet, welchen bestehende Ansätze zugeordnet werden.

Die große Menge an Kennzahlen in Produktionsunternehmen macht den Umgang mit vielen Variablen und deren Visualisierung durch Systeme zur Analyse von UWB notwendig. Zugleich besitzen Kennzahlen, welche diverse Prozesse steuern, heterogene Einheiten (z. B. Stück, Zeiteinheit). Diese Anforderungen stellen mehrheitlich kein Problem dar.

Tabelle 1. Literaturanalyse anhand ermittelter Anforderungen

Quelle	*Ermittlung Relationen*	*Analyse Relationsstärke*	*Relevanz Einheiten*	*Visualisierung*	*Vielzahl Variablen*
Forrester[4] - System Dynamics	Qualitativ	Qualitativ	Ja	Intuitiv	Ja
Youngblood, Collins [22] - MAUT	Qualitativ	Qualitativ	Nein	Keine	Ja
Rodriguez et al. [16] - QRPMS	Qualitativ	Datenbasis	Nein	Intuitiv	Ja
Van der Aalst [6] - Process Mining	Datenbasis	Datenbasis	Ja	Intuitiv	Nein
Peral et al. [7]	Qualitativ	Datenbasis	Nein	Keine	Ja
Kotzanikolaou et al. [23]	Qualitativ	Qualitativ	Nein	Intuitiv	Ja
Rupprecht/Schweinberger [24]	Qualitativ	Keine	Nein	Keine	Nein
Hesse et al. [25] - PlantCockpit	Qualitativ	Datenbasis	Nein	Intuitiv	Ja

Forschungsbedarf hingegen ergibt sich hinsichtlich der Frage der Ermittlung von Kennzahlenrelationen in bisherigen Arbeiten: Bis auf eine Ausnahme werden UWB-Netze mit der Hilfe von Experten qualitativ entwickelt. Dies führt zu einer reduzierten Objektivität und Reproduzierbarkeit der Modelle, da sich diese in Abhängigkeit der beteiligten Personen ändern. Ebenso können getätigte Annahmen als notwendige Bedingungen der jeweiligen Analyse nicht überprüft werden. So entwickeln bspw. [4] und [22] auf Basis von Erfahrungswissen mathematische Formeln zur Beschreibung des Verhaltens von Kennzahlen, um UWB zu simulieren bzw. mittels Korrelationen zu analysieren.

Zwar ermitteln einige Ansätze die Stärke der Relationen datenbasiert, jedoch weisen sie hinsichtlich der Praktikabilität im Kennzahlenkontext Schwächen auf. Beispielsweise rekonstruiert [6] Prozesse anhand der Daten durch den Prozess geflossener Einheiten. Jedoch basiert der Ansatz aufgrund seines Prozessfokus auf Ereignisprotokollen mit den zugehörigen Zeitpunkten, weshalb er sich nicht zur Analyse von Kennzahlenrelationen eignet. Für letzteres nutzen [16] eine Hauptkomponentenanalyse (PCA). Diese „…unterstellt, dass sich die Varianzen der Items einer Itembatterie möglichst vollständig durch einzelne Faktoren abbilden lassen…[,was] aus Sicht der empirischen Forschung wohl kaum realisierbar ist“ [20, S. 221].

Eine gute Grundlage bieten [25], die die UWB zwar qualitativ ermitteln, daraufhin aber ebenfalls paarweise Korrelationsanalysen zwischen den Kennzahlen durchführen. Deren Darstellung erfolgt jedoch nur hinsichtlich einzelner Größen, sodass die datenbasierten Beziehungen des Gesamtsystems nicht abgebildet werden. Somit werden u.a. die initial ermittelten UWB als gegeben angesehen und nicht hinterfragt, was der Objektivität nicht zuträglich ist, und die Prüfung redundanter Kennzahlen erschwert.

3 Analyse der Kennzahlenrelationen eines Automobilwerks

Data Mining beschreibt die Anwendung spezieller Algorithmen zur Identifikation von Mustern in Daten, welche nicht-triviale Vorhersagen bezüglich neuer Daten ermöglichen [26]. Etwas konkreter führt [1, S. 33] aus, dass „…z.B. unter Einsatz statistischer Verfahren nach Clustern oder Korrelationen zwischen den Daten…“ gesucht wird und es der Entdeckung neuer Zusammenhänge und Strukturen dient. Zwei der primären Ziele sind die Deskription und die Prädiktion: Deskription ist die Identifikation von für Menschen verständlichen Mustern, die die Daten beschreiben; Prädiktion nutzt Variablen aus Datenmengen, um unbekannte oder zukünftige Werte anderer Variablen vorherzusagen [27]. Nachstehend wird, wie erwähnt, der deskriptive Teil des entwickelten Modells (Abb. 1) betrachtet.

Entscheidungsträger in großen Unternehmen sind häufig keine Analytikexperten, weshalb die Ratschläge von Analysten für diese schwer zu verstehen und das Vertrauen in die Ergebnisse und daher deren Berücksichtigung bei Entscheidungen unzureichend sein können [26]. „Nur die […] verstandenen und damit auch akzeptierten Resultate können letztlich Entscheidungen und die zukünftige Realität beeinflussen“ [20]. Insofern eignet sich eine Korrelationsanalyse zur Deskription, da sie eine geläufige, akzeptierte und vertrauenswürdige statistische Methode ist.

3.1 Beschreibung des Fallbeispiels

Die Abbildung aller UWB in einem Unternehmen ist [20] zufolge aufgrund der Vielzahl der Prozesse und Kennzahlen zu komplex, was im Rahmen der Fallstudie bestätigt wurde. Daher wurden spezifische Entscheidungsprobleme separat untersucht. Bei deren Abgrenzung zeigte sich, dass sich die Experten häufig intuitiv nach Funktionsbereichen oder Verantwortlichkeiten richteten. Um jedoch ein umfassendes Verständnis der UWB zu erhalten, bewährte sich die Orientierung an potentiellen prozessinternen Einflussgebern ergänzt um Größen aus Vorgänger- / Nachfolgerprozessen, Parallelprozessen oder Zielenehmer-Zielegeber-Beziehungen.

Beispielhaft wurde die Kennzahl Termintreue im Finish-Bereich der Montage des Werks ausgewählt. Sie misst die Quote der Fahrzeuge, die innerhalb der veranschlagten Zeiträume die Montage verlassen, und repräsentiert einen kritischen Erfolgsfaktor für die Werkleitung. Daher ist die Kenntnis, welche Größen sowohl positiven als auch negativen Einfluss auf diese Ergebnisgröße haben, von großer Bedeutung für das Management. Als einflussgebende Größen wurden zunächst weitere Kennzahlen, die vom strategischen Management genutzt werden, ausgewählt und diese im Austausch mit Prozessexperten um operative Größen ergänzt, sodass die Analyse der UWB mit insgesamt zehn Kennzahlen erfolgte. Diese werden im Produktionsprozess parallel oder sequentiell erfasst und im Folgenden nummeriert dargestellt: K-1 bis K-10.

Die Daten, die für diese Größen verwendet wurden, sind die Kennzahlen-Werte für die Serienproduktion von Fahrzeugen mit Verbrennungsmotor auf Tagesbasis von Anfang Januar 2016 bis Ende April 2017. Bei der Auswahl des Zeitraumes wurde berücksichtigt, dass dieser nicht zu lang ist, um Auswirkungen einschneidender Veränderungen der produzierten Fahrzeuge (z. B. Facelifts), Anlaufphasen neuer

Fahrzeuge oder Veränderungen an den Produktionslinien zu vermeiden, da der Regelbetrieb des Werkes ohne Sonderereignisse modelliert werden soll. Andererseits sind zur Erprobung ausreichend Informationen vorhanden (>377 Beobachtungen je Kennzahl), um die UWB realistisch abbilden zu können. Für die Analyse wurde das Package Statsmodels von Python genutzt und das resultierende Netzwerk mittels Neo4J visualisiert.

3.2 Deskription der Kennzahlenrelationen

Mit dem Ziel, die Kennzahlenrelationen derart abzubilden, dass Kennzahlenverantwortliche die Darstellung innerhalb kürzester Zeit verstehen, wurde eine paarweise Korrelationsanalyse nach Bravais-Pearson durchgeführt. Diese ermöglicht die Bestimmung der Stärke des linearen Zusammenhangs zweier Variablen [20]. Zunächst wurde geprüft, ob die Kennzahlen definitorisch unabhängig sind, sodass Korrelationen aufgrund formeller Abhängigkeiten ausgeschlossen sind, und, ob sie logisch und prozessual in Zusammenhang stehen, sodass Kausalbeziehungen möglich sind. Weiterhin sind die Rohdaten der Kennzahlen metrisch skaliert, was ebenfalls für die zulässige Durchführung der Korrelationsanalyse spricht. Unter der Prämisse, die Daten geringstmöglich zu verändern, um deren Objektivität zu erhalten, wurden Schritte zur Datenbereinigung durchgeführt. Fehlende (z.B. NaN) und fachlich unplausible Werte wurden entfernt, wobei ein permanenter Austausch mit Prozessexperten stattfand, um keine für den Prozess charakteristischen Ausprägungen zu beeinflussen: Es besteht prinzipiell die Möglichkeit, dass Anomalien (Ausreißer), welche aus statistischer Sicht vermutlich nicht den Regelbetrieb widerspiegeln, trotzdem in der Analyse betrachtet werden sollen, da Experten diese als normale Prozesserscheinungen identifizieren. Außerdem wurden zum Erhalt der Objektivität der Daten keine Aggregationen oder ähnliches durchgeführt.

Durch das Sichten der resultierenden Streudiagramme aller untersuchter paarweiser Korrelationsbeziehungen wurde sichergestellt, dass die Daten linear beschrieben werden können [28]. Dies verdeutlicht Abbildung 2 beispielhaft anhand der Kennzahlen Termintreue (K-10; x-Achse) und Durchlaufzeit (K-9; y-Achse).

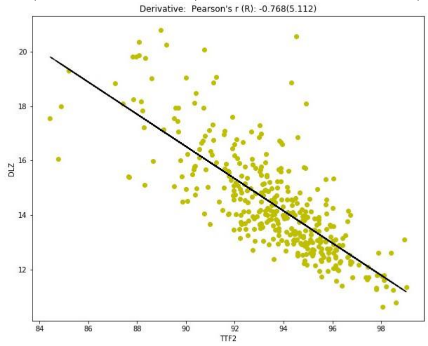


Abbildung 2. Korrelationsanalyse zwischen Termintreue und Durchlaufzeit

Der errechnete Korrelationskoeffizient sagt aus, ob zwischen den Messwertpaaren ein positiver (perfekt positiv: r=1), ein negativer (perfekt negativ: r=-1) oder kein linearer Zusammenhang (Koeffizient - nahe bei - null) besteht [20]. Um ein differenzierteres Maß zur Klassifizierung der Koeffizienten zu erhalten, wurden die in der Literatur üblichen Grenzen[3] angepasst und zur folgenden Graphanalyse Farben zugeordnet (siehe Tab. 2):

Tabelle 2. Farbliche Codierung der Kanten[4]

$\|r\|$	*Korrelation*	*Farbe*
<0,1	keine	keine
0,1<|r|<0,4	Schwach	Rot
0,4<|r|<0,7	Mittel	Gelb
0,7<|r|<0,9	Stark	Grün
|r|>0,9	Sehr stark	Blau

Die Kennzahlen Termintreue und Durchlaufzeit, welche annähernd normalverteilt sind, weisen einen Korrelationsfaktor von -0,768 auf, was bedeutet, dass sie stark antikorreliert sind. Dies ist prozessual schlüssig, da, wenn die Durchlaufzeit je Fahrzeug sinkt, eine Verbesserung der Termintreue sehr wahrscheinlich ist. So wurde durch die Korrelationsanalyse eine plausible Beziehung erkannt. Mittels des P-Wertes kann zudem die Signifikanz aufgezeigt werden[5]. Er gibt an, mit welcher Wahrscheinlichkeit ein Korrelationskoeffizient gleicher Größe entstehen könnte, wenn die Nullhypothese (es existiert keine Beziehung zwischen den Variablen) wahr wäre [28]. In der Literatur wird häufig eine Signifikanzgrenze von 0,05 genutzt, welche sehr deutlich unterschritten wird (P-Wert: 1,23E-72<0,05).[6] Im Anschluss an die Berechnung der Korrelationen erfolgt die Visualisierung des Kennzahlennetzwerks in einem Graphen (siehe Abb. 3 links).

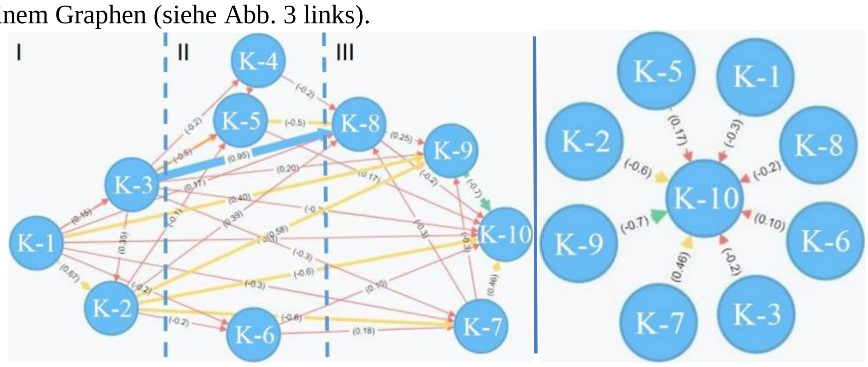


Abbildung 3. Visualisierung der Korrelationen in Beziehungsgraphen

---

[3] |r|<0,5 – schwache Korrelation; 0,5<|r|<0,8 - mittlere Korr.;|r|>0,8 - starke Korr. [18]

[4] Per Konvention gehört die Klassenobergrenze zur jeweiligen Klasse, die Untergrenze nicht [18]

[5] Optional, da auch nicht-signifikante Korrelationen kausal relevant sein können

[6] Bei starken Abweichungen von der Normalverteilung bzw. Nichtlinearität sind andere Korrelationskoeffizienten zu prüfen (bspw. Spearmans Rangkorrelationskoeffizient)

In diesem Graphen repräsentieren die Knoten die jeweiligen Kennzahlen und die Kanten die bestehenden Korrelationsbeziehungen inklusive der Stärke der Beziehungen anhand deren Breite, Färbung und Koeffizienten. Die Pfeile an den Kanten sind zu ignorieren, da mittels einer Korrelation die Wirkrichtung, ob A auf B oder B auf A wirkt, nicht bestimmt werden kann [28]. Dies ist der nicht anpassbaren Darstellungsform von Neo4J geschuldet und beinhaltet keine Aussage. Davon abgesehen eignet sich Neo4J jedoch, um solche komplexen Graphen zu erzeugen.

Die horizontale Sortierung der Kennzahlen erfolgte gemäß der realen prozessualen Sequenz im Werk, welcher auch die Analyse der UWB folgen sollte. Die römischen Zahlen kennzeichnen die drei Erfassungspunkte dieses Beispiels. Mittels einer solchen Darstellung wird die datenbasierte Analyse, welche Kennzahlen statistisch miteinander in Beziehung stehen und wie stark die jeweiligen Korrelationen sind, ermöglicht. Es zeigt sich, dass durch die Vielzahl an Korrelationen zwischen den Kennzahlen die eingangs angenommenen Interdependenzen der Geschäftsprozesse und die Komplexität, der sich das Management gegenübersieht, nachvollzogen werden können.

Des Weiteren können auch einzelne Elemente untersucht werden: Da die Analyse der Termintreue im Fokus steht, werden die Beziehungen hinsichtlich dieser in Abbildung 3 rechts alleinstehend dargestellt. Neben der erläuterten Beziehung mit der Durchlaufzeit weisen weitere sieben der ausgewählten Kennzahlen eine Korrelation mit der Termintreue auf: Vier davon sind schwacher und zwei mittlerer Ausprägung. Darauf aufbauend kann bspw. hinterfragt werden, welche dieser Größen sich zukünftig für die effiziente Steuerung der Termintreue eignen und ob die Zielvorgaben dieser Kennzahlen der Zielerreichung der Termintreue zuträglich sind oder nicht.

Bei der Analyse der Ergebnisse ist zu beachten, dass eine Korrelation keine Aussage bezüglich kausaler Beziehungen trifft, da solche Muster in Daten zufällig entstehen können (sogenannte Scheinkorrelationen). Weiterhin liegt keine statistische Unabhängigkeit zwischen den Größen vor. Dem wird in den Schritten 2 und 3 des Modells (Abb. 1) Rechnung getragen werden, da dieser Beitrag den deskriptiven Teil fokussiert. Auch können einzelne Prozesse mehrere Tage benötigen, weshalb durch die Betrachtung von Tageswerten ohne Berücksichtigung zeitlichen Versatzes ein geringes Rauschen in den Daten entsteht. Aus den genannten Gründen wurden die Ergebnisse mit Kennzahlenexperten auf Plausibilität und mögliche Kausalitäten überprüft und die erkannten Beziehungen bestätigt. Somit konnte dieses Modellelement zur deskriptiven Analyse von Kennzahlenrelationen sowie deren Visualisierung validiert werden.

4 Schlussbetrachtungen

Im BI-Umfeld wird häufig davon ausgegangen, dass Kennzahlen und deren UWB qualitativ beispielsweise im Kontext der BSC [9] betrachtet oder aus dem Business Model Canvas abgeleitet werden [29]. Basis dessen sind i. d. R. Annahmen des Managements, sodass das einbrachte Wissen von den beteiligten Personen abhängt. Der vorliegende Beitrag stellt den deskriptiven Teil eines neuen Analyseansatzes als ergänzende, objektivere Informationsquelle zur Entscheidungsunterstützung vor: Die Beziehungen zwischen Kennzahlen werden mittels empirischer Daten quantitativ ermittelt, was die Reproduktion der Ergebnisse unabhängig von einzelnen Personen ermöglicht und implizites Wissen datenbasiert ergänzt.

Da außer der Auswahl der Größen und deren prozessualer Sortierung keine qualitativen Informationen in die Methode einfließen, ist die Komplexität der vorliegenden Analyse gegenüber anderen quantitativen Ansätzen tendenziell geringer. In der Praxis zeigte sich, dass Korrelationsanalysen den Kennzahlenexperten zumindest terminologisch geläufig waren, wodurch das Vertrauen in die Ergebnisse anstieg. Insbesondere durch die Darstellung beispielhafter Streudiagramme mit unterschiedlich starken Korrelationen, konnte die Aussagekraft des Korrelationsgraphen verdeutlicht und eine sehr gute Nachvollziehbarkeit erreicht werden. Dies ist in den in Tabelle 1 genannten Quellen nur sehr eingeschränkt oder gar nicht möglich. Die Interpretation der Visualisierung der Beziehungen in dem Korrelationsgraphen war für alle Anwender nach wenigen Minuten problemlos möglich.

Darauf aufbauend können zusätzliche Erkenntnisse über untersuchte Teilbereiche komplexer Kennzahlensysteme erlangt werden, um deren Verhalten besser zu verstehen. Eine starke Korrelation zwischen zwei Größen kann z. B. die Wirksamkeit einer angenommenen UWB bestätigen oder ggf. auf Redundanz einer der beiden Kennzahlen aufmerksam machen. Eine fehlende oder geringe Korrelation könnte indes auf eine mangelhafte Einflussnahme einer Größe auf eine andere hindeuten.

Außerdem ist es möglich, unbekannte Beziehungen zwischen Kennzahlen, die bisher nicht bei der Steuerung des Unternehmens berücksichtigt wurden, zu identifizieren: Zeigen zwei Kennzahlen eine hinreichend starke Korrelation, sollte diese Beziehung hinsichtlich ihrer Eignung zur zukünftigen Steuerung untersucht werden. Ferner ist die Analyse großer Datenmengen möglich, sodass durch das versuchsweise Hinzufügen weiterer Kennzahlen möglicherweise neue Erkenntnisse bezüglich bereichs- und hierarchieübergreifende UWB gewonnen werden. So können auch diverse Zeiträume untersucht werden, um bspw. saisonal bedingte UWB zu erkennen.

Dies bietet Potentiale für die Schaffung von mehr Transparenz und die Verbesserung des Verständnisses für die Vernetztheit von UWB. Die folgende Reduktion von Silodenken im Unternehmen dient der Ausrichtung von Maßnahmen und Zielen auf die Unternehmensstrategie. Auch können Lücken im Kennzahlennetzwerk erkannt, neue Kennzahlen zielgerichtet ergänzt und das dynamische Verhalten bestehender Kennzahlen abgeschätzt werden. Insbesondere letzteres vermindert den statischen Charakter bestehender BI-Lösungen. Somit wird entscheidend zur Komplexitätsbewältigung durch effizientere Kommunikation verschiedener Bereiche beigetragen und durch das Abwägen von Handlungsalternativen Unsicherheiten reduziert.

Zusammenfassend stellt dieser in der Praxis validierte Beitrag eine Möglichkeit zur datenbasierten und systemgestützten, deskriptiven Analyse von Kennzahlenrelationen dar, dessen Ergebnisse für Personen ohne Analytikkenntnisse verständlich und anwendbar sind. Zu berücksichtigen ist, dass das dargestellte Kennzahlennetzwerk auf paarweisen Korrelationen basiert, wodurch kombinierte UWB nicht abgebildet und „What-if"-Analysen nicht ermöglicht werden. Eine Lösung dafür wird im zweiten Schritt des entwickelten Modells mittels einer Multiplen Regression realisiert werden.

Außerdem stammen die genutzten Daten aus nur einem Produktionsunternehmen. Da aber keine Spezifika dessen in den Beitrag eingeflossen sind, wird die

Übertragbarkeit auf andere Unternehmen und Prozesse angenommen. Dies wird dadurch gestützt, dass auch aus operativen Daten und Fragestellungen bereits wertvolle Erkenntnisse erlangt wurden.

Auch wenn ein hohes Maß an Objektivität der Analyseergebnisse erreicht wird, ist durch die notwendige Datenbereinigung und die Auswahl der eingebrachten Daten ein subjektiver Einfluss unumgänglich. Gleichermaßen unterliegt die Analyse der Resultate dem individuellen Wissensstand des Anwenders. Dabei ist wichtig, dass sie hinsichtlich möglicher Kausalitäten geprüft werden, da eine Korrelation keine Aussage diesbezüglich treffen kann; der Anwender trägt die Verantwortung, ob die Ergebnisse zur Entscheidungsunterstützung genutzt werden.

Zukünftig wäre zu prüfen, ob die Methode einer verbesserten mathematischen Beschreibung der Beziehungen von Elementen in SD-Modellen (z. B. [5]) und Ansätzen wie MAUT [22] dient. Auch könnte eine Übertragung auf kennzahlenorientierte Prozessmodellierungen wie Visual PPINOT [30] interessante Erkenntnisse liefern. Darüber hinaus bildet die Methode die Grundlage für die prädiktive Analyse und die statistische Betrachtung von Kausalitäten als nächste Schritte des Modells zur datenbasierten Analyse von Kennzahlenrelationen.

5 Quellen

1. Gladen, W.: Performance Measurement. Controlling mit Kennzahlen. Springer Fachmedien Wiesbaden, Wiesbaden (2014)
2. Schwaninger, M.: Systemtheorie. Eine Einführung für Führungskräfte, Wirtschafts- und Sozialwissenschaftler (2004)
3. Halford, G.S., Baker, R., McCredden, J.E., Bain, J.D.: How many variables can humans process? Psychological science 16, 70–76 (2005)
4. Forrester, J.W.: Some Basic Concepts in System Dynamics. Sloan School of Management MIT, Massachusetts (2009)
5. Sterman, J.D.: System Dynamics Modeling. Tools for Learning in a Complex World. California Management Review 43, 8–25 (2001)
6. Aalst, Wil M. P. van der: Process Mining. Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg (2011)
7. Peral, J., Maté, A., Marco, M.: Application of Data Mining techniques to identify relevant Key Performance Indicators. Computer Standards & Interfaces 50, 55–64 (2017)
8. BMW Group: Kennzahlen im Überblick, https://www.bmwgroup-werke.com/leipzig/de/unser-werk.html
9. Kemper, H.-G., Baars, H., Mehanna, W.: Business Intelligence – Grundlagen und praktische Anwendungen. Eine Einführung in die IT-basierte Managementunterstützung. Vieweg+Teubner Verlag / GWV Fachverlage GmbH Wiesbaden, Wiesbaden (2010)
10. Österle, H.: Entwurfstechniken. Springer, Berlin (1995)
11. Alt, R., Puschmann, T.: Digitalisierung der Finanzindustrie. Grundlagen der Fintech-Evolution (2016)
12. Krallmann, H. (ed.): Systemanalyse im Unternehmen. Prozessorientierte Methoden der Wirtschaftsinformatik. Oldenbourg, München, Wien (2007)
13. Schmalzried, D., Cundius, C., Franke, R., Lambeck, C., Alt, R., Zimmermann, W. and Groh, R.: In-Memory basierte Real-Time Supply Chain Planung. Leipzig, 197–211 (2013)

14. Kaplan, R.S., Norton, D.: The Balanced Scorecard. Measures that Drive Performance. Harvard Business Review 70, 71–79 (1992)
15. Hammer, M., Hershman, L.W.: Faster, cheaper, better. The 9 levers for transforming how work gets done. New York (2010)
16. Rodriguez, R.R., Saiz, J.J.A., Bas, A.O.: Quantitative relationships between key performance indicators for supporting decision-making processes. Computers in Industry 60, 104–113 (2009)
17. Strohhecker, J.: Factors influencing strategy implementation decisions. An evaluation of a balanced scorecard cockpit, intelligence, and knowledge. J Manag Control 27, 89–119 (2016)
18. Hügens, T.: Balanced Scorecard und Ursache-Wirkungsbeziehungen. Kausale Modellierung und Simulation mithilfe von Methoden des Qualitative Reasoning. Gabler Verlag / GWV Fachverlage GmbH Wiesbaden, Wiesbaden (2008)
19. Malina, M., Nørreklit, H., Selto, F.: Relations among Measures, Climate of Control, and Performance Measurement Models. Contemporary Accounting Research 24, 935–982 (2007)
20. Cleff, T.: Deskriptive Statistik und Explorative Datenanalyse. Gabler Verlag, Wiesbaden (2015)
21. Norreklit, H.: The balance on the balanced scorecard a critical analysis of some of its assumptions. Management Accounting Research 11, 65–88 (2000)
22. Youngblood, A.D., Collins, T.R.: Addressing Balanced Scorecard Trade-off Issues Between Performance Metrics Using Multi-Attribute Utility Theory. Engineering Management Journal 15, 11–17 (2015)
23. Kotzanikolaou, P., Theoharidou, M. and Gritzalis, D.: Interdependencies between Critical Infrastructures. Analyzing the Risk of Cascading Effects. Springer, Berlin, Heidelberg, 104–115 (2013)
24. Rupprecht, M., Schweinberger, A.: Kausale Zusammenhänge in der Ökonometrie - Die Granger-Kausalität am Beispiel von Immobilienvermögen und Konsum. WIST 37, 666–668 (2008)
25. Hesse, S., Spehr, M., Gumhold, S., Groh, R.: Visualizing time-dependent key performance indicator in a graph-based analysis. In: Grau, A., Martinez, H. (eds.) ETFA'2014. 19th IEEE International Conference on Emerging Techologies and Factory Automation : September 16-19, 2014 : Barcelona, Spain, pp. 1–7. IEEE, [Piscataway, N.J.] (2014)
26. Witten, I.H., Frank, E., Hall, M.A.: Data mining. Practical machine learning tools and techniques (2011)
27. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery in Databases. AI Magazine 17, 37–54 (1996)
28. Field, A.: Discovering statistics using IBM SPSS statistics. And sex and drugs and rock 'n' roll. Sage, Los Angeles, London, New Delhi (2013)
29. Linden, M., Navrade, F.: Effektive Planung und Steuerung erfolgskritischer Komponenten eines Geschäftsmodells. In: Gluchowski, P., Chamoni, P. (eds.) Analytische Informationssysteme. Business Intelligence-Technologien und -Anwendungen, pp. 325–349. Springer, Berlin, Heidelberg (2015)
30. del-Río-Ortega, A., Resinas, M., Durán, A., Bernárdez, B., Ruiz-Cortés, A., Toro, M.: Visual ppinot. A Graphical Notation for Process Performance Indicators. Bus Inf Syst Eng 5, 28 (2017)

The Potential of User Behavioural Data for a Preventive Exception Handling Mechanism

Niklas Eschner[1] and Paul Sutterer[1]

[1] Technical University of Munich, Department of Informatics, Munich, Germany
niklas.eschner@tum.de, paul.sutterer@tum.de

Abstract. Current exception handling mechanisms in Java and .NET frameworks are limited to handle exceptions once they occur. Non-existing or failing exception handling can cause several problems. This study contributes to the development of a preventive exception handling mechanism where the occurrence of an exception is predicted and proactively prevented. A database of 1.5 million user sessions recorded on a stand-alone C# software application was used in this study. We evaluate the potential of user behavioural data to predict the occurrence of exceptions. Five classification methods are benchmarked with stratified 10-fold-cross validation. K-Nearest Neighbour shows superiority to the other methods with an average Matthews correlation coefficient of 0.66. Complementing hardware and system environmental data by user behavioural data improves the prediction quality of exceptions. Our study provides evidence for the ability to predict exceptions regardless of their type. It is a step towards a self-learning mechanism that improves software reliability post-release.

Keywords: Preventive Exception Handling Mechanism, Software Defect Prediction, Exception Prediction, Software Reliability.

1 Introduction

Software is omnipresent in modern society which highly relies on its proper operation. Especially when software is responsible for human lives or has high impact on business decisions, software is expected to operate reliable and flawless. Sometimes this expectation is dismissed as there is no guarantee for a *software defect* free application. Therefore, intelligent technical mechanisms that increase software's maturity, fault tolerance, and recoverability are desirable.

In contrast to material goods, the intangible nature of software enables the occurrence of software defects [1]. Arora et al. analyse the software companies' trade-off between selling software with defects and investments in patching. They conclude that in large markets, software companies release software early, despite unfinished software testing. Reasons include high time-to market pressure, a fixed cost structure to remedy software defects, and marginal costs that are effectively zero for fixing software defects per product post release [1].

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

To prevent a software from crashing due to defects, the exception handling mechanism was invented in 1975 [2]. This paradigm is reactive, defining actions which are triggered when a defect occurs. These actions aim at bringing the software back into a known state and prevent it from crashing. Cabral and Marques empirically evaluate the mechanism for Java and .NET in [3]. Their research reveals two shortcomings. First, software developers misuse the exception handling mechanism. They either throw generic *exceptions*, leave the exception handling code empty, or use the mechanism for writing log-files. This makes error handling impossible and therefore disables its use as an error handling tool. Second, ever since its invention in 1975, the exception handling mechanism's design has not changed much [3].

The exception handling mechanism is subject to current research which addresses the presented shortcomings. The latest developments in machine learning algorithms and the decreasing costs in computational power enable a paradigm shift. Lourenço et al. propose a preventive exception handling mechanism (PreX) [4,5]. A classifier learning from observed exceptions could make predictions about their occurrence. If an exception is predicted, countermeasures can be initiated to bring the software into a state that prevents it from happening. The best classifier's performance ranges from 3 to 5 percent false positive rate (FPR) and false negative rate (FNR) [4].

Besides PreX, the research field of software defect prediction is close to our study. The majority of the studies in this research field use a classification algorithm to predict whether software modules have defects [6]. Lessmann et al. offer a benchmark framework to compare different classification methods for this task [7]. They propose the area under the receiver operating characteristics curve (AUC) as an evaluation metric for comparative studies in this research field. Open issues in software defect prediction are listed in [8].

This study addresses the two research fields: exception handling mechanism and software defect prediction. In contrast to software defect prediction, our classifier predicts the occurrence of exceptions in a post release live system rather than the occurrence of defects in the software code. While Lourenço et al. [4,5] focus on predicting null pointer exceptions only, our classifier predicts all types of exceptions. Furthermore, we evaluate the ability of PreX with recorded data on a C# stand-alone software application. Besides *hardware and system environmental data*, we include user behavioural data for the prediction of exceptions. We aim to answer three research questions with this study:

Research Question 1 (RQ1): What potential does user behavioural data have for the prediction of exceptions in user sessions? Software defects and exceptions are usually predicted with hardware and system environmental data. However, the user's actions trigger exceptions.

Research Question 2 (RQ2): How important is the beginning of a user session for the occurrence of an exception during the session? The first events in a user session should reveal which task a user pursues. We assume that exceptions are related to certain tasks. Therefore, the beginning of a user session should already reveal a later occurring exception.

Research Question 3 (RQ3): What potential does the modelling of user behaviour have for predicting exceptions in a live system? For the development of PreX, it is not sufficient to answer only RQ2, since the exact moment of an occurring exception is unknown. However, for PreX, it is essential to predict that exact moment.

The remainder of this article is structured as follows. In Chapter 2, we provide required definitions and a description of the applied machine learning and evaluation methods. In Chapter 3, we describe the data set design and its features. In Chapter 4, the machine learning methods are applied to two data sets and their performance is evaluated. We conclude with a summary of this study's contribution in Chapter 5, as well as its limitations, and provide some directions for future research.

2 Theoretical Background

In the following, we differentiate software defects from exceptions and provide the reader with the concept of PreX. Next, we present the chosen classifiers and justify the selected evaluation metrics.

2.1 Software Defects and Exceptions

The term software defect describes unintended software behaviour. It includes *software failures* and *software faults* [9]. A software failure describes the software's inability to produce the user's expected result [10]. The cause for a software failure is a software fault [9]. Only a developer can eliminate a software fault in the code before compiling the application. Some software faults remain forever undiscovered in software.

In order to define the term exception, Flaviu distinguishes three different software states that an executed software method can result in [11]. These are: standard domain (SD), anticipated exceptional domain (AED), and unanticipated exceptional domain (UED). To control a software system, it is necessary to know the state of all its variables and objects. If this condition is violated, the software is in an unintended state and terminates. Usually, a called software method terminates at the SD. When a method is invoked and an exception appears, the application terminates in AED or UED. Therefore, Flaviu defines an exception as unintended software state [11]. Flaviu's study is the foundation for current exception handling mechanisms. We use the term exception to refer to an unintended software state [11].

2.2 Preventive Exception Handling Mechanism

The current exception handling mechanisms in Java and .NET follow a `try-catch-finally` syntax. Lourenço et al. [4,5] adopt and extend this syntax as shown in the following.

```
try (<prediction_context>) {
// ... Prediction Block.
// Exceptions can be caught and alarms can be triggered
} prevent ( <exception_name>, <information_object> ) {
// ... Prevention Block.
// Execution follows the resumption model.
} catch ( ... ) {
... Exception Handling code
}
```

The `try` block turns into a prediction block. A classifier predicts the probability for an upcoming exception. When a certain probability threshold is exceeded, an alarm is triggered and the `prevent` block is executed. This block is responsible for preventing the exception from happening by executing the resumption model. When the classifier fails to predict an exception, no alarm is triggered and the traditional exception handling mechanism in the `catch` block is executed.

Lourenço et al.'s aim is to offer developers a sophisticated and powerful tool to develop and apply new kinds of exception handling strategies. The developer is responsible to define what kind of exception should be predicted and what suitable countermeasures for this specific kind of exception are. Lourenço et al. simulate a python client-server application [4,5]. For this environment, they predict connection timeouts between the client and the server. When the probability for such a timeout is high, they slow down the clients execution rate to prevent the connection timeout[4,5].

While Lourenço et al. focus on predicting connection timeouts in this specific environment, we focus on predicting any kind of exception types in an empirical data set. Since the design of the resumption model always is an individual decision of the developer, we do not address its design further.

2.3 Classification Methods and Evaluation Metric

Table 1 displays the seven most frequently used classification methods in software defect prediction [6] and indicates which of these are applied in this study's benchmark.

Table 1. Application of classification methods in the field of software defect prediction

Classification Method	*Number of Studies [6]*	*Applied in Current Study*
Naïve Bayes	14	Yes
Decision Tree	11	Yes
Neural Network	9	No
Random Forest	6	Yes
Logistic Regression	5	Yes
K-Nearest Neighbour	4	Yes
Support Vector Machine	4	No

We do not apply Neural Networks due to the high resource requirements regarding memory and time. Support Vector Machines are excluded also since they require parameter optimisation which is not the focus of this study.
Lessmann et al.'s framework [7] evaluates the classification methods based on the area under the curve (AUC) metric. It describes the area under the receiver operating characteristics (ROC) curve which is defined as

$$ROC = \frac{True\ Postive\ Rate\ (TPR)}{False\ Positive\ Rate\ (FPR)} \quad (1)$$

The AUC value can be interpreted as the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [12]. In recent years, there has been some criticism about the metric. In particular for imbalanced data, it can give a misleading impression about a model's performance by over-estimating a classifiers performance [13]. Due to these shortcomings, we use Matthews correlation coefficient (MCC) [14] as our main evaluation metric. It is defined as follows

$$MCC = \frac{TP\times TN - \quad \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (2)$$

with TP=True Positive, TN=True Negative, FP=False Positive, and FN=False Negative [15]. Since it involves all values of the confusion matrix and considers accuracies and error rates of both classes, it is less biased by an imbalanced data set and considered the best singular assessment metric by some authors [16–18]. It ranges from -1 to 1, with -1 indicating the worst possible, 1 indicating a perfect, and 0 indicating a random prediction performance [17]. Since the MCC is a contingency matrix method of calculating the Person correlation coefficient [15], MCC values can be interpreted as such. Following Evans, a Pearson correlation coefficient between 0.40 and 0.59 is accounted as "moderate", as "strong" between 0.60 and 0.79, and as "very strong" between 0.80 and 1.0 [19]. For comparison reasons, we report both, MCC and AUC.

3 Data Sets Design

The data originates from a software which is primarily used for product portfolio analysis and variant management in the manufacturing industry. The risk of a software crash only results in a software restart and therefore bad user experience. The number of recorded attributes in many user sessions plus the tracking of exceptions, makes it a qualified software application for our study. A database of 1.5 million recorded user sessions was provided for this study. It covers customers' and developers' sessions of two different software products for a period of 35 months. During this period, the recording of sessions has been adjusted several times. We can only ensure the same level of data quality for sessions recorded in the past 12 months. We derive two data sets to answer the research questions. Both contain only data of customers' sessions and the first software product. For technical reasons, they differ in the number of covered months. The two data sets contain 12 hardware and system environmental features, hereinafter referred to as *system features* that are displayed in Table 2. The system features contain information, which is available at the beginning of a user session.

Table 2. Overview of system features

System Feature Category	*Number of Features*	*Example*
Hardware Environment	6	Processor Architecture
Software Environment	5	Operating System Version
Software Product	1	Software Version

In addition to system features, both data sets contain triggered *session events*. These are individual events in a session triggered by the user. A click on the application's start menu is one example for such an event. For this study, we observe 415 unique session events, which cover a click path through the software application. This represents the user's behaviour in a session. Each session event is represented by a unique identifier. Not all 415 session events are necessarily triggered in one single user session. Figure 1 illustrates an exemplary user session with $n \in \mathbb{N}$ session events and $m \in \mathbb{N}$ exceptions.

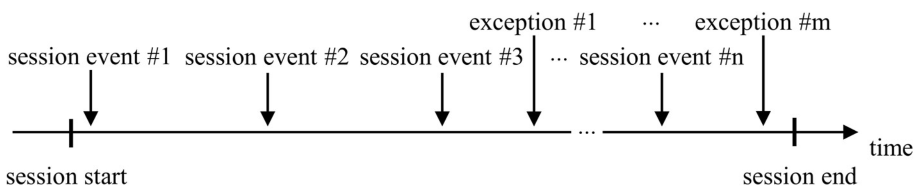


Figure 1. Model of a single user session with session events and exceptions

For both data sets, we model session events as sparse matrices representing the chronological order of events in a user session.

3.1 Modelling of Individual User Behaviour in Data Set I

We design Data Set I to answer RQ1 and RQ2. It covers a time frame of 12 months. A single user session is represented as exactly one instance. The session events only represent sequences of the beginning of a user session with a length of $n \in \{3, 6, 9\}$. For example, assume one single user session consists of the following five session events $(376, 379, 387, 295, 296)$. In case of $n = 3$, only the first three session events (376, 379, 387) are considered.

It is possible to observe more than one exception during a user session. This is irrelevant for the dependent variable, since it only indicates whether an exception will occur at some point during an entire user session.

3.2 Modelling of Individual User Behaviour in Data Set II

We design Data Set II to answer RQ1 and RQ3. It covers a time frame of 3 months. In contrast to Data Set I, a single user session is now represented by several instances. We split all observed session events during one single user session into sequences of length $n \in \{3, 6, 9\}$. For example, assume one single user session consists of the same five session events $(376, 379, 387, 295, 296)$. In case of $n = 3$, three session events sequences $(376, 379, 387)$, $(379, 387, 295)$, and $(387, 295, 296)$ are created. This results in three instances representing the user session.

In case of more than one observed exception in a user session, all occurring exceptions are considered. The dependent variable indicates whether an exception occurs immediately after the session events sequence. In consequence of modelling user behaviour as session events sequences, the system features are replicated and added to each session events sequence.

3.3 Distribution of Data Sets' Dependent Variable

Figure 3 displays both data sets' number of instances and the distribution of the dependent variable.

Table 3. Overview of Data Set I and II

Data	*Number of Instances*	*Percentage With no Exception*	*Percentage With Exception*
data_set_I_$n = \{3, 6, 9\}$	49,085	91.27	8.73
data_set_II_$n = 3$	73,998	99.46	0.54
data_set_II_$n = 6$	73,294	99.46	0.54
data_set_II_$n = 9$	72,352	99.45	0.55

The instances of Data Set I represent 49,085 user sessions covering session events from the beginning of a session only. The number of instances is independent of n. One or more exceptions occur in 8.73 percent of these user sessions.

Data Set II only covers 316 user sessions each with at least one exception. The number of instances depends on n since each user session is divided into session events sequences of length $n \epsilon \{3, 6, 9\}$. This results in three data sets with a slightly different total number of instances. The number of instances is equal to the number of created session events sequences. Compared to Data Set I it is even more imbalanced.

4 Evaluation of Classification Methods

Five classification methods are benchmarked on both data sets using R (version 3.4.1) and the mlr package [20,21]. The benchmarks are computed on cloud computing services, requiring memory in the range of 10 to 72 GB. Stratified 10-fold-cross-validation is chosen for both benchmarks since it is a generally accepted validation method [7,22,23]. At first, only system features are used for both benchmarks.

Afterwards, session events sequences with $n \in \{3, 6, 9\}$ are added to the system features and the benchmarks are performed again.

4.1 Benchmark of Classification Methods with Data Set I

In Data Set I, only the first session events of a session are considered. The classification task at hand is to predict whether at some point during the session an exception will occur. The benchmarking results of all five classification methods on Data Set I are displayed in Table 4.

Table 4. Benchmark I, average MCC (average AUC)

Classification Method	*System Features*	*System Features + Session Events*		
		n = 3	*n = 6*	*n = 9*
Naïve Bayes	0.04 (0.69)	0.11 (0.70)	0.04 (0.70)	0.02 (0.60)
XGBoost	0.04 (0.72)	0.27 (0.77)	0.41 (0.75)	0.51 (0.72)
Random Forest	0.00 (0.53)	0.21 (0.66)	0.58 (0.80)	0.69 (0.86)
Logistic Regression	0.04 (0.59)	0.23 (0.60)	0.48 (0.71)	0.59 (0.78)
K-Nearest Neighbour	0.16 (0.73)	0.33 (0.84)	0.63 (0.92)	0.78 (0.96)

Regardless of the classifier and based on system features only, the prediction whether an exception occurs in a user session is random. Except for the Naïve Bayes classifier, the prediction performances improve, when session events from the session's beginning are added. The more session events at the session's beginning are included, the better the prediction performance of all classification methods is, except for Naïve Bayes. K-Nearest Neighbour based on system features and $n = 9$ session events predicts the occurrence of an exception best, leading to an average MCC of 0.78, which is considered strong. It indicates a very high degree of correlation between the classifiers prediction and the actual outcome. The results of XGBoost are further evidence for the unsuitability of the AUC as evaluation metric. For $n = \{3, 6, 9\}$ the average AUC is almost the same while the average MCC shows considerable differences.

4.2 Benchmark of Classification Methods with Data Set II

In Data Set II, all possible sequences of size $n \in \{3, 6, 9\}$ in a session are considered. The classification task at hand is to predict whether an exception occurs immediately after the corresponding sequence. The Benchmarking results of all five classification methods on Data Set II are displayed in Table 5.

Table 5. Benchmark II, average MCC (average AUC)

Classification Method	*System Features*	*System Features + Session Events Sequences*		
		n = 3	*n = 6*	*n = 9*
Naïve Bayes	0.03 (0.73)	NaN (0.58)	NaN (0.50)	NaN (0.50)
XGBoost	NaN (0.62)	0.36 (0.69)	0.44 (0.69)	0.53 (0.71)
Random Forest	0.00 (0.52)	0.12 (0.73)	0.17 (0.88)	0.27 (0.92)
Logistc Regression	0.05 (0.75)	0.14 (0.83)	0.26 (0.81)	0.42 (0.84)
K-Nearest Neighbour	0.03 (0.57)	0.40 (0.83)	0.60 (0.91)	0.66 (0.88)

Again, regardless of the classifier and based on system features only, the prediction whether an exception occurs in a user's session is random. For XGBoost, there is no MCC available, since FP = TP = 0. As observed in Benchmark I, the prediction performance improves when the actual session events sequences are considered in addition to the system features. The larger the session events sequences, the better the prediction performance is. K-Nearest Neighbour, based on system features and session events sequences of length n = 9, predicts the occurrence of an exception best, leading to an average MCC of 0.66, which is considered strong.

4.3 Discussion of Results

In both benchmarks, the Naïve Bayes classifier shows poor prediction performance. The reason is that session events data violates the Naïve Bayes assumption that all features should be independent. Session events are clearly dependent on previous session events, since they cover the user's click tree.

Both benchmarks differ in the classification task. Benchmark I's prediction performance is slightly better than Benchmark II's. One reason for that difference might be the degree of class imbalance which is worse in Data Set II. With the obtained benchmark results, we can answer the research questions.

RQ1: What potential do user behavioural data have for the prediction of exceptions in user sessions? Both Benchmark I and II show a high potential of behavioural data, here session events, for the prediction of exceptions in user sessions. Complementing system features by behavioural data improves all classifiers' prediction performance, except for Naïve Bayes. We conclude that the reason for this performance improvement is that the user's actions are responsible for triggering exceptions. System features by itself do not contain any information about which software methods are used. This is evidence to include both behavioural data and system features for PreX.

RQ2: How important is the beginning of a user session for the occurrence of a software exception during the session? Benchmark I clearly demonstrates that the beginning of a user session contains information about the later occurrence of exceptions. The first session events in a user session indicate which task a user pursues and this task is assumed to be correlated with the later occurrence of an exception. The larger the sequence of leading session events, the better the prediction performance of all classifiers, except for Naïve Bayes, is.

RQ3: What potential does the modelling of user behaviour have for predicting exceptions in a live system? For the development of PreX, it is essential to predict the exact moment of an occurring exception. Otherwise, it is very difficult to initiate countermeasures to prevent the exception. Modelling session events in sequences of a given length mitigates the time problem of Data Set I. The results of Benchmark II show that modelling user behaviour by session events can lead to a substantial improvement on the prediction performance of exceptions. For all classifiers except Naïve Bayes, session events sequences of length $n = 9$ result in better prediction performance than sequences of length $n = \{3, 6\}$.

This study's results measured with the AUC metric are comparable to Lessmann et al. [7], even though we predict exceptions and Lessmann et al. predict software defects. The results of Benchmark I range between an average AUC of 0.53 and 0.96, while the results of Benchmark II range between 0.50 and 0.91. Lessmann et al. report results between an average AUC of 0.50 and 0.97 in their benchmark study [7]. Overall the benchmark results measured with the MCC show a positive correlation between the classifiers prediction and the actual outcome. K-Nearest Neighbour is identified as the single superior classifier.

5 Conclusion

A self-learning mechanism which improves software reliability post release is of interest in research and practise. PreX is at the starting point of its development and is a new use case for applied machine learning. While PreX remains invisible for software users, they are leveraged as large-scale testing resources who unconsciously contribute to the software's reliability. PreX is a generic mechanism which can be used in any software application regardless of its use case. Different software has different requirements concerning the importance of minimising FP and FN predictions. Therefore, we apply the MCC metric which balances those. When software developers implement PreX they can adjust the costs of FP and FN predictions in line with the application's requirements.

We contribute to PreX's development in three areas. First, this study is evidence for the ability to predict exceptions regardless of their type based on a C# stand-alone software application. Second, the choice of data set design has a substantial impact on the classifier's prediction performance. Third, for the classification task of exception prediction, the choice of classifier also has a substantial impact on the prediction performance. Clearly, the Naïve Bayes classifiers assumptions are violated, making it unsuitable for behavioural data. In both benchmarks, K-Nearest Neighbour shows superiority to the other classification methods. However, the choice of classifier is closely related to the data set design and hence cannot be generalised.

The limitations of this study are a domain specific data set, the application of only five classification methods, and some data set design decisions. Especially, Data Set II is restricted to 316 user sessions. However, more user sessions with exceptions are desirable. Another limitation is the decision to model session events sequences with fixed length of $n = \{3, 6, 9\}$.

This study’s database offers opportunities for further research in the four areas of sequence analysis, modelling of session events, prediction problems and classification runtime.

First, future work should analyse the relationship between session events and exceptions in further detail. One approach might be to identify single session events likely to cause exceptions, another approach is to apply sequence analysis to discover sequences which appear frequently and cause exceptions.

Second, further modelling of user behaviour should be analysed. A third model could combine the approaches of Data Set I and II. Session events sequences always include all session events from the session’s start until the last observed user event. When a new user event is recorded, a new instance is created with a session events sequence covering all user events up to this point.

Third, the exception location and exception type are of interest for PreX’s resumption model. Therefore, future work should predict the exception location as typical for the research field of software defect prediction and the exception type. Nonetheless, the resumption model itself should be topic of future work.

Fourth, the runtime of classifying a user session is essential for implementing PreX. To ensure an unimpaired user experience, future work should study how much time each classifier needs for the predictions in a session.

In summary, PreX aims to improve software reliability based on the occurrence of exceptions only. Exceptions are the effect of software defects. Hence, PreX does not correct any software defects, but rather is a concept to improve software reliability by controlling the effects of the actual cause. While PreX is no substitute for proper software testing, it is a powerful concept to mitigate the effects of software defects.

6 Acknowledgements

This study’s database has been sponsored by a software start-up company. We thank them very much for sharing their data with us and sponsoring computational resources. Furthermore, we would like to thank both reviewers for their valuable comments.

References

1. Arora, A., Caulkins, J.P., Telang, R.: Research Note—Sell First, Fix Later. Impact of Patching on Software Quality. Management Science 52, 465–471 (2006)
2. Goodenough, J.B.: Exception Handling: Issues and a Proposed Notation. Communications of the ACM 18, 683–696 (1975)
3. Cabral, B., Marques, P. (eds.): Exception Handling: A Field Study in Java and .Net. Springer (2007)
4. Lourenço, J.R.: PreX – Preventive Exception Handling. Portugal (2016)
5. Lourenço, J.R., Cabral, B., Bernardino, J.: A Predictive Model for Exception Handling. In: Rocha, Á., Correia, A.M., Adeli, H., Reis, L.P., Mendonça Teixeira, M. (eds.) New Advances in Information Systems and Technologies, pp. 767–776. Springer International Publishing, Cham (2016)

6. Wahono, R.S.: A Systematic Literature Review of Software Defect Prediction. Research Trends, Datasets, Methods and Frameworks. Journal of Software Engineering 1, 1–16 (2015)
7. Lessmann, S., Baesens, B., Mues, C., Pietsch, S.: Benchmarking Classification Models for Software Defect Prediction: A Proposed Framework and Novel Findings. IEEE Transactions on Software Engineering 34, 485–496 (2008)
8. Arora, I., Tetarwal, V., Saha, A.: Open Issues in Software Defect Prediction. Procedia Computer Science 46, 906–912 (2015)
9. Lyu, M.R. (ed.): Handbook of Software Reliability Engineering. IEEE Computer Society Pr, Los Alamitos Calif. (1996)
10. Endres, A. and Rombach, H.D.: A Handbook of Software and Systems Engineering. Empirical Observations, Laws and Theories. Pearson Addison Wesley, Harlow, England, Munich (2003)
11. Flaviu, C.: Exception Handling and Software Fault Tolerance. IEEE Transactions on Computers Vol. C-31, 531–540 (1982)
12. Fawcett, T.: An Introduction to ROC analysis. Pattern Recognition Letters 27, 861–874 (2006)
13. Briggs, W.M., Zaretzki, R.: The Skill Plot. A Graphical Technique for Evaluating Continuous Diagnostic Tests. Biometrics 64, 250–256 (2008)
14. Matthews, B.W.: Comparison of the Predicted and Observed Secondary Structure of T4 Phage lysozyme. Biochimica et Biophysica Acta (BBA) - Protein Structure 405, 442–451 (1975)
15. Powers, D.M.: Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. Journal of Machine Learning, 37–63 (2011)
16. Ding, Z.: Diversified Ensemble Classifiers for Highly Imbalanced Data Learning and Their Application in Bioinformatics. Atlanta, Georgia (2011)
17. Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F., Nielsen, H.: Assessing the Accuracy of Prediction Algorithms for Classification. An Overview. Bioinformatics Review 16, 412–424 (2000)
18. Boughorbel, S., Jarray, F., El-Anbari, M.: Optimal Classifier for Imbalanced Data Using Matthews Correlation Coefficient Metric. PloS one 12 (2017)
19. Evans, J.D.: Straightforward Statistics for the Behavioral Sciences. Brooks/Cole, Pacific Grove, Calif. (1996)
20. Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Jones, Z., Casalicchio, G., Gallo, M., Bossek, J. and Studerus, E., et al.: R Package Machine Learning in R, https://cran.r-project.org/web/packages/mlr/mlr.pdf
21. Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., Jones, Z.M.: R Package mlr. Machine Learning in R. Journal of Machine Learning Research 17, 1–5 (2016)
22. Salzberg, S.L.: On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach. Data Mining and Knowledge Discovery 1, 317–328 (1997)
23. Witten, I.H., Frank, E. and Hall, M.A.: Data Mining. Practical Machine Learning Tools and Techniques. Elsevier, Amsterdam (2011)

Discussing the Value of Automatic Hate Speech Detection in Online Debates

Sebastian Köffer[1], Dennis M. Riehle[1], Steffen Höhenberger[1], and Jörg Becker[1]

[1] University of Münster, European Research Center for Information Systems, Münster, Germany
{sebastian.koeffer,dennis.riehle,steffen.hoehenberger, joerg.becker}@ercis.uni-muenster.de

Abstract. This study discusses the potential value of automatic analytics of German texts to detect hate speech. In the course of a preliminary study, we collected a dataset of user comments on news articles, focused on the refugee crisis in 2015/16. A crowdsourcing approach was used to label a subset of the data as hateful and non-hateful to be used as training and evaluation data. Furthermore, a vocabulary was created containing the words that are indicating hate and no hate. The best performing combination of feature groups was a Word2Vec approach and Extended 2-grams. Our study builds upon previous research for English texts and demonstrates its transferability to German. The paper discusses the results with respect to the potential for media organizations and considerations about moderation techniques and algorithmic transparency.

Keywords: Natural Language Processing (NLP), Hate Speech, Text Analytics

1 Introduction

Online debates have gone off the rails. In a much-noted piece in April 2016, The Guardian published details about user comment behavior on the newspapers' website. Many comments were "crude, bigoted, or just vile". As "xenophobia, racism, sexism and homophobia were all seen regularly", the authors called it "the dark side of Guardian comments" [1]. In Germany, the amount of abusive content on the Internet during the refugee crisis has sparked a national debate on how to deal with online hate speech. German authorities formed a task force that ultimately urged social media providers to apply tougher filtering mechanism for hateful content – an action that was also criticized as excessive political correctness and censorship [2].

Detection of abusive language in user-generated online content has become an important issue for various stakeholders [3]. For instance, it is likely that hate speech and actual hate crimes relate to each other. Benesch [4] reported that hateful language delivered in the media resulted in massive violence in Kenya before and after the elections in 2007 and 2008. Similarly, German commenters proposed a relationship between the hateful online debate on refugees and attacks on homes for asylum seekers [5]. It is apparent that the nature of online debates has changed. They are

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

often characterized by ideological and extreme opinions that frequently discard facts and scientific evidence. As a result, many newspapers and magazines have started fact-checking projects but at the same time, they face a serious criticism towards journalists in particular and the media in general.

Flagging hateful contents is essential for media organizations. Ignoring the problem may lead to less user traffic on their websites and companies pulling advertisements [3]. Among journalists, there is a common sentence: "Don't read the comments" [1]. In contrast to this, news organizations and their community managers try to maintain the conversation with their readers by answering comments, fact checking and explaining journalistic methodologies. As a result, comment moderation is a major manual effort for the media organizations' community managers [6].

However, many news platforms are obviously unable to cope with this demand and have limited the possibilities to comment on articles or at least on articles about contentious topics, such as refugees, conspiracy theories, climate change, and feminism [7]. In Germany, this behavior is also enforced by German law that requires community owners to delete user-generated content immediately as soon as it is known that comments contain so-called "incitement to hatred"[1]. A survey among German newspaper editors found that about 50 percent applied restrictions to the online comment sections [8].

Given the increasing amount of user-generated comments[2], we argue that analytics will ultimately be required to check for and delete abusive content and, at the same time, curate inspiration and wisdom in the debates. Especially small media organizations might be unable to deal with an avalanche of comments after publishing articles on contentious topics. Hence, the goal of this research paper is *to investigate the potential value of automatic analytics of German texts to detect hate speech*. To this end, we aim to inform community managers when thinking about putting in place algorithmic methods to support comment moderation. We draw on existing work in the area of natural language processing (NLP) for English texts. In addition to this paper, we will make public explanations for a broader audience on our projects' website[3]. This will include the possibility to access trained datasets via application programming interfaces (API).

The paper is structured as follows. In the next chapter, we list related studies on our topic. Next, Chapter 3 describes our datasets and data collection methods. Our research methodology with regard to NLP and statistics is explained in Chapter 4. After presenting our results in Chapter 5, Chapter 6 discusses implications for research and practice. The paper concludes by presenting the limitations of our work, combined with a discussion on pathways for future research.

---

1 German Criminal Code in the version promulgated on 13 November 1998, Federal Law Gazette - https://www.gesetze-im-internet.de/englisch_stgb/englisch_stgb.html#p1241

2 For instance, the New York Times receives around 9,000 submitted comments per day [9]. The Guardian receives several tens of thousands comments every day [1].

3 This paper build upon the "Cyberhate-Mining" research project: www.hatemining.de

2 Background

The detection of abusive content, including hate speech, is not trivial: Different dictions, a huge variety of special terms for insults, a context-specific meaning, and a lot of sarcasm in the text make the task rather difficult [3]. Nonetheless, several researchers have tried to detect abusive contents in user-generated content automatically by means of text analytics [10].

Studies tried to identify hate speech in particular. Waseem and Hovy [6] used n-grams to detect hateful content in Twitter posts. Similarly, Burnap and Williams [11] detect hate speech in Tweets using a "Bag-of-Words" approach. Warner and Hirschberg [12] aim to identify hate speech in online texts by means of website and user comment annotations combined with word-sense disambiguation. Finally, Nobata et al. [3] – a group of Yahoo Labs researchers – are using deep-learning inspired methods to detect abusive comments, including hate speech. The authors apply to set of feature exploration techniques that we overtake for our study.

3 Data

3.1 Primary Dataset

We collected user-generated comments that were publicly available on news platforms on the Internet. The extraction of data included mainstream journalistic news websites as well as websites of so-called alternative media. Since most German news platforms do not offer an API to collect the comments programmatically, we used web scraping technologies. For the implementation of web scraping techniques, we used a Python framework called Scrapy[4], which has been regularly used for data collection in research projects, e.g., [13].

To select appropriate news platforms for our data collection, we rated 41 news platforms using the following criteria:

- *Comments allowed:* Platforms were excluded which did not allow user comments on articles related to the refugee crisis or which did not allow user comments at all.
- *Bots allowed:* We adhered to international standards on accessing websites with bots by respecting all bans specified in a robots.txt file. Therefore, any platforms disallowing access to bots were discarded.
- *Expected number of comments:* Since we focused on collecting large amounts of comments, we only considered platforms with a reasonable amount of articles and, even more important, with a reasonable amount of user-generated comments.
- *Estimated complexity:* We evaluated all platforms regarding their complexity. For instance, web scraping becomes more complex if websites make use of dynamic web technologies, such as AJAX or JavaScript. We followed a "low-hanging fruits" approach, starting with platforms, where comments were rather easy to scrape.

[4] https://www.scrapy.org/

Out of the 41 platforms, 16 did not allow comments to articles related to the refugee crisis. Further, two platforms denied to scrape their content using a robots.txt file. Seven of the remaining platforms did not have an active community, i.e., only very few comments could be found. And last, three platforms were discarded because of difficulties regarding their use of DDoS protection mechanisms or comments being loaded asynchronously with JavaScript. Thus, our dataset comprises 13 platforms.

Since website structures tend to vary a lot, it is necessary to implement an individual scrape mechanism for each platform. Three different scraping methods were used to ensure that only articles related to the refugee crisis were selected:

- Take articles from *topic pages*, which only list articles related to the refugee crisis. Some platforms offered special dossiers on the refugee crisis.
- Use of the news platforms' *search* function. Sometimes, this is forbidden by means of the robots.txt and then was discarded.
- Searching the websites using 79 keywords within the articles. Exemplary keywords were *asylum seeker*, *immigrant*, *integration*, *refugee*.

In total, 376,143 comments and 21,740 articles have been collected. Table 1 shows that the number of articles per platform varies greatly from 210 for "Alles Schall und Rauch" to 5,812 for "Zeit" with an average number of 1,672 articles per platform. As regards the comments, the amount per platform also varies greatly. With 182,625 comments, the platform "Welt" has by far the most comments, followed by "Focus" with 75,857 comments. All other platforms have less than the average number of comments per platform which is 28,933 comments.

Table 1. Primary dataset and scraping method

Platform	*Method*	*Articles*	*Comments*
Alles Schall und Rauch	Keyword	210	4,617
Cicero	Keyword	260	4,415
Compact	Search	328	11,764
Contra Magazin	Search	543	7,984
Epoch Times	Search	4,584	27,497
Focus	Keyword	3,959	75,857
Freie Welt	Search	1,944	13,628
Junge Freiheit	Keyword	333	2,745
NEOPresse	Search	626	11,054
Rheinische Post	Topic page	991	3,678
Tagesspiegel	Topic page	229	4,478
Welt	Keyword	1,921	182,625
Zeit	Topic page	5,812	25,792
Total		21,740	376,143

We also collected additional meta information for articles and comments, including the date of publication. Peaks are visible in late summer of 2015 and shortly after

New Year's Eve in January and February of 2016. This development corresponds to substantial events that occurred during the refugee crisis.

3.2 Evaluation Dataset

An important part of this study was to classify collected comments and to determine whether they are perceived as hateful or not. To gather a substantial collection of ratings, we collected ratings via an online survey. Using Crowdsourcing to obtain labeled training data is a common approach in research projects that deal with natural language processing to detect emotions in texts [3, 14].

Inspired by previous work on detecting hateful speech [3, 6], we used a binary categorization, so that study participants rated comments as "hate" or "no hate". In addition, study participants could also decide to skip a comment if they were unsure whether it contained hate or not.

From May to June 2016, study participants rated randomly selected comments on the project website. The selection of comments ensured that we got a similar amount of labeled data for each platform. Each comment needed to be rated by multiple participants before a final scoring decision was taken. Thus, a comment was labeled as "hate" only, if there were three hate ratings and at most one "no hate" rating and vice versa. In addition, comments that were skipped two times more than they were rated, or comments that received a 2:2 rating, were discarded.

Throughout the whole time span of our study, we received 11,973 ratings from 247 individual users in total. Among these, there were 3,875 hate, 6,073 no hate, and 2,025 unclassified ratings. According to the rules described above, this led to 2,983 labeled comments in total as depicted in Table 2. With 50 %, the largest amount of comments perceived as hate was found on "Contra Magazin", while lowest amount of perceived hate was found on "Tagesspiegel" (11 % of all comments).

Table 2. Evaluation data overview (scores) per platform

Platform	*# Hate*	*# No Hate*	*# Unclassified*	*% Hate*
Alles Schall und Rauch	54	119	61	23
Cicero	46	122	42	22
Compact	68	121	41	30
Contra Magazin	117	75	42	50
Epoch Times	93	113	39	38
Focus	45	147	57	18
Freie Welt	90	88	42	41
Junge Freiheit	74	91	47	35
NEOPresse	46	111	53	22
Rheinische Post	59	108	42	28
Tagesspiegel	26	170	40	11
Welt	55	136	48	23
Zeit	28	154	48	12
Total	801	1555	602	27.15

The overall share of hateful comments (27.15 %) is comparable high. The datasets used by Nobata et al. [3] only contain about 10 % of abusive comments. Several aspects might have contributed to the high share of hateful comments. For instance, our selection of comments is limited to articles on the refugee crisis that triggered very emotional debates. Also, our demographics of our survey participants are biased towards young people. During the rating process, all participants were asked to submit their gender age voluntarily. Out of the 247 participants, 169 did provide their age and gender; the remaining 78 users submitted neither age nor gender. The users' demographic structure is depicted in Table 3.

Table 3. Demographics of rated comments

Age group	*Male*	*Female*	*Total*
Below 25	24.7%	8.2%	33.0%
25-30	28.6%	12.7%	41.4%
31-35	5.2%	7.7%	12.9%
Over 35	6.5%	6.3%	12.7%
Total	65.0%	35.0%	100.0%

4 Methodology

4.1 Research Approach

The adoption of algorithms for comment moderation challenges the norms of transparency in journalism [15]. Originally, analytical methods to detect sentiments used vocabularies that contain sentiment words assigned with particular emotions and opinions [16]. One advantage of vocabularies is that their functioning is more comprehensible also for non-technical people.

Our study builds upon previous work by Nobata et al. [3] who evaluated several classification methods of NLP features to detect abusive content. Furthermore, we were inspired by a Kaggle competition on predicting online movie ratings from review texts [17]. Similarly to the competition, we juxtapose the vocabulary-based approach with deep-learning inspired methods that focus on the meaning of words. Most NLP studies for detecting emotions in user-generated content are using English texts only. Nobata et al. [3] note that "it remains to be seen how our approach […] would fare in other languages" (p. 152). Our study shall contribute to transfer efforts of NLP techniques with respect to German language.

4.2 Feature Extraction

We overtook the feature classification from Nobata et al. [3] who grouped their features into n-grams, linguistics, and distributional semantics (Word2Vec, Doc2Vec). In addition, we use a bag-of-words model to create a vocabulary of hateful and non-hateful words. Besides its simplicity, n-gram techniques have produced good and effective results. Thus, we decided to develop an additional feature group named

“Extended n-grams” that combines n-grams and distributional semantics. In the following, we describe the extracted feature groups in more detail[5].

Bag-of-words. To build up our vocabulary we first removed or substituted special characters, such as ä, ö, ü, and ß. Subsequently, a stop word list[6] was used to remove words from the vocabulary that are insignificant for hate speech. We also considered stemming and lemmatization for preprocessing using algorithms from the Snowball[7] project. To get numeric representations for our classifiers, we used the inverse document frequency (tf-idf). This approach yielded slightly better results than the CountVectorizer that was used in the tutorial for the Kaggle competition [17].

N-grams. We used character 2- and 3-grams. Regarding the German alphabet with 26 letters, the special characters ä, ö, ü, ß, and the space character, we obtain at most 31^2 (31^3) different 2-grams (3-grams). We used the normalized tf-idf value to determine the relative importance in the text corpus.

Linguistics. We extracted 20 features with comparatively low computational complexity. Exemplary features include the count of words, sentences, capital letters, punctuation (!?.,”), smileys, and URLs as well as the average word length and the average number of words per sentence. In order to ensure comparability between comments of different length, features were scaled in relation to the appropriate metric of the comment, i.e., number of sentences, words, characters.

Word2Vec / Doc2Vec. We used the 376,143 collected comments as training data for the Word2Vec model [18]. For feature extraction, we first transformed each word that appeared at least two times in the training into its vector representation. Subsequently, we determined the mean vector of all word vectors which is used as inputs for the features. The number of features is determined by the dimensionality of the vector. Here, we followed Nobata et al. [3] to select 50 dimensions. Similarly, we trained the Doc2Vec [19] model with all collected comments. Then, the trained model returned vector representations for all new comments. Again, we used 50 dimensions for the size of the Doc2Vec vector.

Extended n-grams. N-grams techniques cannot consider semantically equivalent but syntactically divergent texts. For instance, the words “Merkel” and “Bundeskanzlerin” most likely have a similar meaning, but the related n-grams are rather different. Our extended n-grams make use of the Word2Vec model to enrich original comment texts with nearest neighbors that are derived from the word vector representations (cosine similarity). To this end, we determined the normalized tf-idf value for each word except stop words. The higher the tf-idf measure, the more words were appended to the original comment for emphasizing words. The extended comments were then used to derive the n-gram feature.

---

5 For a detailed explanation of feature extraction approaches, please refer to Nobata et al. [3].

6 The list is available as package of the Python Natural Language Processing Toolkit (NLTK) via https://pypi.python.org/pypi/stop-words. It is maintained by Alireza Savand.

7 A collection of stemming algorithms for several languages: http://snowballstem.org

4.3 Supervised Learning

The numerical features of the distinct feature groups (created only from the comment text itself) served as input for the classification models. For this task, only labeled comments (811 hate, 1,561 no hate) were considered. These were applied on logistic regression and evaluated to identify the best classification model[8]. The implementation was performed in Python using packages of the scikit-learn[9] module. A train and test set validation approach was chosen using a split of 75:25 between train and test set. Furthermore, we used undersampling to have equal sample size for the two classes. Thus, only 811 non-hateful comments were sampled, and the complete evaluation dataset was composed of 811 + 811 = 1,622 comments.

5 Results

Table 4 depicts our results. We report accuracy (ACC) and F-score for our models. We also tried whether combinations of two feature groups perform better. The bag-of-words approach obtained the best ACC value with 67.8 percent. The highest F-score was obtained using the Word2Vec with 0.67. The best performing combination of feature groups was Word2Vec and Extended 2-grams (ACC = 0.7068, F-score= 0.70).

Table 4. Performance of feature groups for classification task

Feature group	*ACC*	*F-score*
Bag-of-words	0.6780	0.51
2-grams	0.6206	0.64
3-grams	0.6551	0.65
Linguistics	0.5689	0.53
Word2Vec	0.6650	0.67
Doc2Vec	0.6477	0.63
Extended 2-grams	0.6009	0.61
Extended 3-grams	0.6059	0.61

For the bag-of-words approach, Table 5 shows the words that are most indicative for hateful comments, i.e., the appearance of the word “Europe” mostly increases the probability that a comment is considered as hateful.

Some of the hate indicative words can be related to political topics. For instance, chancellor Merkel promoted an open culture for refugees and faced a lot of criticism in online debates. The list of the non-hateful indicative words contains auxiliary words (gar, vielen) that could have been part of the stop word list. However, for the purpose of this study, we stick to the lists that we overtook from previous work.

[8] We also applied Support Vector Machines. Since the results were similar but slightly worse compared to the logistic regression, we do not report the figures as part of this paper.

[9] https://www.scikit-learn.org

Table 5. Five most indicative words for hateful and non-hateful comments[10]

Hate		*No hate*	
German	*English*	*German*	*English*
europa	Europe	finde	find
verbrecher	criminals	artikel	article
luegen	lies	integration	integration
duerfen	may / can	vielen	many
merkel	merkel	gar	even

6 Discussion

In this study, we examined the value of text analytics for an automatic detection of hate speech in German texts. Therefore, we conducted a preliminary study in which we collected a dataset of user comments on German news articles, focused on the refugee crisis in Germany in 2015/16. A crowdsourcing approach was used to label a subset of the data to be used as a training and evaluation dataset. We then selected feature groups that are anchored in related other scientific work to evaluate a classification model using a logistic regression approach. Furthermore, a vocabulary has been created containing the words that are indicating hate and no hate.

Our study demonstrates that previously used concepts by other researchers [3, 6, 17] can be transferred to German texts. However, German language specifics, like irregular plural forms, compound nouns or anglicisms complicate the process of stemming and lemmatization. As a final result, we achieved best results with an accuracy of approximately 70 % and an F-score of 0.7. Thus, our results are slightly outperformed by recent academic work that used similar methods with English texts[11]. Given the limitations of our work, particularly the available datasets, we rate the preliminary study's outcome as promising and satisfactory.

Our results partly confirm the strength of character-based NLP techniques. Despite their simplicity, character-based (and also word-based) techniques do not perform considerable worse in our study than more complicated mechanisms, such as distributional semantics. However, we think that distributional semantics and other deep-learning inspired techniques have the potential to outperform character-based and word-based techniques as soon as the datasets are big enough. For instance, training Word2Vec on a lot more text should improve performance [17].

As regards our vocabulary, we argue that our list of most indicative words may contribute to an increased transparency of analytical methods. Algorithmic methods for comment moderation are soon criticized as automatic censorship to repress political opponents [15]. If analytical approaches are able to share intermediate results and explanations, they may have the potential to be more comprehensive and more

[10] For the purpose of this paper, we translated the words into English.

[11] Nobata et al. [3] achieved F-scores up to 0.81 to detect abusive content in a news dataset by combining similar feature groups that are used in this study. Waseem and Hovy [6] reached F-scores up to 0.74 using character n-grams to detect hateful comments in Tweets.

objective than any netiquette that is used as a guideline for manual comment moderation, which usually happens behind the scenes. However, while our hate-indicative word lists may slightly open the analytical black box, they may also be a target for criticism itself if people do not agree with certain elements of the list.

7 Limitations and Outlook

First and foremost, the biggest shortcoming of this study is the relatively small size of the dataset. To train our algorithms, we were confined to a set of 2,372 labeled comments. Related studies that apply machine learning with NLP use massively bigger datasets with hundreds of thousands labeled texts [3, 20]. We plan to further increase our labeled dataset in the future and are confident that this will increase the chances to obtain the same evaluation scores like other researchers. This is particularly important to gain more acceptances for algorithmic approaches by journalists and news organizations.

Second, although web scraping is the (only) suitable approach for us, it has several pitfalls. For instance, we cannot guarantee that our scraping method has collected all relevant articles on the refugee crisis or whether there have been errors when scraping the comment texts. Furthermore, we had to discard many news platforms that would have been worth to analyze. However, our scripts worked reliable so that we were able to obtain the data rather easily. Nevertheless, the fact that larger international newspapers (e.g., New York Times) offer APIs might encourage German news platforms to follow at some point.

Third, our study is limited to comments on news platforms and articles on the refugee crisis. Thus, its findings cannot be transferred directly to other topics and platforms, such as social media platforms. We chose this focus because we believe that media organizations will ultimately need analytical approaches to maintain online debates on their websites. If online debates continue to move from journalistic media to social media platforms, journalists will lose their opportunity to steer and enrich the debates, and ultimately be ever more dependent on social media platforms [21]. To work on bigger datasets and better data labeling, we encourage German media organizations and researchers to join forces. The Coral Project[12] where New York Times, Washington Post, developers, and researchers team up to "build better communities around their journalism" is the prime example.

Fourth, our dataset is already pre-filtered by the news platforms that use very different moderation strategies to delete hateful comments before they are publicly visible. We do not precisely know which semi-automatic techniques for comment moderation are already in place[13]. It would be interesting to have access to the raw data which is likely to contain more hateful contents. Hence, the results must be interpreted carefully, because the dataset does not directly represent what people write in the online comment sections.

---

[12] The Coral Project: https://www.coralproject.net

[13] The Guardian revealed that 1.4 million (2% of the total) comments had been blocked by February 2016 using manual moderation [1].

Fifth, the study's participants to label the comment texts as hateful or non-hateful are not representative of the whole population. They were recruited via social media platforms among people in a University context. Hence, participants are rather young with a relatively high level of education. Furthermore, it is unlikely that many participants possess journalistic expertise or experience with community moderation. Future studies should use a more representative sample of the population.

Sixth and finally, our decision to use a binary classification between hate and non-hate is problematic, since every individual might have a different understanding what hate means. In a subsequent study, we plan to further detail the comment ratings to be able to distinguish between different aspects of hate speech such as insults, xenophobia, and threats.

To conclude, an analytical tool for comment moderation must deliver a high level of accuracy to meet high journalistic standards. Accuracy values around 70 % as in our preliminary study or around 80 % like in related studies are probably still insufficient. But even if better datasets and algorithms allow better prediction rates, they do not necessarily call for an automatic deletion of hateful comments. Since no analytical approach is likely to guarantee almost zero failure in a foreseeable future, false positives may continue to trigger discussions of undesirable censorship by media. From the discussions of our study results with several stakeholders we conclude that semi-automatic approaches, where moderators review the analytical outcomes are more feasible. Such approaches can also include the commenters themselves since they could get immediate feedback about the submitted comment text. This example is just a fraction of potential pathways that can be envisaged through the use of analytical methods. We hope that our paper contributes to enabling the use of text analytics to bring online debates back on track – pursuing fruitful and enriching discussion on the web. It is an effort worth making.

Acknowledgements

This paper was written in the context of the research project "Cyberhate-Mining" (www.hatemining.de). We gratefully acknowledge the constructive comments and cooperation provided by the participating students, advisors, and colleagues.

References

1. Gardiner, B., Mansfield, M., Anderson, I., Holder, J., Louter, D., Ulmanu, M.: The dark side of Guardian comments, https://www.theguardian.com/technology/2016/apr/12/the-dark-side-of-guardian-comments.
2. Faiola, A.: Germany springs to action over hate speech against migrants, https://www.washingtonpost.com/world/europe/germany-springs-to-action-over-hate-speech-against-migrants/2016/01/06/6031218e-b315-11e5-8abc-d09392edc612_story.html.

3. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., Chang, Y.: Abusive Language Detection in Online User Content. In: International World Wide Web Conference Committee (IW3C2). Montreal, Quebec, Canada (2016).
4. Benesch, S.: Countering dangerous speech to prevent mass violence during Kenya's 2013 elections. (2014).
5. Lobo, S.: Wie aus Netzhass Gewalt wird und was dagegen hilft, http://www.spiegel.de/netzwelt/netzpolitik/netzhass-und-gewalt-was-man-dagegen-tun-kann-lobo-kolumne-a-1048799.html.
6. Waseem, Z., Hovy, D.: Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. In: NAACL Student Research Workshop. San Diego, CA, USA (2016).
7. The Coral Project Community: Shutting down onsite comments: a comprehensive list of all news organisations, https://community.coralproject.net/t/shutting-down-onsite-comments-a-comprehensive-list-of-all-news-organisations/347.
8. Siegert, S.: Nahezu jede zweite Zeitungsredaktion schränkt Online-Kommentare ein, http://www.journalist.de/aktuelles/meldungen/journalist-umfrage-nahezu-jede-2-zeitungsredaktion-schraenkt-onlinekommentare-ein.html.
9. Etim, B.: The Most Popular Reader Comments on The Times, http://www.nytimes.com/2015/11/23/insider/the-most-popular-reader-comments-on-the-times.html.
10. Schmidt, A., Wiegand, M.: A Survey on Hate Speech Detection using Natural Language Processing. Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, Valencia, Spain. 1–10.
11. Burnap, P., Williams, M.L.: Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. Policy and Internet Vol 7(2), pp. 223–242 (2015).
12. Warner, W., Hirschberg, J.: Detecting Hate Speech on the World Wide Web. In: Proceedings of the 2nd Workshop on Language in Social Media, pp. 19–26. Montréal, Canada (2012).
13. Wang, J., Guo, Y.: Scrapy-based crawling and user-behavior characteristics analysis on Taobao. In: International Conference on Cyber-Enabled Distributed Computing and Knowledge Discover,. pp. 44–52 (2012).
14. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word-emotion association lexicon. In: Computational Intelligence, pp. 436–465 (2013).
15. Diakopoulos, N., Koliska, M.: Algorithmic Transparency in the News Media. Digital Journalism. (2016).
16. Liu, B.: Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies Vol. 5, pp. 1–167 (2012).
17. kaggle.com: Bag of Words Meets Bags of Popcorn, https://www.kaggle.com/c/word2vec-nlp-tutorial.
18. Mikolov, T., Corrado, G., Chen, K., Dean, J.: Efficient Estimation of Word Representations in Vector Space. International Conference on Learning Representations (ICLR 2013), pp. 1–12 (2013).
19. Le, Q., Mikolov, T.: Distributed Representations of Sentences and Documents. International Conference on Machine Learning - ICML 2014, pp. 1188–1196 (2014).
20. Rao, Y., Lei, J., Wenyin, L., Li, Q., Chen, M.: Building emotional dictionary for sentiment analysis of online news. World Wide Web Vol. 17, pp. 723–742 (2014).
21. Brodnig, I.: Der Hass im Netz – und was dagegen zu tun ist, http://www.carta.info/81616/hass-als-instrument/.

Recommendation-based Business Intelligence Architecture to Empower Self Service Business Users

Safwan Sulaiman[1], Jorge Marx Gómez[1]

[1]Oldenburg University, Business Informatics/VLBA, Oldenburg, Germany
{safwan.sulaiman, jorge.marx.gomez}@uni-oldenburg.de

Abstract. Enterprises live today in a very competitive business environment, which is influenced by globalization, technical progress and volatile markets. Therefore, business intelligence systems are widely adopted to support enterprises to ensure a competitive edge against competitors by providing their decision makers with the required information at the right time. However, business users still face significant difficulties while performing both ad-hoc and sophisticated data analysis. In this paper, we propose a new approach for self-service business intelligence as a major support to business users by providing them with guided recommendations while they do their own data analysis. The main idea behind this approach is to transfer the knowledge of the power users, which is represented as analysis paths, to the business users in form of recommendations or suggestions to help them when struggling with business intelligence systems and their complex data analysis.

Keywords: Recommender System, Self-Service, Business Intelligence, Knowledge Transfer Model.

1 Introduction

In the current economic environment, enterprises face constant challenges to run their businesses in a very competitive business environment within a dynamic market [1]. Due to the rapid changes in the external and internal conditions of present day economic life, information has become a very important production factor for enterprises, who have become very dependent on reliable information [2]. Therefore, to ensure continued success in their businesses and to guarantee a sustained competitive edge, enterprises recognize the importance of providing their employees with relevant and timely information by using Business Intelligence (BI) tools effectively [3–5]. To respond to this situation enterprises have widely implemented and integrated BI systems [6].

In spite of the big market and the importance of BI and analytics for enterprises, various market studies and surveys have shown that the usage of BI systems is still very low. The Gartner Report showed that less than 30% of the potential users use BI systems [7]. Eckerson explain in his study about business driven BI that the BI penetration is minimal, only 26% of employees use BI tools [8]. According to the BI

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

survey 15 conducted by BARC Research, the median percentage of employees using BI in companies reached 13% [9].

In the literature, several user acceptance models for BI showed that the complexity of the tools is considered as biggest challenge for adoption of BI [10, 11]. In addition, due to their flexibility and powerfulness, BI systems are still too complex for business users, who face significant challenges in performing adequate ad-hoc analyses [12]. Alpar and Schulz pointed that when you offer business users more flexibility this requires more BI skills from them [13]. Consequentially, due to the high demand for reporting and data analysis, power users become bigger bottlenecks for an enterprise. The large amount of business users' requests leads to the inability of power users to react sufficiently rapidly and accordingly business users will make decisions without relying on information derived from data analysis, possibly making decisions based on their good feeling or experience, which lead sometimes to wrong decisions [8, 13, 14].

To react to this situation, a new trend in the BI domain has emerged called self-service BI. The Data Warehouse Institute (TDWI) defined self-service BI as the facilities within the BI environment that enable BI users to become more self-reliant and less dependent on the IT organization [15]. In this paper, we propose a new approach towards the realization of self-service BI, where the business users will be empowered and supported while performing sophisticated data analysis to get the information they need at the time they need it, without relying on or asking the IT department or the power users. This approach mainly supports business users while they do their own data analysis by providing them with guided recommendations. These recommendations are derived from the extracted knowledge of the power users (analysis paths). This is not to be confused with other similar researches, its goal being to help and support business users while using BI systems to improve the adoption and usage rate of BI in enterprises and not the automation of the decision making process based on old decisions from power users.

The remainder of this paper is structured as follows. First, an overview of the theoretical foundation and related work are provided in Section 2. After that, the conceptual architecture of the recommendation-based self-service BI will be explained in detail in Sections 3. Next, a prototype as a proof of concept and the evaluation of this research will be presented in Section 4. Finally, in Section 5, the conclusion of this paper and its result are summarized and the outline for future work presented.

2 Theoretical Foundation and Related Work

In this section, self-service BI is first described in more detail, to indicate its importance. After that, the recommender systems will be also described as an important part of this research. Finally, relevant related works are introduced.

2.1 Self-Service BI

According to the result of BI trend monitor 2017, self-service BI is one of the most important trends for BI practitioners in their work [16]. To meet the time-to-insight required by the current competitive business environment, many enterprises want to democratize analytics capabilities via self-service BI [6]. Self-service BI has been on organizations' wish lists for a long time as IT departments struggle to satisfy the ever-growing demands and requirements from business users for faster changes and flexibility for ad-hoc reporting and analysis to meet their BI needs [14, 16]. Imhoff and White argued that BI providers have already achieved a high level of ease of use for reporting and simple analytics. However, more complicated analytics still need be enhanced and made easy to use and create. The realization of self-service BI should focus on one or more of the following four main objectives: providing easy access to source data for reporting and analysis, building easy-to-use BI tools and improved support for data analysis, providing fast-to-deploy and easy-to-manage data warehouse options such as cloud computing, and designing simpler and customizable user interfaces [15].

2.2 Recommender System

Due to the ever-increasing amounts of data and the heterogeneity of the users, a Recommender System (RS) is of great scientific importance and is described in numerous researches [17] . The main goal of the RS is to provide recommendations for items most likely to be of interest to a specific user for helping him within various decision making processes, such as what product to buy, what movie to see, etc. [17]. RS was already wide-spread in the E-commerce domain, however, it is also used in numerous application domains like web search, computational advertisements, social media recommendations [17]. In addition, RS is increasingly being used in the area of Technology Enhanced Learning (TED) to provide learners with appropriate and individualized knowledge that supports all their learning activities and helps them work independently and more flexibly to achieve a specific learning goal [18]. This area of research is close to the problem defined above. Therefore, in the following section we will introduce some related work that has some similarities with our approach.

2.3 Related Work

In this section, we present some interesting work, which plays a special role for our research. These are published in the area of TED.

Zaiane described a concept to help e-learning users through intelligent recommendations navigation and operation of the software system [19]. The recommendations should be generated by means of an access history, which is derived from the server log files. This work tried to transfer the RS functionality from E-commerce to the E-Learning domain by using web-mining techniques to extract and find learning activities to be suggested to the learner for improving the course material navigations within the E-learning system. In the method described, these patterns are analyzed dynamically, so that the user receives recommendations for further possible steps directly from the current situation. In addition, the log sequences are compared by the user with other log sequences from other users in order to be able to provide him with additional interesting content specific to his needs.

Khrib et al. describe a concept for recommendations, which provides automatic and personalized learning material in E-learning systems [20]. The learning material is referenced via a URL in the e-learning system and archived via a log file or a database so that it can be used for later recommendations. This work describes two approaches to modeling user models. On the one hand, there is the creation of a user model by explicit information about the properties of a user. On the other hand, an implicit approach is used for the automatic generation of a user model based on the navigation history and the activities within a user's session. Users with the same interests are then grouped so that they can benefit from the behavior of others.

These approaches contain important characteristics for our work. However, analyzing log files of web servers to develop a RS for BI systems is not sufficient because many queries contain important information related to the usage of a BI system in the POST and GET parameters, which must also be considered for further recommendations. Furthermore, the focus of these approaches is to personalize a web site or E-learning system based on the user model of the learner or the user, while in our case the focus is to extract the analysis paths of the power users and then to suggest them to the business users independent of a specific business user.

3 Recommendation-based Architecture for Self-Service Business Intelligence

In this section, the conceptual architecture of the new self-service BI system will be explained. First, the main concept of the knowledge transfer model, which represents the main incentive of this architecture, will be described. After that, the first part of the knowledge transfer model, which is the capturing and extracting the knowledge of the power users, will be illustrated. Following this, the second part of the knowledge transfer model, which is the application of the extracted knowledge to the business user, will be explained in detail, this includes the description of the main components and algorithm of the RS.

3.1 Knowledge Transfer Model

The main idea behind this architecture emerged from the transferring of the knowledge from the power user (source of knowledge) to the business user (receiver of knowledge) [21]. The model of transferring knowledge consists of two phases. The first one is the externalization of the knowledge by conversion from tacit to explicit mode. The second phase is the internalization of the knowledge by conversion from explicit to tacit mode [21, 22]. The extraction and capturing of the knowledge of power users were already done: First, due to the shortcomings of the methods of analyzing the log files of web servers to extract the knowledge of the power user, a new tracing mechanism for the power users' interactions while they use the BI system was conceptualized and developed [23]. The interactions and their metadata are stored in trace files. Then, based on applying sequential pattern mining algorithms on the stored trace files, the analysis paths of the power users were extracted [24]. An analysis path is defined as a sequence of steps required to reach a specific goal (e.g. multidimensional view of business data) or to perform a specific data analysis within a BI system. In addition, analysis paths also include BI related metadata to enrich the business user with more information about what steps should be done within BI systems [24]. Analysis paths are stored in a knowledge repository. One of the main advantages of this approach is that the knowledge repository is remains up-to-date. As a consequence of power users use of the BI system, new analysis paths are always extracted and added to the knowledge repository. In the next sections, we will focus on the second phase of the knowledge transfer model, which is applying the extracted knowledge to business users through recommendations to help and empower them while using a BI system.

3.2 Conceptual Architecture

BI system have become more and more accessible via the web browser (web based BI system) and so are easily available for all employees an enterprise without the need of any installation on the client side. This influenced our conceptual architecture of the new self-service BI system. Accordingly, the new system is conceptualized and developed as a client-server architecture and typically consists of three layers, the presentation, the application (logic) and the data layer. Figure 1 depicts the architecture of the system and its components. On the left side of the figure, we can see the internet browser, which represents the access point of the BI system for power and business users. It was decided was to integrate the RS within the browser as a plug-in. In the following three sections, the three layers and the functionality of the architecture's components are explained.

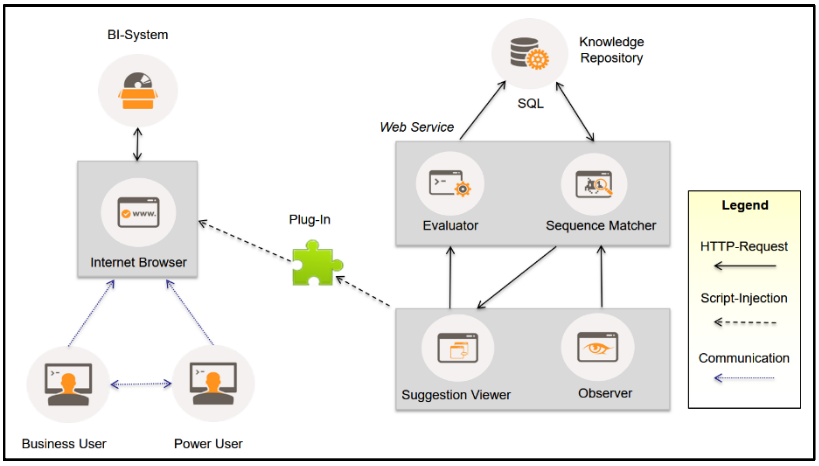


Figure 1. Conceptual architecture of the recommendation-based self-service BI

Presentation Layer. This RS layer is represented by the browser plug-in. This latter dynamically inserts the resources of the RS into the BI system so that the RS can be directly initialized in the browser by the business user to enable her/him to work with it interactively. Especially in the presentation layer, communication with the observer and the visualization of the recommendations play an important role. Furthermore, interaction options such as clicking on or evaluating a recommendation are also made in this layer. In this case, the interactions with the event listener are dynamically recorded and passed on to the application layer. The presentation layer essentially represents the user interface as well as the interaction possibilities of the RS. It includes the following main components:

- **Observer:** is responsible to trace the interactions with its BI metadata of the business user. After this, these are passed by a connector to the sequence matcher to be compared with the analysis paths from the knowledge repository. Observer aims to know what the business user is trying to do within the BI system.
- **Suggestion Viewer:** is responsible for the visualization of the resulting recommendations for the business user. It gets the recommendations list sorted by its priorities from the sequence matcher. After every interaction of the business user with the BI system, suggestion viewer first removes all previous recommendations and then adds the new recommendations to the current page dynamically sorted by their priorities. It also allows the business user to evaluate the last recommendation and to send the evaluation to the evaluator in the application layer.

Application Layer. This layer is represented by the web services and it is the link between the presentation layer and the data storage layer, in this case the knowledge repository. In this layer, the main logic and thus the core functions of the RS, which are the sequence matcher and the evaluator, are executed.

- **Sequence Matcher**: is responsible to match the interactions obtained from the observer with the analysis paths from the knowledge repository, and then to build a list of recommendations to be sent later to the business user. The function and the algorithm of the sequence matcher will be described in detail later in this section.
- **Evaluator**: is used to evaluate the recommendations so that future results can be generated better with the help of collected evaluations of the business users. The evaluator represents the evaluation module of the RS. At this point, the analysis paths suggested to the business user by the RS can be evaluated directly or indirectly. Indirect evaluation is done by clicking on a recommendation, while the direct evaluation takes place via a conscious evaluation by the business user who clicks on a recommendation to evaluate the last interaction performed with a thumb up or down. After the feedback for the recommendation made by the business user is prepared for the data layer, after that, the feedback is passed on to the knowledge repository.

Data Layer. The data storage layer of the recommender engine is represented by the knowledge repository. The data retention layer thus provides the RS with all the information necessary to generate new recommendations.

- **Knowledge Repository**: stores all analysis paths and their Metadata, which are already extracted from the power users. This component could be realized via a relational database.

Sequence Matcher Algorithm. Sequence matcher represent the heart of the RS, therefore, we will explain its functionality and the algorithm behind it in more detail. It is responsible for the search for matches between the analysis paths of the power users and the sequence of business user's interactions that are passed from the observer.

When the sequence matcher is called after a business user interaction, the call contains information about the current business user session, such as the user role, the dimension name, the support level, and the interaction sequences of the business user from the Observer. The support level specifies how many interactions must be matched between the analysis paths and the business user interactions. This helps to reduce the amount of recommendations when lot of analysis paths come from the knowledge repository as a result of the matching. Figure 2 illustrates the functionality of the support level and sorting of the recommendations. Each interaction of a business user with the user interface of the BI system is represented with a unique ID "PXY" like "P05".

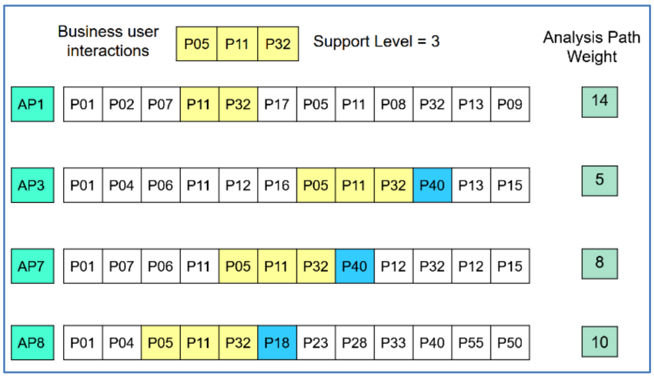


Figure 2. Recommendations matching and sorting algorithm

Figure 2 shows that the click sequence of the business user requires a support level of three interactions. In this example P05, P11 and P32 must occur in the analysis paths from the knowledge repository. If this is not the case, the analysis paths from the knowledge repository are not considered further; all other analysis paths and their meta-information are temporarily stored in an array for the further generation of recommendations. If all analysis paths from the knowledge repository have been searched successfully, the array with the analysis paths that have fulfilled the support level is analyzed and the entries are prioritized. This analysis and prioritization process will be described with the help of Figure 2 in the following steps:

1. **Same recommendations are summarized.** In this step, all recommendations will be combined into a recommendation that would return the same value. In Figure 2, it can be seen that the click sequence of the business user occurs in three analysis paths from the knowledge repository, which are AP3, AP7 and AP8. Two analysis paths would return the same recommendation "P40" and the third analysis path would have the recommendation "P18" (marked blue), In order to avoid a double recommendation for the user, the two analysis paths (AP3 and AP7) are combined

with the same recommendations in this step of the sequence matcher. Please note the Meta information, such as the AP weight (click frequency and business user's evaluation), which is used to prioritize the recommendations. For this reason, the AP weights are added when the same recommendations are combined, so that the weighting for the analysis paths remains summarized during the summarization process. The result of this step is that P40 will be recommended with an AP weight 5+8 =13. Then the step P18 will be recommended as second recommendation with the AP weight 10.

2. **The recommendations are sorted according to the call / click frequency.** In this step, the entries with the highest click frequency are sorted in descending order, since it can be assumed that entries with a high click frequency also provide particularly good results because they have already been called up several times by other business users. The click frequency is, on the one hand, a good indicator as to whether a click sequence has often been clicked on by business users, and on the other hand, it represents the individual evaluation of a business user as this user has the possibility to evaluate it directly after clicking on an analysis path. In the case of a positive evaluation, the analysis path is increased by one counter, while a negative value is reduced by two. The reduction by two counters is because the clicking frequency is already increased by clicking on the analysis path and this step is first undone after a negative evaluation and an additional counter has to be subtracted.
3. **The recommendations are sorted by Metadata.** If the recommendations according to step 2 still have the same prioritization because of the same click frequencies, it is checked whether there are any recommendations that have the metadata with the same cube or dimension names, with which the business user is currently interacting. If so, they are preferred because the probability is greater that the recommendations from the same cube are more appropriate for the business user. Moreover, it is also checked whether there are any recommendations made by the same user group (user role) under the same results. If this is the case, these are preferred since the probability is that recommendations from the same user group will be more adequate for the business user.

4 Prototypical Implementation and Evaluation

In this Section, the prototype and its technologies will be first explained, followed by the description of process flow of the RS. After that, a business scenario is defined to demonstrate the artifact based on the prototype in workshops to validate the utility and applicability of the artifact.

4.1 Prototype of the Recommendation based Self Service BI

Based on the conceptual architecture explained above, a prototype for recommendation based self-service BI was implemented as a proof of concept. Pentaho Business Analytics Platform was chosen because it is an open source BI system under the GPL license. Moreover, it includes different BI tools like reporting, dashboard and OLAP analytics, which can be accessed via a Web browser. For the implementation of the server-side functionalities, PHP web services were used for sequence matcher and the evaluator. In addition, knowledge repository was realized using MySQL database. The client application has been implemented as an extension for the Google Chrome browser. Moreover, the plugin itself was implemented using JavaScript, CSS3, HTML5 and JQuery. In addition, JSON has been used as a lightweight format for exchanging data between the Web server and Web pages.

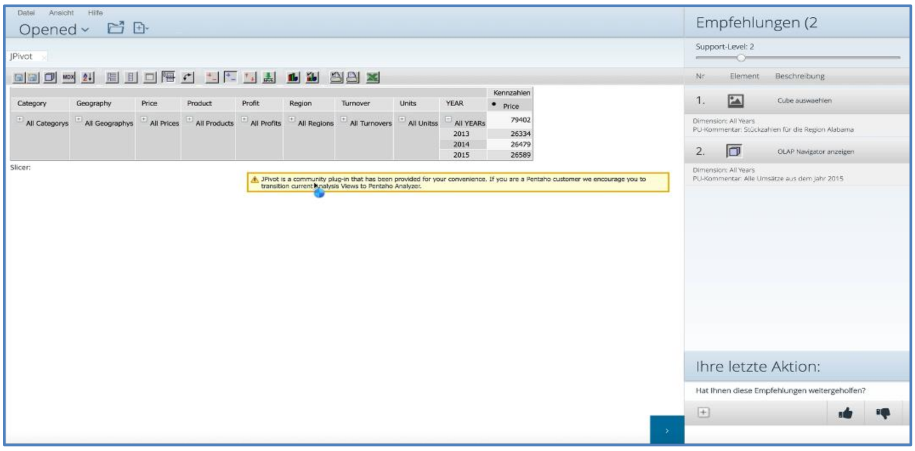


Figure 3. Recommender System for Pentaho Business Analytics

Process Flow: After the business user has logged-in to the BI system, the plug-in is called and all resources such as the frameworks used, the program code or the libraries are loaded and then integrated into the frontend of the BI system. As soon as the RS has been successfully loaded, the business user in the BI system receives a fixed area on the right side of the browser as we can see in Figure 3, which will be used to display the recommendations in the future. This part could be also hidden if the business user does not want to see it. As we can see at the top on the right side in Figure 3 a business user can change the support level that leads changing the number of steps are sent to the sequence matcher. Then, below this, we can see the recommendations sorted by the priority, as explained above. We see in each recommendation first the number of the recommendation followed by an element that is the same symbol as in the Pentaho Business Analytics and can be clicked, because it was implemented to allow the business user to call the function within the RS. Finally, we can see the description of the element. In the lower right hand of the figure, the business user can evaluate the previous recommendation with a thumb up or down.

4.2 Evaluation and Discussion

This research follows the design science research framework [25]. Therefore, the evaluation of this work and its utility and applicability was done based on the selection of the appropriate evaluation methods suggested by Hevner et al.. To achieve that, functional (black box) and structural (white box) testing of the artifact were done by performing various tests of the whole functionalities and interfaces of the prototype to ensure the correctness and consistency of the artifact. After that, an illustrative business scenario was defined to demonstrate the utility, applicability and the organizational impact of the artifact using the prototype. These were presented in form of workshops in two companies. The first was done with one of the directors in an IT company. The second workshop was addressed to two BI power users from a production company [25, 26].

Based on a fictive sales dataset that includes the following (product, category, city, region, country and year as dimensions and price, units, turnover and profit as facts), a multidimensional data model was created and loaded into Pentaho Business Analytics. The business scenario for the expert interviews was to perform data analysis to let the business users display the turnover of all products on a graph every year. As we already mentioned in Section 3, the first phase of our approach is to extract the analysis paths of the power users. This is done by training the system to extract the analysis paths of its power users that are stored in the knowledge repository. After that, the workshops were organized in three parts. Firstly, the new concept was presented in front of the experts. Secondly, based on the business scenario, the prototype was demonstrated to show how the business user is guided by providing him with the appropriate recommendations to accomplish his data analysis. Finally, using a questionnaire-based expert interview, the architecture and its prototype was evaluated.

In terms of the utility of the new approach, the feedback that we got in the first workshop was that the recommendations were user-friendly and represented to support to the business user while performing data analysis. In addition, this will improve the analytical skills of business users every time they use the system. In the second workshop, the experts said that business users always forget how to get the information they need, even if they already got the information before from the power users. Therefore, the experts concluded that our approach could help business users to do their ad-hoc analysis by themselves without asking the power users every time.

In term of the applicability of the approach, the director confirmed the applicability of the artifact and said this can offer several potentials. In the second workshop, the application of the artifact in their company was intensively discussed. The conclusion was that this approach could be applicable, but with extra operational costs, because our prototype is implemented using Pentaho Business Analytics and it cannot be transferred one to one to their company. This requires analyzing the interfaces of their BI system to adjust our prototype to suit their BI system. We were aware of this point from the beginning, but in the research, the prototype was implemented with the goal of demonstrating the new approach in many companies regardless what specific BI system they use. The conceptual architecture can be instantiated for every Web-based BI system.

Regarding the improvement of the recommendations: In the first workshop, the expert suggested that the RS could be extended by enabling business users to write comments for other business users on specific recommendations to enhance the functionality of the system. In the second workshop, based on their experience with the business users, the experts suggested improving the system with a search functionality to allow business users to find a specific result. These recommendations and suggestions from the two workshops will be discussed and considered in our future work.

5 Conclusion

In this paper, we have presented a novel approach to realize the concept of self-service BI. It is designed to transfer the knowledge of power users to business users via recommendations. Such recommendations guide the business users while they perform their own ad-hoc data analysis to empower them to be more self-reliant and less depending on power users or IT department. This will help enterprises to improve the usage rate and the adoption of their BI systems. The conceptual architecture was described in details including the algorithm of generating the recommendations. Furthermore, the prototype and a business scenario were used in demonstration and evaluation phases to firstly illustrate the functionalities of the new approach and secondly to evaluate its results. The evaluation was done using a questionnaire-based expert interview. The target group of this interview were experts in two companies in IT and production domains. The feedback gained from the answers of the experts in the interviews will be incorporated in future work. Finally, we plan to evaluate this work in further companies and to obtain more feedback to enable the improvement of our artifact and to prove its applicability to a wide variety of BI companies.

References

1. Kemper, H.-G., Rausch, P., Baars, H.: Business Intelligence and Performance Management: Introduction. In: Rausch, P., Sheta, A.F., Ayesh, A. (eds.) Business intelligence and performance management. Theory, systems and industrial applications / Peter Rausch, Alaa F. Sheta, Aladdin Ayesh, editors, pp. 3–10. Springer, London (2013)
2. Gluchowski, P., Chamoni, P.: Analytische Informationssysteme. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
3. Vasconcelos, Andre Ferreira Ferrao Couto E., Bras, T.J.G.: Information Reference Architecture for the Portuguese Health Sector. In: 2015 IEEE 17th Conference on Business Informatics (CBI), pp. 72–79 (2015)
4. Hawking, P., Sellitto, C.: Business Intelligence Strategy: A Utilities Company Case Study. Int. J. Enterp. Inf. Syst. 11, 1–12 (2015)
5. Wixom, B., Watson, H.: The BI-Based Organization. In: Herschel, R.T. (ed.) Organizational Applications of Business Intelligence Management, pp. 193–208. IGI Global (2012)
6. Susan Moore: Gartner Says Worldwide Business Intelligence and Analytics Market to Reach $16.9 Billion in 2016, http://www.gartner.com/newsroom/id/3198917
7. James Richardson: Business Intelligence Adoption Trends (2011)
8. Eckerson, W.: Business-Driven BI. Using New Technologies to Foster Self-Service Access to Insights (2012)
9. BARC Research: Pervasive Business Intelligence: Still a Dream. The BI Survey 15 (2016)
10. T. E. Yoon, B. Ghosh, Bong-Keun Jeong: User Acceptance of Business Intelligence (BI) Application: Technology, Individual Difference, Social Influence, and Situational

Constraints. In: System Sciences (HICSS), 2014 47th Hawaii International Conference on, pp. 3758–3766 (2014)
11. Grublješič, T., Jaklič, J.: Business Intelligence Acceptance. The Prominence of Organizational Factors. Information Systems Management 32, 299–315 (2015)
12. Mertens, M.: KNOBI - knowledge-based business intelligence for business user information-self-service. OlWIR, Oldenburger Verl. für Wirtschaft, Informatik und Recht, Edewecht (2013)
13. Alpar, P., Schulz, M.: Self-Service Business Intelligence. Bus Inf Syst Eng 58, 151–155 (2016)
14. Daradkeh, M., Al-Dwairi, R.M.'d.: Self-Service Business Intelligence Adoption in Business Enterprises. Int. J. Enterp. Inf. Syst. 13, 65–85 (2017)
15. Imhoff, C., White, C.: Self-Service Business Intelligence. TDWI Best Practices Report, Third Quarter (2011)
16. BARC Research: BI Trend Monitor 2017. The world‘s largest survey on BI trends (2017)
17. Ricci, F., Rokach, L., Shapira, B.: Recommender Systems Handbook. Springer US, Boston, MA (2015)
18. Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., Koper, R.: Recommender Systems in Technology Enhanced Learning. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook, pp. 387–415. Springer US, Boston, MA (2011)
19. Zaiane, O.R.: Building a recommender agent for e-learning systems. In: Ed, K. (ed.) International conference on computers in education, pp. 55–59. IEEE Comput. Soc (2002)
20. Khribi, M.K., Jemni, M., Nasraoui, O.: Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval. In: 2008 Eighth IEEE International Conference on Advanced Learning Technologies, pp. 241–245
21. Liyanage, C., Elhag, T., Ballal, T., Li, Q.: Knowledge communication and translation – a knowledge transfer model. Journal of Knowledge Management 13, 118–131 (2009)
22. Sulaiman, S., Mahmoud, T., Marx Gómez, J., Kurzhöfer, J.: Automatic Knowledge Transfer-based Architecture towards Self-Service Business Intelligence. In: The Seventh International Conference on Advances in Databases, Knowledge, and Data Applications, p. 90 (2015)
23. Sulaiman, S., Mahmoud, T., Robbers, S., Marx Gómez, J., Kurzhöfer, J.: A Tracing System for User Interactions towards Knowledge Extraction of Power Users in Business Intelligence Systems. In: 8th International Conference on Knowledge Management and Information Sharing, pp. 199–207
24. Sulaiman, S., Marx Gómez, J.: Application of Sequential Pattern Mining to Extract the Knowledge of the Power Users in Business Intelligence Systems. In: Eureka International Virtual Meeting Eureka 2016 OPTISAD 2016 (2016)
25. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. Mis Quarterly 28, 75–105 (2004)
26. Sonnenberg, C., Vom Brocke, J.: Evaluations in the Science of the Artificial - Reconsidering the Build-Evaluate Pattern in Design Science Research. In: Peffers, K., Rothenberger, M., Kuechler, B. (eds.) Design Science Research in Information Systems. Advances in Theory and Practice: 7th International Conference, DESRIST 2012, Las Vegas, NV, USA, May 14-15, 2012. Proceedings, pp. 381–397. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)

CSCW & Social Computing

Teilkonferenzleitung

Michael Koch
Michael Prilla
Alexander Richter

Coordination Across Open Source Software Communities: Findings from the Rails Ecosystem

Alexander Eck

University of St.Gallen, Institute of Information Management, St.Gallen, Switzerland
alexander.eck@unisg.ch

Abstract. While coordination of work within open source software (OSS) communities is well-researched, it is virtually unknown how work is coordinated across community boundaries. However, as OSS projects are often part of a larger digital ecosystem of interdependent artifacts and communities, cross-community coordination is a pertinent topic. We turn to the ecosystem around Ruby on Rails to empirically explore this research gap. To this end, we scrutinize 96 coordination episodes among five interrelated OSS projects and identify four cross-community coordination mechanisms: adaptation, upgrading, positioning, and departure. Each mechanism describes a distinct and stable arrangement to integrate contributions across community borders. After presenting our findings, we reason about the significance of the results on explaining generative change in digital ecosystems.

Keywords: cross-community coordination, open source software, digital ecosystems, generative change, case study

1 Introduction

The open source software (OSS) model of designing and changing complex artifacts based on a publicly accessible codebase [1] has created new forms of coordinating which have been investigated thoroughly during the last two decades [2]. Despite this work, little is known about coordination *across* communities of interdependent artifacts, as opposed to coordination *within* the community around an individual OSS artifact [3]. Cross-community coordination is of practical importance, because many OSS artifacts are coupled with other artifacts to form larger digital ecosystems and as such their communities have to interact [4]. One example is the ecosystem around *Ruby on Rails*, a web application framework. Thousands of add-on modules extend utility of this artifact [5]. Cross-community coordination is also of theoretical relevance, as it provides an apt setting to study the phenomenon of generative change, that is unanticipated change produced by contributions from broad and varied audiences [6], which is believed to drive innovation in the digital age [7]. Therefore, we ask:

How do open source software communities coordinate with other communities to change their respective artifacts within a shared ecosystem?

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

OSS communities usually introduce change through granular, incremental steps [8]. We harness this mode of working to empirically inquire coordination episodes. By recognizing patterns among these granular episodes, we identify mechanisms with which actors from different communities integrate their respective contributions – that is, we identify cross-community coordination mechanisms that lead to change, or potential for change, in OSS artifacts.

Our findings from five communities that are part of the *Ruby on Rails* ecosystem suggest four mechanisms. First, cross-community coordination reflects the hierarchy between artifacts. If one artifact serves as the basis of the second artifact, it is likely that the community of the latter readily seeks to *adapt* to any relevant change of the outside artifact. Second, communities anticipate fundamental changes to an artifact they depend on by contemplating possible consequences and *upgrading* their artifact accordingly. Third, cross-community coordination often revolves around locating and eliminating a design flaw that is identified through artifact coupling. The communities work together to *position* the appropriate location for fixing the design flaw. Fourth, members of an outside community may argue for a *departure* from existing artifact design, based on experiences within the context of their own artifact.

The remainder of this paper is organized as follows. We review the relevant literature on coordination across open source software communities. We then present the research design and findings of the empirical study, followed by a discussion how the findings fit into the broader topic of generative change in digital ecosystems. The paper concludes with a summary of main limitations and directions for further research.

2 Background

2.1 Coordination, Coordination Episodes, and Coordination Mechanisms

Coordination refers to the extent to which actors that need to integrate their respective contributions due to interdependencies do so consistently and coherently [9]. A major purpose of an *OSS project*, i.e. an OSS artifact and the community supporting it, is to advance artifact design via source code contributions [10]. Therefore, a large part of coordination consists of managing change in software artifacts through resolving interdependencies that emanate from source code contributions [11]. Consequently, this paper investigates coordination episodes during which possible source code contributions are discussed and integrated. A *coordination episode* is a logically connected series of activities with a trigger-activities-resolution structure [12]: a *trigger* that prompts coordination need; sequential *coordination activities* between actors during which knowledge is exchanged and source code contributions are discussed; and an eventual *resolution* whether and what kind of change is implemented. We can expect that over time, OSS communities develop stable arrangements to coordinate, or distinct *coordination mechanisms* [13]. It is reasonable to believe that coordination mechanisms exist also when two OSS communities repeatedly interact to integrate contributions across community boundaries, which makes it possible to identify mechanisms for cross-community coordination.

2.2 Coordination Mechanisms Across Communities of Open Source Software

Often, a software artifact is coupled with other artifacts to enable and augment its own capabilities [e.g., 14]. As such, it is typically part of a larger *digital ecosystem*, or a set of interdependent, co-evolving artifacts and the communities supporting them [4]. Artifacts belonging to an active ecosystem commonly change, which induces the regular need to integrate contributions originating from outside communities. Such *cross-community coordination* is difficult, because the OSS projects making up a digital ecosystem are heterogeneous and distributed, both in terms of resources and in terms of control [cf. 15]. For example, social practices how artifacts are changed may differ between communities, and whether an outside artifact changes so that depending artifacts are forced to be adapted is ultimately beyond the control the focal community.

Extant research on cross-community coordination mainly highlights the role of boundary spanners, i.e. actors who participate in multiple projects and who are capable to integrate heterogeneous contributions and reconcile dependencies between distributed artifacts [16–19]. Hence, boundary spanners coordinate by leveraging their personal knowledge, not through processes that span actors in disparate communities.

Beyond the attention given to the boundary spanning role, there is little research on cross-community coordination. This might be attributed to the property of software artifacts to be loosely coupled via defined interfaces [cf. 20]. Because interfaces codify the rules and protocols of exchange, they commonly avoid the need for direct interaction [21]. Yet, interfaces neither resolve all interdependencies nor do they remain unchanged over time [cf. 22]. For those situations in which coordination is necessary, we found weak cues for two distinct coordination mechanisms in the literature, which we label *adaptation* and *departure*. First, a community may adapt its artifact when an interdependent outside artifact changes, which is a rather transactional mechanism: a community accepts the change produced by another community and adapts accordingly [22–24]. Second, a more involved process is to advocate a departure from existing artifact design. This mechanism exploits the affordance of OSS communities to receive contributions from outsiders: a community influences another one to deviate from its existing artifact design and implement a change, usually one that helps the outsiders themselves [22]. Overall, our knowledge of cross-community coordination is limited. In the following, we report on an exploratory case study, addressing this research gap.

3 Research Design and Methods

We conducted an explanatory multiple case study of cross-community coordination in the digital ecosystem around Ruby on Rails (or simply *Rails*), an artifact that simplifies development of web applications written in the Ruby programming language. The goal of this examination was to describe and explain how disparate communities coordinate to change their OSS artifacts in light of interdependencies. Our choice of the Rails ecosystem was based on three main considerations. First, Rails was designed to be coupled with other artifacts written in the Ruby programming language – *Ruby gems* in jargon – which created the potential for cross-community coordination needs. Second, with its history reaching back to 2004 [25], both the codebase and the community of

Rails, and in extension many related projects, can be assumed to be mature. This was important to us, because we did not intend to trace the dynamics of how coordination mechanisms emerge, but rather identify established coordination mechanisms. Third, through reconstructing many digital ecosystems, including the Rails ecosystem [26], we noticed that the Rails artifact maintained an unusually large number of couplings with other artifacts. This made Rails an ideal starting point to investigate cross-community coordination in OSS settings.

3.1 Data Collection

We selected a sample of five OSS projects that are part of the Rails ecosystem, listed in Table 1, based on four major considerations. First, the project had to be hosted on Github, a popular service for collaborative OSS development, as we collected empirical data from there. Second, we wanted to ensure that the projects had many contributors, measured by *Github forks* [27], to increase the chances of observing coordination. Third, we chose popular projects, measured by *Github stars* accumulated [28]. As an added benefit, these criteria ensured that the selected projects were relatively mature, because it takes time to accumulate a sizable number of forks and stars.

Table 1. Artifacts included in the case study

Artifact	*Description*	*Forks*	*Stars*	*Ep*	*First activity*	*Last activity*
Ruby on Rails (04 Nov 2008)	Framework for web application development	13,073	32,195	74	24 Aug 2009	06 Jul 2016
Devise (16 Sep 2009)	Provides advanced authentication functionality	3,499	15,576	38	05 Apr 2010	29 Jun 2016
Activeadmin (15 Apr 2010)	Simplifies creation of admin interfaces for web apps	2,497	6,763	41	01 Aug 2011	11 Jul 2016
Kaminari (06 Feb 2011)	Provides pagination functionality for web apps	779	5,909	17	04 Mar 2011	08 Jan 2016
Formtastic (07 Apr 2008)	Domain-specific language for designing web-based forms	604	4,894	22	24 Aug 2009	11 Jul 2016

Artifact — *Artifact name (in brackets: day artifact became available on Github)*
Forks — *Number of forks, a measure of contributing community size (as of 15 Jul 2016)*
Stars — *Number of stars, a measure of artifact popularity (as of 15 Jul 2016)*
Ep — *Total number of episodes with artifact participation (after exclusion criteria)*
First/last activity — *Start day of first episode and end day of last episode with artifact participation*

Fourth, based on our knowledge of projects belonging to the Rails ecosystem [26] and employing self-written software, we systematically searched the discussion threads of projects in the Rails ecosystem, with data ranging from 29 Oct 2007 to 01 Jul 2016. Specifically, we captured discussions during which an outside project was referenced, and took this referencing as cue that two artifacts were coupled, which potentially led to the need of repeated cross-community coordination [3]. Consider this example: In June 2011, an issue on the *Devise* discussion board triggered a coordination episode: *"I have just updated devise 1.4.0 => 1.4.1 in my project, and now it throws an exception"*. After some causal theorizing, the root cause was identified: *"it's broken by* *this commit in Rails: rails@0ca69ca"*. Notably, a reference to Rails (that follows a well-defined syntax) was included. Less than two hours after the initial issue description, Devise was adapted to accommodate the change made by the Rails community, which resolved the episode.

We counted the number of outside references, and retained pairs of projects that shared many references. For example, we counted 109 instances in which either Rails or Devise referenced the other project. In total, the five OSS projects that we included in the case study shared 394 references between them. We then manually inspected the full discussion threads that contained those references. A discussion thread on Github typically revolves around a specific and usually self-contained *issue* for which coordination is needed to sort it out [29]. In our data sample, we observed that an issue discussion may spill over to another project, or that it may span multiple threads within the same project. Therefore, we followed the activities pertaining an issue even if it spanned multiple discussion threads and grouped them together. Overall, we collected 245 grouped discussion threads, which we regarded as tentative empirical evidence for cross-community coordination episodes, our unit of analysis. To ensure that the data indeed represented episodes of cross-community coordination, we inductively elaborated a list of exclusion criteria. For example, we excluded an episode if it was still ongoing at time of data collection. After applying the exclusion criteria, we were left with 96 episodes of cross-community coordination.

3.2 Construct Operationalization

The selected episodes served as empirical basis for analyzing how cross-community coordination is carried out in the Rails ecosystem and what effect it has on the artifacts involved. Our unit of analysis were individual coordination episodes. Therefore, we needed to operationalize the trigger of an episode, individual coordination activities, and the eventual episode resolution, as summarized in Table 2.

To operationalize the *trigger* construct we relied on open coding and axial grouping [30], eliciting five possible triggers from our data.

As for *coordination activities*, we drew upon the coding book of [29], which distinguishes between knowledge integration and direct implementation. *Knowledge integration* comprises activities needed to compile and integrate knowledge among the involved actors. *Direct implementation* covers activities needed to create and evaluate source code that may change the artifact. Just as in [29] our data set showed other activities as well, such as acknowledgment (e.g., *"You, sir, are awesome."*) which we did not analyze as they were not material to our research question.

Each coordination activity was performed by an actor, whose *actor status* we classified as either insider or outsider following a simple heuristic: We counted how often an actor contributed to each community across our full data set and made her member of the community to which she contributed most. Then we compared assigned membership of this actor (e.g., Rails) with the community in which she performed a coordination activity (e.g., Devise). Matching pairs were coded as *insider*, mismatches were coded as *outsider*.

Table 2. Operationalization of constructs and number of data points per construct

Construct	*Operationalization*	*Count*
Trigger	Identified trigger: issue, not offering a possible root cause (N=16); issue, offering a possible root cause (49); anticipated outside change may require change in own artifact (18); outside artifact offers new functionality (3); change proposal (10).	96
Coordination activity	Discussion entry that was coded as either knowledge integration (709) or direct implementation (599).	1,308
Actor status	Actor was either insider (1,178) or outsider (130) of the community in which she performed the coordination activity.	1,308
Resolution	Identified resolution: business rules change (47); software properties change (29); workaround found (5); no change (15).	96
Changed artifact	Episode was resolved by changing: …upstream artifact (16); …downstream artifact (53); …both artifacts (7); none (20).	96
Episode length	Number of activities in a coordination episode.	96
Episode duration	Duration of coordination episode.	96
Actors involved	Number of different actors who participated in an episode.	96
Files involved	Number of files for which source code changes were discussed.	96
KI ratio	Share of knowledge integration activities per episode.	96
OA ratio	Share of activities performed by outsiders per episode.	96

KI: knowledge integration
OA: outsider activity

A coordination episode in our context resolves by (not) implementing a source code change. To describe the *resolution* in more detail we were interested to learn also about the kind of (non-)change that resolved the episode. Therefore, we drew upon [31], who distinguish two types of source code change: change of *business rules* are changes to the behavior of a software artifact such as fixing a design flaw, whereas change of *software properties* alters nonfunctional metrics of a software artifact such as increasing the readability of source code. Furthermore, we knew from prior literature that agreeing on a *workaround*, i.e. a temporary and localized solution that does not change the artifact, was yet another way to resolve a coordination episode [32].

In addition, we captured which of the two artifacts involved in a coordination episode were changed: upstream, downstream, both, or none, following common terms in software engineering [e.g., 22]. An *upstream* artifact is one on which others depend, while the depending artifact is called *downstream*. From manual inspection of the five artifacts in our case study we established the upstream/downstream relationships, depicted in Figure 1. For example, Rails is an upstream artifact to Devise. By analogy, the Rails community is upstream to the Devise community.

We elicited distinct cross-community coordination mechanisms through identifying patterns across coordination episodes. To support this process we selected six additional constructs that characterized the episodes as a whole [cf. 12]. *Episode length* captures the number of coded activities in a coordination episode. *Episode duration* measures how much time passed between the first and the last activity in a coordination episode. *Actors involved* counts how many different actors were involved in an episode. *Files* *involved* counts the number of files for which source code changes were discussed during the episode. *Knowledge integration ratio* sets the number of knowledge integration activities in relation to the number of all activities in an episode. Finally, *outsider activity ratio* measures the percentage of activities that outsiders performed.

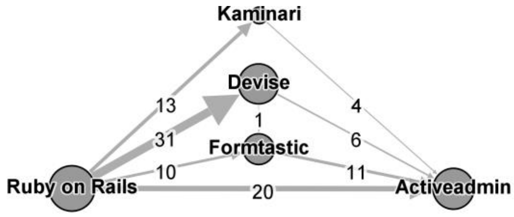


Figure 1. Artifact dependencies and number of episodes per artifact pair

3.3 Data Analysis

Data analysis consisted of two logical stages of coordination episode coding and coordination mechanism identification, through which we moved back and forth as we learnt more about the empirical setting. First, we *coded coordination episodes* according to the scheme summarized in Table 2. It struck us how straightforward the coding scheme could be applied. For example, when source code was changed it was easy to see whether functionality was added or fixed, readability was increased, etc. We attribute this to two reasons: In OSS communities it is customary to carry out fine-grained activities so that other community members can easily comprehend them [8]. In addition, the coordination episodes were usually focused on solving the specific issue at hand and did not divert into other directions. This demonstrated a certain level of discipline across the studied communities, which was an indicator that stable arrangements to coordinate, i.e. coordination mechanisms, existed.

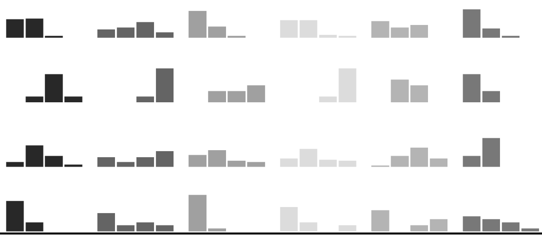
Second, we *identified coordination mechanisms* in cross-community settings by seeking commonalities and differences between the individual episodes and matching them with preliminary coordination mechanism descriptions. We would switch to this stage of analysis whenever we identified tentative patterns or peculiarities that were potentially helpful to meet our research objective. As prior research mentioned *adaptation* and *departure* mechanisms, these were our natural starting points for candidate mechanisms, for which we wrote initial descriptions. We found evidence for these mechanisms and discovered two additional coordination mechanisms, which we eventually labeled *upgrading* and *positioning*.

4 Findings: Mechanisms of Cross-Community Coordination

Through analyzing 96 episodes of cross-community coordination, our empirical analysis unearthed four mechanisms that could comprehensively explain the observed coordination work. Summary statistics for each mechanism is provided with Table 3. In what follows, we show how each mechanism describes a different arrangement of cross-community coordination.

Table 3. Cross-community coordination mechanisms

	Trigger		*Activities*						*Resolution*		
Mechanism	*N*	*I*	*Length*	*Duration*	*Actors*	*Files*	*KI ratio*	*OA ratio*	*C*	*B*	*U*
Adaptation	43	72%							91%	42%	0%
Upgrading	7	14%							86%	43%	0%
Positioning	33	100%							91%	76%	67%
Departure	13	0%							8%	8%	8%



N	Number of episodes attributed to mechanism
I	Percentage of episodes that were triggered by an issue (with or without root cause analysis)
Length	Number of activities per episode (histogram categories: 2-5; 6-20; 21-50; 51-92)
Duration	Duration per episode in days (<1; 1-10; 10-100; 100-1,297)
Actors	Number of different actors involved per episode (2-3; 4-8; 9-16; 17-24)
Files	Number of source code files involved per episode (<2; 2-4; 5-9; 10-87)
KI ratio	Share of knowledge integration activities per episode (<25%; 25%-50%; 50%-75%, >75%)
OA ratio	Share of outsider activities per episode (<5%; 5%-50%; 50%-95%, >95%)
C	Percentage of episodes that resolved with an artifact change
B	Percentage of episodes that resolved with a change of business rules
U	Percentage of episodes with change in upstream artifact

4.1 The Adaptation Mechanism

Our data analysis shows how one community seeks to adapt its artifact to changes in an upstream artifact as to restore functionality or utilize new functionality, which we define as *adaptation* mechanism. Typically, a coordination episode of this type is triggered by a malfunction, and its root cause is found to be a recent change in an upstream artifact. This instigates change in the downstream artifact, which resolves the initially reported issue. The adaptation mechanism leverages hierarchy of artifact couplings as its logic to coordinate work across communities. The upstream community introduces changes without reaching out to downstream communities. Instead, it assumes that downstream artifacts will be adapted to restore compatibility. This is an efficient arrangement for the upstream community, because it mitigates the costs of verbal exchange with outside communities. It is also efficient for the downstream communities, given that most adaptation episodes resolve after few activities and involve few actors. Furthermore, the low number of files involved in an adaptation episode indicates that the necessary adaptations are rather small, which in turn implies that the changes to the upstream artifact cannot be extensive as well.

4.2 The Upgrading Mechanism

Our data analysis reveals how one community seeks to understand major changes in an upstream artifact and how it upgrades the downstream artifact to appropriate these changes, which we define as *upgrading* mechanism. Typically, a coordination episode of this type is triggered by upcoming major changes (e.g., changes to interfaces) that the community of an upstream artifact announced. After discussing the implications for its own artifact, the downstream community systematically changes its artifact to account for these implications. The upgrading mechanism resembles the adaptation mechanism, as it follows the same logic of artifact hierarchy. However, it is distinct in that coordination activities typically commence in anticipation of change to the upstream artifact, as opposed to a usually reactive adaptation episode. This mechanism is efficient for the upstream community, as indicated by the very low share of outsider activities. Thus, the coordination burden is mostly on the downstream community. Not only must it implement many changes to its artifact, it also has to spend effort on collating and integrating knowledge, as evidenced by the comparatively high knowledge integration ratio of most upgrading episodes.

4.3 The Positioning Mechanism

Our data analysis surfaces how two communities seek to track down an issue observed from coupling their respective artifacts and introduce change at the position they jointly agree on, which we define as *positioning* mechanism. Typically, a coordination episode of this type is triggered by an issue that is observed from the combination of upstream and downstream artifacts. Both communities then investigate why the issue happens, debate which artifact should be changed and implement the changes deemed appropriate. The positioning mechanism is distinct as it always sees one community reaching out to another due to an issue it cannot solve on its own. Creating a shared understanding between the two communities is a major concern, suggested by the comparatively high knowledge integration ratio combined with a high outsider activities ratio. A positioning episode mainly increases artifact quality, because a previously unknown issue is identified and resolved: 76% of all changes were to the functionality of the artifact, usually fixing a design flaw. Remarkably, the main beneficiary is the upstream artifact, as 67% of all positioning episodes in our data resolved with changes to the upstream artifact. This suggests that a growing pool of downstream artifacts effectively increases the number of collaborators who contribute with technical expertise to the upstream artifact, which extends the community borders.

4.4 The Departure Mechanism

Our data analysis exposes how an actor, based on experiences made with a coupled artifact, proposes a change and the focal community decides whether and how to implement the change despite it being a departure from existing design, which we define as *departure* mechanism. Typically, a coordination activity of this type has few activities, involves just a couple of actors, and no or just one source code file. Usually, the initially proposed departure of artifact design is not accepted by the focal community. The departure mechanism incorporates what is commonly considered a hallmark of open source software: anybody may modify an OSS artifact and see this change become part of the ‘official release’ if the community deems it beneficial [3]. However, in our case study of five rather mature OSS projects, fewer than 15% of all episodes could be attributed to the departure mechanism. What is more, all episodes but one resolved with a rejection of the proposed change by the focal community, and typically did so after few coordination activities. This suggests that OSS communities managing mature artifacts are reluctant to depart from existing artifact design.

5 Discussion: Coordination as Means of Generative Change

By analyzing how the communities of five interdependent OSS projects in the Rails ecosystem resolve coordination needs stemming from changes in coupled artifacts, we identified four cross-community coordination mechanisms, namely *adaptation*, *upgrading*, *positioning*, and *departure*. Each mechanism describes a mode of integrating contributions from coupled OSS projects that instigate change in the focal artifact. Thus, cross-community coordination organizes the co-evolutionary production of change caused by broad and varied audiences absent of central control, that is cross-community coordination is a means of *generative change* in a digital ecosystem [6].

The strength of a generative system lies in its capacity to allow contributions from heterogeneous actors, and to turn these contributions into productive changes. Because they are the result of a creative dialogue between insiders and outsiders, some of these changes are unanticipated and innovative [6]. Due to perpetual incompleteness of digital artifacts [33] and their ability to be coupled with other incomplete artifacts to form new combinations and variations, digital ecosystems produce constant co-evolutionary changes. Arguably, this dynamic is a main driver of digital innovation [7].

To this end, the identified cross-community coordination mechanisms help us better explain and rationalize digital innovation processes in complex digital ecosystems, particularly in an OSS setting. The adaptation and upgrading mechanisms show how upstream OSS projects effectively influence the evolutionary trajectory of downstream projects. In turn, the positioning mechanism demonstrates how autonomous, yet interdependent OSS communities create synchronization opportunities to sort out design flaws for mutual benefit. Finally, the departure mechanism suggests that mature OSS projects have a strong sense of direction, which reduces the propensity of radical design changes. Overall, the existence of coordination mechanisms that operate across OSS communities are further proof of the self-organizing capability of heterogeneous systems characterized by distributed resources and distributed control [15].

6 Limitations and Conclusion

The contributions presented here are naturally limited by our research approach and the scope of our data, which points at opportunities for future research. First, we cannot claim generalizability of the findings beyond the studied context. Therefore, additional research that studies different empirical contexts might complement our results. Second, stemming from our selection of mature communities, we did not regard path dependencies between the individual episodes. Future research might identify dynamic aspects of cross-community coordination. Third, we did not observe coordination activities directly, but derived them from verbalized contributions on discussion boards.

Hence, future research could employ ethnographic methods to draw a richer picture of the micro-foundations of cross-community coordination. Fourth, we obtained data only from the discussion boards on Github. Other tools to coordinate work are available, such as email, chat, and (virtual) whiteboards. As such, further examinations could tap into alternative data sources to identify coordination mechanisms that we did not observe.

In summary, much of the existing research on cross-community coordination assumes that boundary spanners and loose coupling are sufficient to resolve interdependencies between OSS projects. In this paper, we employed an explanative case study to elicit four mechanisms for coordinating work across interdependent OSS communities, which is the main accomplishment of this paper. In addition, we argued that cross-community coordination mechanisms are means of generative change in digital ecosystems, and thus help us better explain processes of digital innovation in heterogeneous systems.

References

1. Raymond, E.: The Cathedral and the Bazaar. Know Techn Pol 12, 23–49 (1999)
2. Crowston, K., Wei, K., Howison, J., Wiggins, A.: Free/Libre Open-Source Software Development: What We Know and What We Do not Know. ACM Computing Surveys 44, Article 7 (2012)
3. Faraj, S., von Krogh, G., Monteiro, E., Lakhani, K.R.: Special Section Introduction—Online Community as Space for Knowledge Flows. Information Systems Research 27, 668–684 (2016)
4. Selander, L., Henfridsson, O., Svahn, F.: Capability Search and Redeem Across Digital Ecosystems. J Inf Technol 28, 183–197 (2013)
5. RubyGems: Gems, https://rubygems.org/gems (Accessed: August 12, 2017)
6. Zittrain, J.L.: The Future of the Internet and How to Stop It. Yale University Press, New Haven, CT (2008)
7. Yoo, Y.: The Tables Have Turned: How Can the Information Systems Field Contribute to Technology and Innovation Management Research? JAIS 14, 227–236 (2013)
8. Howison, J., Crowston, K.: Collaboration Through Open Superposition: A Theory of the Open Source Way. MIS Quarterly 38, 29–50 (2014)
9. Faraj, S., Xiao, Y.: Coordination in Fast-Response Organizations. Management Science 52, 1155–1169 (2006)
10. Grewal, R., Lilien, G.L., Mallapragada, G.: Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems. Management Science 52, 1043–1056 (2006)
11. Crowston, K., Li, Q., Wei, K., Eseryel, U.Y., Howison, J.: Self-organization of Teams for Free/Libre Open Source Software Development. Information and Software Technology 49, 564–575 (2007)
12. Annabi, H., Crowston, K., Heckman, R.: Depicting What Really Matters: Using Episodes to Study Latent Phenomenon. ICIS 2008 Proceedings (2008)
13. Okhuysen, G.A., Bechky, B.A.: Coordination in Organizations: An Integrative Perspective. The Academy of Management Annals 3, 463–502 (2009)
14. Haefliger, S., von Krogh, G., Spaeth, S.: Code Reuse in Open Source Software. Management Science 54, 180–193 (2008)

15. Lyytinen, K., Yoo, Y., Boland, R.J.: Digital Product Innovation Within Four Classes of Innovation Networks. Information Systems Journal 25, 47–75 (2016)
16. Daniel, S., Stewart, K.: Open Source Project Success: Resource Access, Flow, and Integration. The Journal of Strategic Information Systems 25, 159–176 (2016)
17. Singh, P.V.: The Small-World Effect: The Influence of Macro-Level Properties of Developer Collaboration Networks on Open-Source Project Success. ACM Transactions on Software Engineering and Methodology (TOSEM) 20, Article 6 (2010)
18. Vasilescu, B., Blincoe, Kelly, Xuan, Qi, Casalnuovo, C., Damian, D., Devanbu, P., Filkov, V.: The Sky is Not the Limit: Multitasking on GitHub Projects. ICSE 2016 Proceedings, 994–1005 (2016)
19. Weiss, M., Moroiu, G., Zhao, P.: Evolution of Open Source Communities. OSS 2006 Proceedings, 21–32 (2006)
20. Baldwin, C.Y., Clark, K.B.: Design Rules: The Power of Modularity. MIT Press, Cambridge, Mass. (2000)
21. Ghazawneh, A., Henfridsson, O.: Balancing Platform Control and External Contribution in Third-party Development: The Boundary Resources Model. Information Systems Journal 23, 173–192 (2013)
22. Bogart, C., Kästner, C., Herbsleb, J., Thung, F.: How to Break an API: Cost Negotiation and Community Values in Three Software Ecosystems. FSE 2016 Proceedings, 109–120 (2016)
23. Decan, A., Mens, T., Claes, M., Grosjean, P.: When GitHub Meets CRAN: An Analysis of Inter-Repository Package Dependency Problems. SANER 2016 Proceedings, 493–504 (2016)
24. Gonzalez-Barahona, J.M., Robles, G., Michlmayr, M., Amor, J.J., German, D.M.: Macro-level Software Evolution: A Case Study of a Large Software Compilation. Empir Software Eng 14, 262–285 (2009)
25. Hansson, D.H.: Ruby on Rails, http://rubyonrails.org/ (Accessed: August 10, 2017)
26. Eck, A., Uebernickel, F.: Reconstructing Open Source Software Ecosystems: Finding Structure in Digital Traces. ICIS 2016 Proceedings (2016)
27. Biazzini, M., Baudry, B.: May the Fork Be with You: Novel Metrics to Analyze Collaboration on GitHub. WETSoM 2014 Proceedings, 37–43 (2014)
28. Borges, H., Valente, M.T., Hora, A., Coelho, J.: On the Popularity of GitHub Applications: A Preliminary Note (2015)
29. Lindberg, A., Berente, N., Gaskin, J., Lyytinen, K.: Coordinating Interdependencies in Online Communities: A Study of an Open Source Software Project. Information Systems Research 27, 751–772 (2016)
30. Corbin, J.M., Strauss, A.L.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications, Los Angeles, Calif (2008)
31. Chapin, N., Hale, J.E., Khan, K.M., Ramil, J.F., Tan, W.-G.: Types of Software Evolution and Software Maintenance. Journal of Software Maintenance and Evolution: Research and Practice 13, 3–30 (2001)
32. Ma, W., Chen, L., Zhang, X., Zhou, Y., Xu, B.: How Do Developers Fix Cross-project Correlated Bugs? A Case Study on the GitHub Scientific Python Ecosystem. ICSE 2017 Proceedings, 381–392 (2017)
33. Garud, R., Jain, S., Tuertscher, P.: Incomplete by Design and Designing for Incompleteness. Organization Studies 29, 351–371 (2008)

The Influence of Emoticons on the Perception of Job-Related Emails: An Analysis Based on the Four-Ear Model

Claus-Peter H. Ernst[1], Martin Huschens[2], Sebastian Herrmann[3], and Laura Hoppe[2]

[1] Frankfurt University of Applied Sciences, Frankfurt am Main, Germany
cernst@fb3.fra-uas.de
[2] Johannes Gutenberg-Universität Mainz, Mainz, Germany
{huschens,lhoppe}@uni-mainz.de
[3] Hochschule RheinMain, Wiesbaden, Germany
sebastian.herrmann@hs-rm.de

Abstract. Non-verbal communication cues and their surrogates in computer-mediated communication, emoticons, can influence how a message is understood. Based on the four-ear model, we examine in detail at which levels of communication emoticons affect message perception. Using a factorial survey with a treatment control group design (N = 104), our findings suggest that emoticons do not influence the understanding of a message at the content level. However, we show that they significantly influence the metamessage. These findings hold important theoretical and practical implications: First, our results help to explain the mixed results of previous studies on the effects of emoticon usage on message interpretation. Second, we show that emoticons are a means to convey contextual information over and above the raw content of the message, especially at the self-revelation and relationship level. However, this does not come without a price: Emoticons dilute the effects of job-related emails at the appeal level.

Keywords: Emoticons, Computer-Mediated Communication, Four-Ear Model, Factorial Survey

1 Introduction

How a recipient perceives and understands a spoken message depends, on the one hand, on the verbal content, and, on the other hand, on the contextual interpretation of non-verbal elements such as facial expressions and body language [1]. In the workplace context, computer-mediated communication (CMC), and especially email, has positioned itself as an alternative to face-to-face communication. However, this communication style has one important disadvantage: It lacks non-verbal communication cues. Since emails are a text-based form of CMC, a sender has no direct opportunity to use non-verbal communication elements to provide contextual information over and above the written information and, hence, the context of a

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

message might be difficult to interpret. Emoticons, however, as text-based symbolizations of facial expressions, emotional states, and feelings [2], can serve as corresponding cues.

Many studies show that emoticons have a certain impact on message interpretation. They are used and perceived as surrogates for non-verbal cues and are, for example, able to soften the illocutionary force of speech acts in emails [2-6]. In contrast, other studies claim that emoticons can only complement messages, and that they are far less important for message interpretation than the written content [7]. The aim of our research paper is to further clarify the effect of emoticons on message perception in job-related emails by drawing on the four-ear model of Schulz von Thun [8]. Indeed, we aim to explain why and how emoticons affect the illocutionary force of written (speech) acts and are able to explain the mixed results of previous studies.

The four-ear model postulates that every type of communication has an underlying anatomy that is a combination of four communication levels at which a message can be sent and received: the factual information level, the self-revelation level, the relationship level, and the appeal level. We hypothesize that in CMC, emoticons such as the smiley-face [:-)], influence the recipient's perception of a job-related message at all these communication-defining levels.

To test our hypotheses, we conducted a factorial survey [9] with 104 respondents and a treatment control group design in which one group was shown an email with an emoticon and another group was shown an email without an emoticon. Our main finding is that emoticons only influence the receiver's perception of the metamessage, namely at the self-revelation level, relationship level, and appeal level, but that they do not have an impact on the receiver's perception at the factual information level. First, this shows that emoticons are indeed perceived as surrogates for non-verbal communication cues and are able to convey information over and above the pure content level, however they are not strong enough to shape or even contradict the content level of a message. Second, we showed that emoticons are able to soften the illocutionary force of speech acts because they convey further information at the relationship, self-revelation and appeal levels. However, this does come with a certain price since the effect of the message at the appeal level is mitigated.

The paper is structured as follows: In the following section, we will introduce the four-ear model of Schulz von Thun [8] and we will also provide the theoretical foundations of CMC and of emoticons, both in general and in the professional context. Following this, we will present our research model and will introduce our research methodology, more specifically, the design of the factorial survey. Finally, we will present and discuss our results before concluding our article with the limitations of our study and the practical implications of our results.

2 Theoretical Background

2.1 The Four-Ear Model of Schulz von Thun

Schulz von Thun [8] postulated that each message has an underlying anatomy that is a combination of four different communication levels at which a message can be sent and received, respectively: the factual information level, the self-revelation level, the relationship level, and the appeal level. This model is also commonly termed the "four-ear" model, which refers to the ways in which the recipient understands (or hears) the message. The general process of communication and the four layers of a message are depicted in figure 1.

At the factual information level, i.e., the content aspect of the message, pure factual information is passed from the sender to the recipient. This communication layer is conveyed by the pure spoken word or the written textual word (logical digital language). At the factual information level, the recipient assesses whether a message is true or false, relevant or irrelevant, and reliable or unreliable.

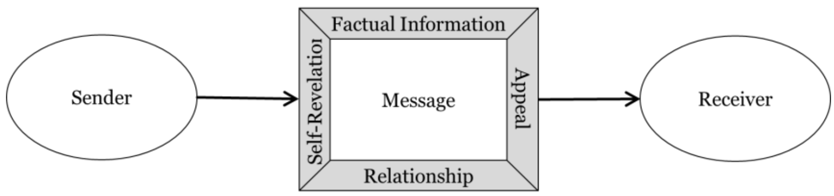


Figure 1. Communication Process and Four-Ear Model [8, p. 30]

Besides the pure words used, Schulz von Thun [8] argues that a message inherently consists of an additional subtext or metamessage. This metamessage is only partly influenced by the pure textual information delivered by the sender. Rather, it is conveyed via non-verbal communication means such as facial expressions, gestures, tone, speech speed and general body language, etc.

First, at the self-revelation level, the sender discloses information about themselves and their current motives, values, and emotions (so called I-messages). This level is described as a *small sample of personality*, since information about the communicator is inevitably revealed. Second, at the relationship level, the sender indirectly expresses a position towards the recipient (so called we-messages). Lastly, the appeal level provides information about the response expected by the sender.

As an example, imagine a conversation between two colleagues where one tells the other "your report is not here". At the factual information level, the recipient may interpret the message as raw information regarding the current state of the process. At the self-revelation level, the recipient may get the impression that the sender was irritated by the delay. At the relationship level, the recipient may understand the message as an accusation of incompetence. At the appeal level, the recipient may feel that the sender expects him/her to work more thoroughly in the future.

2.2 Emoticons as Text-Based Computer-Mediated Communication Cues

Computer-mediated communication (CMC) is now an established form of communication, and continues to steadily develop itself as such. CMC can be defined as “[a]ny communicative transaction that takes place by way of a computer …” [10, p. 552]. In contrast to real-life face-to-face communication, much of CMC today is founded on text such as emails and instant messages. In these contexts, analogue language normally used to clarify messages is inaccessible [1]. As a result, a sender’s ability to show emotions, for example, is limited.

However, as a substitute for these missing elements, text-based elements have established themselves as non-verbal cues in written communication. These CMC cues can be equally effective as regular analogue language [11] and are thus able to help clarify messages [12] as well as provide information about the type and strength of the emotions that the sender wishes to convey with the message [13]. One popular form of CMC cues are emoticons, i.e., text-based symbolizations of facial expressions, emotional states, and feelings [2].[1]

Emoticons are not only used in the private domain but also increasingly in job-related communications within companies and organizations [16]. Indeed, it has been shown that positive emoticons in the professional context provide three functions: (1) marking positive attitudes, (2) marking jokes/irony, (3) acting as hedges, i.e., strengthening expressive speech acts (such as thanks or greetings) or softening directives and criticism (i.e., requests, rejections, corrections, and complaints) [2].

Multiple studies have examined the effects of emoticons in CMC. Huang et al. [17] found that the use of emoticons in instant messaging has a positive effect on the enjoyment, personal interaction, perceived information richness, and perceived usefulness of messaging applications. Some studies have proven that emoticons serve the function of clarifying textual messages by accentuating a tone or meaning [3, 6], thus, helping to communicate more clearly. Derks et al. [4] find that emoticons can help to create ambiguity and express sarcasm online. However, they conclude that emoticons are not strong enough to outweigh the verbal content of a message or even change the valence of a message. Also, Walther and D'Addario [7] find that emoticons can only complement messages and do not have the strength to enhance messages.

Our study seeks to specify these results. By drawing on the four-ear model, we were able to identify the level of communication at which emoticons unfold their effects and to what extent they are doing so.

---

[1] In addition to text-based CMC cues, there are also pictographic-based cues such as emojis (e.g., ?, ?). However, they are currently primarily used in private communication such as in instant messaging services and in social media [14], and especially on touch-based mobile devices [15]. Since we are interested in job-related communications, we thus refrained from using emojis and focused on their text-based predecessors, i.e., emoticons [14].

3 Research Model

While in face-to-face communication facial expressions might be used by the sender to provide information about their point of view as to how their message should be interpreted by the recipient [1], in emails, emoticons can be used to provide this information, since they are a form of text-based analogue language. We thus assume that subjects will understand a message significantly differently if emoticons are used. Our dependent variable, which is the receiver's perception of a message, is divided into four subconstructs, each representing one of the four communication levels of the four ear-model [8]. We expect to find a direct effect of our independent variable (assignment to the treatment group or the control group) on all these communication-defining levels, namely the factual information level, the self-revelation level, the relationship level and the appeal level. Figure 2 presents our research model.

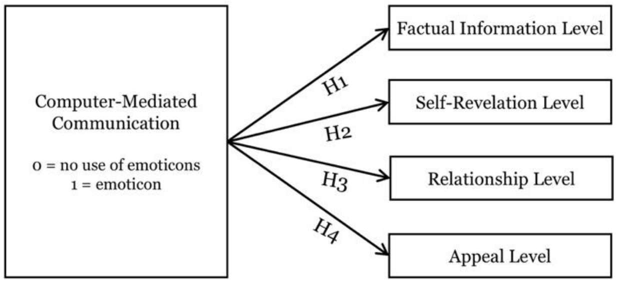


Figure 2. Research Model

First, we believe that emoticon usage influences the understanding of job-related emails at the factual information level. Indeed, non-verbal cues, such as facial expressions, gestures, and body language, influence the understanding of a message [8] by amplifying, weakening or even changing the entire verbal content — especially if there is a discrepancy between the verbal content and non-verbal cues [1]. In line with this, emoticons, which act as surrogates of facial expressions in CMC, can be expected to act similarly in the written word. However, following the insights of Derks et al. [4] and Walther and D'Addario [7], we assume the effect to be only weak, as the meaning and relevance of a message seems to be primarily conveyed by the spoken or written words. We hypothesize that:

*The usage of emoticons influences the recipient's perception of a job-related message at the factual information level (**H1**).*

Similarly, we argue that emoticons also significantly shape the metamessage of a message. More specifically, we believe that emoticons can convey information over and above the pure content level, providing enhanced information in terms of how a message should be understood and, by that, ultimately shape a message's perception.

Indeed, as Schulz von Thun [8] stated, messages' metamessages are generally conveyed to a large proportion by non-verbal cues. In line with this, we assume that emoticons as facial expressions surrogates influence all levels of communication that carry a message's metamessage — namely the self-revelation, relationship and appeal levels. We hypothesize that:

*The usage of emoticons influences the recipient's perception of a job-related message at the self-revelation level (**H2**), the relationship level (**H3**), and the appeal level (**H4**).*

4 Research Design

4.1 Factorial Survey

To test our hypotheses, we conducted a factorial survey (also called vignette study). Factorial surveys use "short, carefully constructed description[s] of a person, object, or situation, representing a systematic combination of [the research-relevant] characteristics" [18, p. 128]. Respondents are then presented with these different fictional descriptions, and assess them on the basis of a questionnaire. Such descriptions may consist of a textual description, a video, illustrations or any other form of explanation.

In our context, we asked respondents to put themselves in the position of a company intern that receives an email from their supervisor, in which criticism is expressed (see table 1). We used the vignette character of an intern because we expected to recruit a quite young sample of people — indeed, we posted the invitation to our survey on Facebook, in special interest groups that are mainly used by students. We assumed that these subjects could quite easily put themselves in the position of an intern or might even have experienced such a situation themselves. Moreover, we used the scenario of receiving a criticism-filled email, since criticism in the workplace is typically expressed with the goal of improving work performance [19] and senders commonly use "positive" analogue language such as facial expressions to soften criticism and directives and to reduce the recipient's negative feelings [2].

Since the defining factors of our research hypotheses are emoticons, we chose the presence and absence of a positive smiley-face emoticon as the factor levels of our vignettes (in the control group, the emoticon shown in table 1 has been omitted). Finally, we chose to implement a between-subjects design in which subjects were randomly assigned to the treatment and control group.

In order to evaluate the influence of an emoticon on each of Schulz von Thun's communication levels, we developed two context-specific items for each layer in the subsequent questionnaire. All items were measured using a seven-point Likert-type scale ranging from "strongly disagree" to "strongly agree". Table 2 presents the items of our questionnaire.

Table 1. Introductory Text and Treatment Vignette

Introductory text	Place yourself in the following situation: You are currently an intern with the company Krueger GmbH. Your tasks include, among other things, the preparation of presentations, which requires your participation in certain meetings. In the following section, you will receive an email from your supervisor, in which your work method is discussed. Please share your impressions by evaluating the given statements on a scale from 1 ("strongly disagree") to 7 ("strongly agree").
Vignette treatment group Critical email with emoticon	Dear intern, It has come to my attention that you often arrive late for our meetings, and that the presentations you have edited are not submitted in a timely manner. Time management is crucial for good work performance. I would kindly ask you to ensure to be on time in future and to adhere to your deadlines :-) Best regards, Kim Krueger

Table 2. Questionnaire Items

Communication level	*Item*
Factual information (FI)	I interpret the message as being an advice-giving message on how to improve my job performance (FI1) The message is informative (FI2)
Self-revelation (SR)	Mr. Krueger is assertive (SR1) Mr. Krueger is upset (SR2)
Relationship (R)	Mr. Krueger is disappointed by me (R1) Mr. Krueger does not appreciate me (R2)
Appeal (A)	I question my own way of working, and will work more conscientiously in the future (A1) I will consider the topics raised (A2)

4.2 Data Collection

As described above, we recruited German-speaking users via Facebook. The participation was incentivized by a raffle of 20 € Amazon vouchers for three of the participants. In this manner, we obtained 104 complete online questionnaires with 55 respondents evaluating the vignette without the emoticon and 49 respondents evaluating the vignette with the emoticon. As expected, our sample was quite young with 79% of the respondents aged 18 to 25. Our sample also consisted of a higher proportion of women (58%) and a majority of subjects (65%) had a high-school certificate as their highest current educational achievement.

5 Results

5.1 Descriptives

Table 3 presents the descriptives per questionnaire item and the average composite score for every communication level. For all items, the scores in the treatment group were lower than those in the control group.

We also examined the distribution properties of our sample. The variances of our two groups were equal, and the sample size of both groups were large (>30) and approximately the same size. However, a first examination of the QQ-plots and a subsequent Shapiro-Wilk-test ($W_{FI} = .962$, $p < .05$; $W_{SR} = .969$, $p < .05$; $W_R = .974$, $p < .05$; $W_A = .873$, $p < .05$) showed that our variables were not normally distributed. Still, since the t-test is relatively insensitive to a violation of the normality assumption, we were able to use it for our dataset analyses [20].

Table 3. Descriptives

Item	*Control group (CG)*			*Treatment group (TG)*		
	Mean	SD	SE	Mean	SD	SE
Factual information (FI)*	5.882	.990	.133	5.561	1.139	.163
FI1	5.691	1.318	.178	5.449	1.430	.204
FI2	6.073	1.184	.159	5.673	1.144	.163
Self-revelation (SR)*	5.091	1.097	.148	4.184	1.093	.156
SR1	5.636	1.310	.177	4.918	1.351	.194
SR2	4.545	1.317	.178	3.449	1.355	.193
Relationship (R)*	4.164	.938	.126	3.398	1.020	.146
R1	4.782	1.066	.144	3.694	1.278	.183
R2	3.545	1.288	.174	3.102	1.159	.166
Appeal (A)*	6.273	1.022	.138	5.806	1.025	.146
A1	6.182	1.188	.160	5.714	1.061	.151
A2	6.364	.988	.133	5.898	1.046	.149

* Composite score per communication level, normalized with item count (= 2)

5.2 Hypothesis Testing

To test our hypotheses, we performed multiple independent-means t-tests. Table 4 presents the t-values, degrees of freedom (df), the 2-tailed significance levels (sig.), and the effect sizes for all item means.

Hypothesis 1 assumed that we would find an effect of emoticon usage at the factual information level. However, we found no significant difference between the treatment and control group ($mean_{TG} = 5.561$, $mean_{CG} = 5.882$, $t_{(102)} = 1.5226^{n.s.}$). Both in the treatment and control group subjects evaluated the message similarly strong as being informative and as being an advice-giving message to do better in the future. Hence, hypothesis 1 was not confirmed. We explain this finding by the fact that factual information is predominantly conveyed by the digital language (the raw text) and not by analogue language (e.g., facial expressions, emoticons) [1]. Indeed, recipients mostly use the content aspect of the message to evaluate the relevance, importance and trustworthiness of the information provided [8]. This is in line with Derks et al. [4] who also described that "emoticons do not have the strength to turn around the valence of a message" [4, p. 386].

Table 4. Significance and Effect Sizes

Communication level	*t*	*df*	*Sig.*	*|Hedges' g|*
Factual information	1.5226	102	.131	.300
Self-revelation	4.2171	102	.000	.828
Relationship	3.9668	102	.000	.784
Appeal	2.3207	102	.022	.456

Concerning the self-revelation level, we found that participants evaluated the items measuring the self-revelation level significantly lower when the message contained a positive emoticon ($mean_{TG}$ = 4.184, $mean_{CG}$ = 5.091, $t_{(102)} = 4.2171^{***}$). We also evaluated the effect size. Concerning the effect at the self-revelation level, the effect size was large ($|g| = .828$). As a result, hypothesis 2 was confirmed.

Likewise, the negative associations of receiving criticism were lower at the message's relationship level when the message contained a positive emoticon than when the message did not contain a positive emoticon ($mean_{TG}$ = 3.398, $mean_{CG}$ = 4.164). The difference was significant ($t_{(102)} = 3.9668$, $p < .001$) and represented a medium-sized effect ($|g| = .784$). Thus, hypothesis 3 was confirmed.

Finally, we also found a significant group difference ($t_{(102)} = 2.3207$, $p < .05$) for the appeal level of the email, which represented a small-sized effect ($|g| = .456$). More specifically, subjects in the treatment group evaluated the appeal level lower than the subjects in the control group did ($mean_{TG}$ = 5.806, $mean_{CG}$ = 6.273). Hence, hypothesis 4 was confirmed. These results confirm on the one hand that the emoticon has a significant influence on the appeal level, and on the other hand, that the positive smiley emoticon softens a message's illocutionary force.

6 Discussion

Whereas Walther and D'Addario [7] concluded that emoticons do not have the strength to enhance messages, Derks et al. [4] showed that emoticons may act as surrogates for non-verbal communication cues in CMC, however, they "do not have the strength to turn around the valence of a message" [4, p. 386]. Our study helps to explain those mixed results concerning the effects of emoticons.

More specifically, previous studies did not differentiate between messages' multiple communication levels and rather understood messages as one-dimensional entities. In contrast, we differentiated between four communication levels and showed that emoticons have a significant influence on the metamessage of messages, namely on messages' self-revelation levels, relationship levels and appeal levels. Still, at the level of factual information, the influence of emoticons was insignificant. Based on these findings, we can conclude that the use of emoticons helps recipients interpret CMC messages by increasing information richness while neither diluting nor enhancing the factual information of the message. Our findings also explain why emoticons are able to soften the illocutionary force of speech acts: They are able to convey information within the metamessage, namely at the self-revelation level, relationship level, and appeal level.

In addition to these insights into the general effects of the emoticon on communication levels, we also provided insights into the direction of these effects. More specifically, in the specific case of an email expressing criticism, we showed that a positive emoticon can help soften criticism while simultaneously retaining the factual information of the message. In general, criticism can cause negative associations at the relationship and self-revelation levels in the recipient. Indeed, people can take criticism personally [21], even though in the workplace it is typically expressed with the goal of improving work performance and not meant personally [19]. Since a positive emoticon such as the smiley face conveys a sender's positive attitude [22], its usage in an email expressing criticism reduces negative associations of the recipient at the relationship and self-disclosure levels [8].

However, this does not come without a price. We also observed lower scores at the appeal level: Positive emoticons also seem to reduce the recipient's understanding of the overall intention of the sender and their expected response, i.e., the improvement of their work performance. Thus, expressing criticism along with the usage of emoticons may not be effective at evoking the awaited response. This effect can be explained by the fact that emoticon usage, especially in the professional domain, also comes along with decreased perceptions of competence and authority [23]. However, whereas the effects of the positive emoticon at the relationship and self-disclosure levels were large and medium-sized, they were only small-sized at the appeal level, which implies that the price paid is probably not all too high.

7 Conclusions

In this article, we evaluated the influence of emoticons on recipients' interpretation of Schulz von Thun's four communication levels, in the context of workplace emails expressing criticism. Based on a factorial survey with, our results suggested that emoticons exert an effect on the communication levels linked to the metamessage of a message, that is, the relationship level, the self-revelation level, and the appeal level. However, we did not detect a significant influence of emoticons on the interpretation of the content aspect of the message, that is the factual information level.

These findings hold important practical implications. Our study emphasizes that emoticons can be a useful means for conveying information about the metamessage of a message, thus helping the recipient of a message correctly interpret it as it was intended, while not diluting its factual information. Furthermore, we showed that positive emoticons can be used in workplace emails that express criticism, in order to mitigate recipients' negative associations with the messages at the relationship and self-revelation levels. In other words, a recipient takes criticism less personally if the sender uses a positive emoticon. At the same time, although positive emoticons do not have an influence on the perception of the message at the factual information level, they do soften the illocutionary force of the message at the appeal level. Hence, the mitigation of recipients' negative associations comes with a price.

There are some limitations to our study that should be noted. Certainly, a situation that is described in a vignette can never be fully realistic and is especially prone to individual misperceptions. Thus, the external validity of our study might be limited, even though we tried to select an appropriate vignette situation that fit our expectedly young sample of people. Additionally, we only used two specific emails (which were identical with the exception of the additional emoticon in one of them). Although the scenario and the wording were carefully constructed and examined by third parties, it is still possible that the respondents might have misinterpreted the emails. Indeed, the situation that was described in vignette was not equally realistic for all respondents. For example, in the case of the employed people in our sample, it was maybe more difficult for them to imagine themselves as interns than it would be for students.

Furthermore, there are certain limitations to between-subjects designs when it comes to perceptions, opinions and situational judgments as is the case in factorial surveys. It can be argued that in between-subjects designs, each respondent judges only a single vignette, which can lead to measurement problems due to individually different vignette contexts. However, results from a within-subject design would be seriously flawed as subjects would not have been blind to condition and, thus, memory effects, sponsorship effects, and sequence effects would have come up.

Moreover, we only included one specific positive emoticon, :-). Hence, different results might be found when other positive emoticons are used, such as :) or :-D. In addition, we did not look at emoticons that introduce irony into the communication, such as ;-).

Finally, we only surveyed German-speaking people and 79 percent of our respondents were between 18 and 25 years old. Hence, differences might be found for non-German speaking people as well as for other age groups. Overall, due to these shortcomings, the external validity of our study might be limited.

As a next step, we intend to take a look at the influence of multiple kinds of emoticons (positive, negative, ironic, etc.) on recipients' interpretation of different kinds of job-related emails (positive, critical, etc.).

References

1. Watzlawick, P., Bavelas, J.B., Jackson, D.D.: Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes. W. W. Norton & Company, New York, NY (2011)
2. Skovholt, K., Grønning, A., Kankaanranta, A.: The Communicative Functions of Emoticons in Workplace E-Mails: :-). Journal of Computer-Mediated Communication 19, 780-797 (2014)
3. Krohn, F.B.: A Generational Approach to Using Emoticons as Nonverbal Communication. Journal of Technical Writing and Communication 34, 321-328 (2004)

4. Derks, D., Bos, A.E.R., von Grumbkow, J.: Emoticons and Online Message Interpretation. Social Science Computer Review 26, 379-388 (2008)
5. Lo, S.K.: The Nonverbal Communication Functions of Emoticons in Computer-Mediated Communication. CyberPsychology & Behavior 11, 595-597 (2008)
6. Dresner, E., Herring, S.C.: Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force. Communication Theory 20, 249-268 (2010)
7. Walther, J.B., D'Addario, K.P.: The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. Social Science Computer Review 19, 324-347 (2001)
8. Schulz von Thun, F.: Miteinander Reden 1 - Störungen und Klärungen: Allgemeine Psychologie der Kommunikation. rororo, Reinbek, DE (2001)
9. Rossi, P.H., Anderson, A.B.: The Factorial Survey Approach: An Introduction. In: Rossi, P.H., Nock, S.L. (eds.) Measuring Social Judgments: The Factorial Survey Approach, pp. 15-67. Sage, Beverly Hills, CA (1982)
10. McQuail, D.: Mcquail’s Mass Communication Theory. Sage, London, UK (2010)
11. Walther, J.B.: Theories of Computermediated Communication and Interpersonal Relations. In: Knapp, M.L., Daly, J.A. (eds.) The Handbook of Interpersonal Communication, pp. 443-479. Sage, Thousand Oaks, CA (2011)
12. Vandergriff, I.: Emotive Communication Online: A Contextual Analysis of Computer-Mediated Communication (CMC) Cues. Journal of Pragmatics 51, 1-12 (2013)
13. Harris, R.B., Paradice, D.: An Investigation of the Computer-Mediated Communication of Emotions. Journal of Applied Sciences Research 3, 2081-2090 (2007)
14. Marengo, D., Giannotta, F., Settanni, M.: Assessing Personality Using Emoji: An Exploratory Study. Personality and Individual Differences 112, 74-78 (2017)
15. Pavalanathan, U., Eisenstein, J.: Emoticons Vs. Emojis on Twitter: A Causal Inference Approach. In: Proceedings of AAAI Spring Symposium on Observational Studies through Social Media and Other Human-Generated Content (OSSM) (2016)
16. O’Kane, P., Hargie, O.: Intentional and Unintentional Consequences of Substituting Face-to-Face Interaction with E-Mail: An Employee-Based Perspective. Interacting with Computers 19, 20-31 (2007)
17. Huang, A.H., Yen, D.C., Zhang, X.: Exploring the Potential Effects of Emoticons. Information & Management 45, 466-473 (2008)
18. Atzmüller, C., Steiner, P.M.: Experimental Vignette Studies in Survey Research. Methodology 6, 128-138 (2010)
19. Ang, S., Cummings, L.L., Straub, D.W., Earley, P.C.: The Effects of Information Technology and the Perceived Mood of the Feedback Giver on Feedback Seeking. Information Systems Research 4, 240-261 (1993)
20. Pagano, R.R.: Understanding Statistics in the Behavioral Sciences. Wadsworth, Belmont, CA (2012)
21. Kim, E.J., Yamaguchi, A., Kim, M.-S., Miyahara, A.: Effects of Taking Conflict Personally on Conflict Management Styles across Cultures. Personality and Individual Differences 72, 143-149 (2015)
22. Mitchell, A.A.: The Effect of Verbal and Visual Components of Advertisements on Brand Attitudes and Attitude toward the Advertisement. Journal of Consumer Research 13, 12-24 (1986)
23. Glikson, E., Cheshin, A., van Kleef, G.A.: The Dark Side of a Smiley: Effects of Smiling Emoticons on Virtual First Impressions. Social Psychological and Personality Science (2017)

Global Adoption of Unified Communication Technologies as Part of Digital Transformation in Organizations: A Cross-Cultural Perspective

Christian Meske[1], Tobias Kissmer[2], Stefan Stieglitz[2]

[1] Freie Universität Berlin, Department of Information Systems, Berlin, Germany
{christian.meske}@fu-berlin.de
[2] University of Duisburg-Essen, Department of Computer Science and Applied Cognitive Science, Duisburg, Germany
{tobias.kissmer,stefan.stieglitz}@uni-due.de

Abstract. Unified Communication (UC) is increasingly introduced in global enterprises to support the digital interaction of employees. Cultural differences can influence such adoption of corresponding technologies and need to be considered by IT managers. In this study, we show how cultures with specific characteristics (according to Hofstede and Hall) adopt UC differently. This research paper is based on a unique dataset, consisting of communication data from over 6,000 users of a multi-national enterprise. During the period of six months, 4.6 million records were collected with the purpose of analyzing the adoption after the rollout of a UC technology. Eight subsidiaries in Asia, North America and Europe were analyzed. The results show that especially Hofstede's factors Individualism/Collectivism and Long-Term Orientation as well as Hall's factor Monochronic/Polychronic culture helped to explain post-implementation usage activity. The results help IT managers to understand potential cultural influences on introductions of according technologies in global enterprises.

Keywords: Communication Technology, Unified Communication, Digitization, Diffusion, Adoption

1 Introduction

The digital transformation is prospected to have a significant impact on both private and professional life [1]. Taking a closer look at how enterprises deal with it, one can see companies in a permanent struggle, striving to be more digitally mature [2]. On the one hand, this is naturally driven by economic reasons, maintaining the ability to cope with other competitors in terms of new digital products or providing other beneficial processes. On the other hand, and just as important, companies have to transform their working environment, since employees are more attracted by digital leaders [3]. Information Systems need to adjust to different architectures to serve the people's increasing need for efficiency and mobility [4]. As [2] state, the digital transformation is not purely about the technology that is being implemented. It's about how companies integrate it into their global organizations, which is why the adoption process for digital

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

oriented technology is crucial for this transformation. The digitalization of a company does also require facilitating communication and "will only succeed with the aid of modern, professional corporate communication" [5]. To do so, enterprises need to adjust their processes and adopt digital novelties. Physical presence is more and more replaced by new communication solutions [3]. It will be of great relevance which channel of communication is going to be used in future to foster knowledge within the organization, whereas solely the choice of channel delivers a certain context of the information itself [4]. It is assumed, that the importance of communication and collaboration across geographical spans increases, emphasizing the need to implement intelligent digital communication concepts such as Unified Communication.

Long before the actual dissemination and implementation of Unified Communication (UC) solutions in companies, various investigations have predicted a high impact of this technology on business. As the everyday work becomes easier to manage, more efficient and the increasing need for communication is easier to handle for office workers [6]. UC is understood as „ […] communication services across geographical boundaries and networks based on rules and policies that provide seamless integration between services" [7]. Companies are increasingly confronted with the change due to globalization and digitization. An international collaboration between employees of different countries becomes more and more important and can be affected by many factors. Two examples are the available type of communication technology and the prevailing culture of that country [8]. For IT-managers, it is therefore valuable to understand the mutual influence of certain *cultural* aspects and the diffusion of communication technologies in this context. It will furthermore help decision makers to cluster and focus respectively, since implementation of new technologies will come in a higher frequency. Employees must cope with this and adopt. This is directly affecting their efficiency, which is why enterprise decision makers should prepare the implementation of new technologies in an optimal way and estimate the adoption speed. For this, it is of importance to know the mutual influence that culture comes along with in different countries. Though we have extensively reviewed the literature in this field, we did not come across a study that is analyzing the correlation between an IS adoption and culture based on a large dataset. Therefore, in this paper we focus on the following two research questions (RQ):

RQ1: How does the usage activity develop over time in different countries after the system's implementation?

RQ2: To which extent can commonly applied cultural factors serve as an explanation for the development of system usage in different countries?

To answer this, our paper is structured as follows. In section 2, we provide an overview of related work on the digital transformation and unified communication. This is followed by an overview of existing research on cultural influences on IS adoption. In section 3, we describe the research design, methodology and underlying data basis in more detail. The results are presented in section 4, followed by a discussion with the study's limitations in section 5. This paper ends with a summary and an outlook to further research in section 6.

2 Related Work

2.1 Digitalization and Unified Communication

"Employees want to work for digital leaders." [2]. Enterprises strive to digitally adapt to not only support digital goods and services that secure their economic future, but also provide the environment to ensure the continuance of their human capital. A more digital life comes with more convenience and saves time, and most of all it is more efficient [3]. As [9] postulate, this transformation brings the need of higher connectivity between individuals globally, where people are using a manifold variety of devices. Hence, they strive to be mobile while increasing their own communication activity [9], [10]. More and more people are used to rapid technological changes in society (see e.g. e-business, social media etc.). This is increasingly considered by businesses and as [11] predicts, the frequency of changes in enterprise IS will also increase with the digital transformation process [12], [13]. Thus, many companies have undertaken efforts to start digital initiatives to change their IS portfolio [9]. The software "Skype for Business", which is a UC tool, is an exemplary technology for a modern, cloud enabled architecture. [14] and [15] consider communication technologies as digital technologies, which are the core of digital transformation processes. They argue that the IT strategy needs adjustments to more future-oriented architectures, based on key external digital trends, such as "pervasive connectivity" or "growth of cloud computing". [16] are indicating that this has to be taken a step further and include a digital transformation of cross-company-cooperations. [6] predicted an enormous market potential for the still developing technology of UC for private as well as enterprise usage, crediting it the possibility to change the market. Later, [17] continued this thought and even brought it even further. They concluded that integrated technologies such as UC are one of the most influencing factors on daily work processes. UC represents a combination of different communication services, leading to an improved user experience, since multimedia and communication requirements from users can be covered [17]. Since this technology is becoming of more and more relevance, a vast variety of tools exist, though the enterprise market is dominated by only a few vendors such as Microsoft (Skype for Business), Cisco (Unified Communication Manager) and Mitel [18]. UC focusses on real-time and non-real-time communication, whereas it should be distinguished from enterprise social networks which foster the collective wisdom in organizations.

2.2 Cross-Cultural Adoption of Technology

Regarding the adoption of Information Systems (IS) in different cultural environments, several studies with each a certain focus exist. They all focus on the general acceptance of IS. Especially [8] have reviewed the literature and concluded that several studies base themselves on or mention the cultural dimensions developed by Hofstede. A total of 16 studies existed at the time of their research, focusing on the interaction between culture, technology and adoption. Most of these studies concentrate on the theoretical combination of these aspects, further exploring the model of

Hofstede. However, the literature review brought up just a few studies that support their hypothesis with a dataset of varying size. An example for this is [21]. [22] builds his studies on the cultural dimensions of Hofstede as well and rudimentarily touches the combination of both adoption and technology. Furthermore, an investigation of [23] revealed that also the management of organizational culture is important for an adoption of technologies, because it influences on the one hand the acceptance by individuals and on the other hand the actual use of information technology within an organization. Another affecting factor of the technology diffusion might be the perceived risk and social influence [24]. Particularly social influence can have impact on the decision regarding adoption or rejection of technology, which was indicated by results of [24]. They examine moderating and predictive factors and their impact on trust regarding the adoption of corporate cloud technology. Trust as a basis for cooperation within organizations affects the IT adoption as well, which is suggested by research outcomes of [25]. The combination of all these aspects play an important role in the adoption of a system. However, these influences on adoption are not equally distributed globally but rather varying for each cultural environment. For example, performance oriented cultures tend to compare a new system with the known and try to derive a direct advantage out of this regarding efficiency [26]. In this context, the technology chosen as foundation for this study (UC) may play an important role here, since it does not only support pure content but also provide the possibility to adjust media richness (e.g. chat, audio, video). The argument that the culture could affect the acceptance of technology was confirmed by [27] as well. Therefore [27] conclude(s/d) that such factors impact the technology adoption in dependence of the country and because of their culture. South Korea for example is classified as a culture with a constant fear of not being within the given specification. Another dimension is collectivism and individualism. The United Kingdom can be understood as a strongly individualistic society, where a high complexity is prevailing and the single individuals are in a loose connection with each other. South Korea and China on the contrary are known to be collectivistic and show strong ties among individuals [28], [29]. As previously mentioned, we did not come across research that combined the cultural insights of previous studies (Hofstede, Hall) with the adoption of technology in a digital context. Based on the review of existing literature that we undertook, we contribute with the results of our study towards the closure of this existing gap and provide insights to a large dataset, linking the adoption with culture.

3 Research Design

3.1 Research Model

Based on previous research and the derived knowledge, we chose to underlay Halls' and Hofstede's factors for this study. We gather the input and factors solely based on the studies of [29], [33] and [34] as reference for Hall's research. Out of currently six existing factors based on Hofstede's research, we have chosen five due to the following reasons. Based on his initial research [29], collectivism, uncertainty avoidance, power distance and masculinity are the oldest factors which date back to 1984. According to the findings of [8], these factors play an important role in the context of user adoption. Accordingly, a wide spread collectivism will either result in the quicker acceptance of the technology or a collectivistic restraint from usage [8]. Furthermore, [35] assume(s/d) that cultures that react cautiously to unknown, classify new technologies as unsafe and adapt relatively slow. The power distance is an often-investigated factor in science and may influence a more agile exchange between hierarchical levels, resulting in a different speed of adaptation [8]. Masculinity refers to the orientation of specific masculine or feminine values within a culture. Hence countries with a high value tend to be oriented to masculine values, such as higher focus on competition. The opposite orientation is when cultures foster social components [8], [20], [22], [34], [36]. Long-Term orientation is included in more recent research of Hofstede and catches a culture's handling of past and future with the respective influence on work [37]. Hall mainly describes two factors. The first, low and high context cultures, is capturing a culture's tendency for the need of context to make decisions. High context cultures would gather more information and context before deciding on something than low context cultures. The second factor, monochronic and polychronic cultures, describes another variation of time, more presence-based than Hofstede's long-term orientation. Monochronic cultures are better organized and usually on time, whereas polychronic cultures tend to be organized differently and e.g. tend to start multiple tasks at the same time [34], [38]. For Hall's factors, we were in the need of transcoding them into a suitable format for further analysis. Hence, we made use of the existing research of [34] who already operationalized the classifications of Hall. For this, figure 1 provides an overview of the chosen cultural factors.

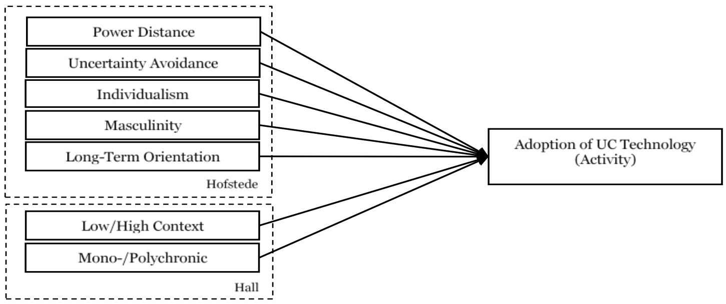


Figure 1. Research Model based on cultural factors

3.2 Methodology and Data Basis

Our foundation is a dataset from a large enterprise headquartered in Germany with approx. 100.000 employees. Being in the manufacturing business, the company decided to switch globally from various communication channels to a unified solution: Skype for Business. Throughout the basic data set, all communication channel activities were considered to conduct comparing analysis. However, since the chat functionality instantly came with the prerequisite of the software installation, we could not make a clear distinction of when the chat feature was used, hence we focused only on the voice component. The voice feature though could be assigned to a specific implementation date and hence we could mitigate possible interferences. Accordingly, the local management officially pronounced the order to use the solution as soon as possible. In all countries, the internal marketing and advertisement for the new technology has been conducted in the same way - before and after the introduction. We based our analysis on 6,309 users in 8 countries and could draw from 4.6 million records which represented nearly 100% of the employees working in the respective locations. To answer the research questions, we aim to find a pattern in the activity distribution from day 1 to day 180 (six months). The evaluation of this final dataset has been done using SPSS, providing us with all statistical variables. To calculate the necessary correlations, we chose to use the Pearson correlation because all our scales were single-item scales [43].

4 Results

Table 1 provides the output of the Pearson correlation and the linear regression analysis that was done on the dataset. The table indicates the underlying n per country (*Users*) representing the active user count. It is sorted by the Pearson coefficient in an ascending order. The results show that countries with a negative and positive correlation coefficient are balanced. Negative coefficients show, that activity decreases while the time increases (increase of days after go-live). All results show a high significance with $p<0.01$ (indicated with ** in table 1) and mostly a strong correlation (all coefficients >= .279). Categorizing the results by positive and negative correlation coefficients, we observe that countries prevailing with a descending activity (negative correlation) are all located in the Asian region. The opposing pattern - a positive correlation - is reflected by a mix of countries. Three quarter are in the European or American region (Germany, United States and United Kingdom) and the fourth country is Thailand, located in the Asian region.

Table 1. Explorative analysis (activity over Time; significance indicated with ** = p<.01)

Country	*Users*	*Total Voice Conversations*	*Pearson's Corr. Coeff.*	*Gradient*	*Regression Constant*	R^2
China	785	459,971	-.725**	-.012	6.682	.526
South Korea	414	10,100	-.674**	-.011	5.590	.454
Japan	185	163,190	-.422**	-.004	5.093	.178
Singapore	68	43,251	-.279**	-.003	5.510	.078
United States	223	75,269	.485**	.010	3.089	.236
United Kingdom	162	61,931	.569**	.014	4.557	.323
Thailand	65	19,434	.642**	.011	4.662	.117
Germany	4407	3,501,534	.650**	.005	4.874	.422

The results show that countries with a negative Pearson correlation also have a negative gradient, ranging from -.012 (China) to .014 (United Kingdom), due to the similar nature of the analysis. The initial activity on day 1, expressed by the constant of the regression, is showing the highest values for countries with the lowest gradient (or the lowest correlation coefficient respectively). We found R^2 values averaging at .292, with a maximum of .526 (China) and a minimum of .078 (Singapore). We base further analysis mainly on Pearson's coefficient, descriptively supported by the regression gradient and the constant. To visualize the dataset and the undertaken statistical analysis, figure 2 and figure 3 show opposing examples for the activity distribution.

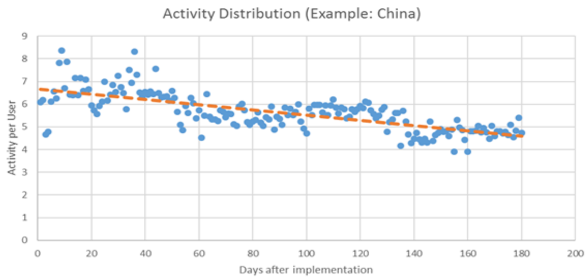


Figure 2. Negative correlation coefficient with decrease of activity (example China)

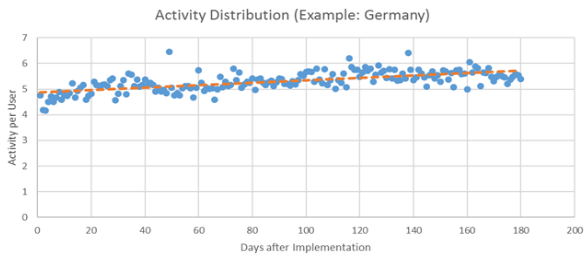


Figure 3. Positive correlation coefficient with increase of activity (example Germany)

The regression function is visualized in orange. The blue dots each represent the averaged activity data (per user) gathered for one day as introduced earlier. As the analyzed features were fully available to all users right from day one, a comparable situation is given. However, we observed that in each country the number of users increased with a different speed. Further statistical tests have shown a significant positive correlation between the increasing user count and the user activity. Countries with a positive correlation of *activity* and *time* (table 1) show a reluctant initial adoption (as in actively using the technology) which is strongly increasing over time. Countries with a negative correlation tend to adopt quickly with a higher initial user count and continue with a steady (slow) linear increase of additional active users over time.

As figure 1 indicates, we aim to exploratively analyze the association between Halls' and Hofstedes' factors and the results of our previous analysis. To do so, we collected the factors (table 2) Power Distance (PDI), Individualism/Collectivism (IDV), Masculinity (MAS), Uncertainty Avoidance (UAI) and Long-term Orientation (LTO) from Hofstede, as well as Low-/High Context (LHC) and Mono-/Polychronic (MPC) from Hall [29], [33], [34], [39]. A limitation is that the most recent index of Hofstede, the *Indulgence vs. Restrain* index was not included, because it was just established in year 2010, a sufficient amount of data could not be collected and it has not been verified by as many studies as the other indexes to be representative [40]. It must be noted that there was no MPC value for Thailand derivable from [34].

Table 2. Cultural factors for regression analysis

Country	*PDI*	*IDV*	*MAS*	*UAI*	*LTO*	*LHC*	*MPC*
China	80	20	66	30	87	15	12
South Korea	60	18	39	85	100	10	10
Japan	54	46	95	92	88	16	6
Singapore	74	20	48	8	72	10	10
United States	40	91	62	46	26	3	2
United Kingdom	35	89	66	35	51	5	4
Thailand	64	20	34	64	32	10	-
Germany	35	67	66	65	83	1	1

Especially for Hall's indexes we observe, that a negative correlation reflects a high LHC and MPC value. Hofstede's indexes MAS and UAI show mixed values and no pattern towards a correlation with the first analysis. IDV and LTO on the other hand show possibilities for a pattern. IDV is low for countries with a negative correlation of activity and time and LTO is high for those countries. PDI shows a weak pattern, where high power distance reflects in a negative activity correlation.

5 Discussion and Limitations

With the first part of the analysis we could show that for different countries the adoption process looks different. Our explorative analysis returned attributes for each country that indicate a decrease or an increase of activity after the introduction of a new technology. Hence, we can cluster the countries accordingly, which represents the focus of this study. The goal was to analyze the activity over time and find patterns directly after the go-live. With the explorative approach, we applied a correlation analysis and found a tendency for four Asian countries to start with a high activity and then decrease. Additionally, we observed that the stronger the correlation was between *time* and *activity*, the higher (for negative correlation) or lower (for positive correlation) an initial activity level on day 1 was. With the mentioned results, we can answer research question RQ1, since 2 patterns have been found: (1) countries with low activity at the beginning and an increase over time and (2) countries with a high start-activity and a decrease.

Even though all countries from (2) are Asian countries, we draw from a limited dataset and for a conclusion overall Asian region, further validation is necessary. Furthermore, the shortcomings of a regression analysis with *time* as an independent variable were known and accepted, because the focus was to generate as many descriptive parameters as possible. Since the results of the linear regression were not used for prediction but rather for description, we believe the parameters contribute to understanding the various activity pattern. As limitation it is to say, that the values we found for Singapore did not show a very strong correlation (-.279), although being highly significant. This is also reflected in the regression that was applied, where Singapore and Thailand had a noteworthy lower R^2 and lower number of users, forming the base for our data. Results for those two countries are therefore not as representative as the others.

As stated earlier, digital transformation is heavily depending on a global collaboration of individuals and hence a globally synchronous usage of such is desirable. Given this, our study contributes in a unique way, since it is significant to IT decision makers, that they need to assume different usage adoptions in different regions. Additionally, with the digital transformation, all enterprises will have to struggle with new technologies that are emerging [41]. Hence the frequency of technology adoption that a user is going to face will increase. As new technologies are to be piloted first [42], our results suggest that this should be done in Asian countries preferably. Since a quick initial adoption was noticed and the activity was at first relatively high, we conclude that this carries the benefit of immediate pilot feedback regarding the technology. As initially stated, RQ1 aimed to provide descriptive insights into the adoption phase after go-live, whereas RQ2 provides an approach to identify reasons for the analyzed patterns. However, it was not possible to answer RQ2 with significant reliability, since the dataset we drew from did not contain enough countries to run representative statistical tests. But we could see a strong tendency which helps to understand the results of our first analysis. For our research, we derived and compared especially Hofstede's factors from various sources, as well as his website. Interestingly, the newest factor long-term orientation also reflects the highest correlation and r^2 in our results.

Nevertheless, like every research, our analysis also comes with limitations, such as a seasonal or project based bias. We could not ensure that the observed regions do not have seasonal variability or activity peaks due to projects. We filtered those to the best of our knowledge and made sure with each local responsible, that no unusual dynamic prevails. Furthermore, we were only able to collect a limited amount of time and hence focused on the first six months after go-live. Therefore, we were not able to make further reliable long-term prognosis and further strengthen our results. The underlying data which was used to derive assumptions, is based on a large multi-national enterprise with a European headquarter. Therefore, the results might differ with a dataset from various smaller companies or different technologies for example. As mentioned earlier, to generate further descriptive parameters we applied a linear regression analysis which is feasible for the limited dataset we drew from. However, the results and common sense indicate that a logarithmic model will fit better for longer time frames as the activity will not reach a negative level, but will get close to a threshold. Also, our findings are derived from one case company only. Therefore, even though based on a large data set, the results cannot automatically be transferred to all organizations (or regions as mentioned earlier) without considering the individual contexts.

6 Conclusion and Further Research

In this study, we focused on the adoption pattern of a digital technology in a global, cross-cultural context. Analyzing eight different countries across the globe, we found that employees react differently, depending on the region they come from. Our results indicate that Asian countries adopt the technology initially quickly, which is reflected in a high initial activity among users. It then decreased steadily over time and settled at a certain level. European and American countries (together with Thailand which is not fitting the pattern) show an opposing behavioral pattern. Due to limited data availability, we were not able not extent our research to the level at which the activity finally settles. We encourage other researchers to pick up on this research area and investigate the long-term usage of communication technology with regards to user activity.

Furthermore, we found that within a digital transformation, an isolated look at individual countries does not make sense anymore, which is in line with [42]. Global actions need to keep local cultures in mind. The contribution of this study indicates that not all cultural factors are equally vital in the adoption process and its different characteristics. We propose that several cultural factors do correlate with the findings of our first research question. Attributes like mono-/polychronic [38] seem to predict the user adoption well. We could provide descriptive insights on the adoption pattern after go-live which we believe is a novel contribution to the community and an important first step for further studies in this direction. Due to this, our study is relevant to scholars for further research, as well as for practitioners in enterprises of global scale. Even though we were not able to reliably answer the second research question – examining the correlation between cultural indexes and the adoption patterns – we encourage for further studies while picking up these results.

The contribution for decision makers indicates that particularly the change management process within their digital transformation should not look the same globally. Our results suggest adjusting the marketing actions regionally, depending on the region.

References

1. Belz, C.: Marketing Heatmap I – digitale Transformation und strategische Optionen. Swiss Marketing Digital Transformation (2016).
2. Kane, G., Palmer, D., Phillips, A.N., Kiron, D., Buckley, N.: Strategy, not Technology Drives Digital Transformation: Becoming a Digitally Mature Enterprise, http://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/ (Accessed: 15.01.2017).

3. Atkinson, R.D., Castro, D.D.: Digital Quality of Life: Understanding the Personal & Social Benefits of the Information Technology Revolution. Information Technology and Innovation Foundation,(2008).
4. Kaak, E.-J.: Mention Communication—Think Organisation: Agile Communication in the Digital Era. In: Klewes, J., Popp, D., Rost-Hein, M. (eds.) Out-thinking Organizational Communications, pp. 129-143. Cham: Springer International Publishing (2017).
5. Klewes, J., Popp, D., Rost-Hein, M.: Digital Transformation and the Challenges for Organizational Communications: An Introduction, pp. 1-6. Cham: Springer International Publishing (2017).
6. Hulme, R.: One for all: Unified messaging comes of age. Work Study 52, 141–144 (2003).
7. Fikry, A., Abdul G., Mukhtar, Z.: Unified communication: It's all between you and me. Business Strategy Series 13, 168–172 (2012).
8. Leidner, D.E., Kayworth, T.: A Review of Culture in Information Systems Research: Toward a Theory of Information Technology Culture Conflict. MISQ 30, 357–399 (2006).
9. Neumeier, A., Wolf, T., Oesterle, S.: The Manifold Fruits of Digitalization - Determining the Literal Value Behind. In: Internationale Tagung Wirtschaftinformatik (2017).
10. Hanelt, A., Piccinini, E., Gregory, R.W., Hildebrandt, B., Lutz M.K.: Digital Transformation of Primarily Physical Industries – Exploring the Impact of Digital Trends on Business Models of Automobile Manufacturers. In International Conference on Wirtschaftsinformatik (2015).
11. Pagani, M. Digital Business Strategy and Value Creation: Framing the Dynamic Cycle of Control Points. MISQ 37, 617-632 (2013).
12. Fichman, R.G., Dos Santos, B.L., Zheng, Z.: Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. MISQ 38 (2014).
13. Röglinger, M., Urbach, N.: Digitale Geschäftsmodelle im Internet der Dinge. Forum für Verbraucherrechtswissenschaft (2013).
14. Klewes, J., Popp, D., Rost-Hein, M.: Out-thinking Organizational Communications. Cham: Springer International Publishing (2017).
15. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.V.: Digital Business Strategy: Toward a Next Generation of Insights. MISQ 37, 471-482 (2013).
16. Kessler, T., Buck, C.: How Digitization Affects Mobility and the Business Models of Automotive OEMs. In: Khare, A., Stewart, B., Schatz, R. (eds) Phantom Ex Machina, pp. 107-118, Cham: Springer International Publishing (2017).
17. Shen, Z., Fluker, J.: Impact of Unified Communications on Communication, Relationship Building and Performance. In: ICIS Proceedings (2013).
18. Elliot, B., Fernandez, M.M., Blood, S.: Magic Quadrant for Unified Communications. https://www.gartner.com/doc/reprints?ct=160714&id=1-3BON1KW&st=sb
19. Chen, C.C., Ractham, P.: Adopting Internet Telephony Technology as a Cross-Cultural Communication Training Tool. In: PACIS Proceedings (2013).
20. Mitchell, A., Chen, C., Medlin, B.D.: E-Learning to Improve Intercultural Communication. In: AMCIS Proceedings (2010).
21. Caporarello, L., Magni, M., Pennarola, F.: Does Culture Make the Difference? Technology Acceptance and Diffusion in India. In: MCIS Proceedings (2014).
22. Triandis, H.C.: Culture and Conflict. International Journal of Psychology 35, 145–152 (2000).
23. Dasgupta, S., Gupta, B.: Antecedents of Internet Technology Adoption in an Emerging Economy (2015).
24. Ocasio Velázquez, M., Ho, S.M.: Do You Trust the Cloud? Modeling Cloud Technology Adoption in Organizations (2015).

25. Talebi, M.: Role of Trust, Innovation and Attraction in Adoption of Informational Systems within Organizations. Indian Journal of Science and Technology 8 (2015).
26. Ryschka, S., Bick, M.: Drivers of Cultural Differences in Information System Adoption - A Case Study. In: AMCIS Proceedings (2013).
27. Khan, A.I., Al-Shihi, H., Al-khanjari, Z.A., Sarrab, M.: Mobile Learning (M-Learning) adoption in the Middle East: Lessons learned from the educationally advanced countries. Telematics and Informatics 32, 909–920 (2015).
28. Fernandez, D.R., Carlson, D.S., Stepina, L.P., Nicholson, J.D.: Hofstede's Country Classification 25 Years Later. The Journal of Social Psychology 137, 43–54 (1997).
29. Hofstede, G.: Cultural dimensions in management and planning. Asia Pacific Journal of Management 1, 81–99 (1984).
30. Lee, S.-G., Trimi, S., Kim, C.: The impact of cultural differences on technology adoption. Journal of World Business 48, 20–29 (2013).
31. Faqih, K.M.S., Jaradat, M.-I.R.M.: Assessing the moderating effect of gender differences and individualism-collectivism at individual-level on the adoption of mobile commerce technology. Journal of Retailing and Consumer Services 22, 37–52 (2015).
32. Jing, P., Rong, G.: Exploring Cross-Cultural Factors Affecting Mobile Commerce Adoption: Theories and Empirical Comparison between U.S and China. In: International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII), Kunming, China, 324–327 (2010).
33. Hofstede, G.: Country Comparison. https://geert-hofstede.com/countries.html. (Accessed: 15.01.2017).
34. van Everdingen, Y.M., Waarts, E.: The Effect of National Culture on the Adoption of Innovations. Marketing Letters 14, 217–232 (2003).
35. Straub, D., Keil, M., Brenner, W.: Testing the technology acceptance model across cultures: A three country study. Information & Management 33, 1–11 (1997).
36. Ou, C.X.J., Davison, R.M., Zhong, X., Liang, Y.: Empowering employees through instant messaging. Information Technology & People 23, 193–211 (2010).
37. Hofstede, G.H.: Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations / Geert Hofstede. Thousand Oaks, SAGE: London (2001).
38. Hall, E.T.: Beyond culture. New York: Anchor books (1976).
39. Bandyopadhyay, K., Fraccastoro, K.A.: The Effect of Culture on User Acceptance of Information Technology. CAIS 19 (2007).
40. Kohun, F.G., Burčik, V., Skovira, R.J.: Research into Hofstede's Thesis. In: Management, Knowledge and Learning (2012).
41. Agarwal, R., Gao, G., DesRoches, C., Jha, A.K.: Research Commentary —The Digital Transformation of Healthcare: Current Status and the Road Ahead. Information Systems Research 21, 796–809 (2010).
42. Arayici, Y., Coates, P., Koskela, L., Kagioglou, M., Usher, C., O'Reilly, K.: Technology adoption in the BIM implementation for lean architectural practice. Automation in Construction 20, 189–195 (2011).
43. Zhu, K., Dong, S., Xu, S.X., Kraemer, K.L.: Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of European companies. European Journal of Information Systems 15, 601–616 (2006).
44. Eisinga, R., Grotenhuis, M. Te, & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? International Journal of Public Health, 58(4), 637–642.

Enablers and Inhibitors of Engagement in Enterprise Social Networks from the Viewpoint of Executives – A Case Study in a Technology Company

Janine Hacker[1], Natalie Florian[1], and Freimut Bodendorf[1]

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute of Information Systems, Nürnberg, Germany
{janine.hacker,natalie.florian,freimut.bodendorf}@fau.de

Abstract. Research on Enterprise Social Networks (ESN) agrees on the significance of top management commitment in the implementation of ESN in companies and identifies a lack thereof as one of the most common reasons for ESN initiatives to fail. While middle- and lower-level managers are suggested to be just as, or even more, important in promoting ESN, only few studies focus on their usage and perception of ESN. Against this background, in this paper, we explore enablers and inhibitors of ESN adoption and usage based interviews with executives at a technology company. We find the number of inhibitors to strongly exceed the number of enablers. In particular, technological and organisational inhibitors affect ESN adoption and usage. Our study contributes to literature on factors influencing whether and how employees use ESN. As such, it highlights the importance of a fit between organisational culture and the principles of social technologies.

Keywords: Enterprise social networks, adoption and usage, middle managers, digital leadership, governance.

1 Introduction

Enterprise Social Networks (ESN) have emerged as important technologies to support collaboration and knowledge work [1]. In the context of Enterprise 2.0, ESN promote the digital transformation of companies by enabling employee participation, continuous organisational learning, and overall, a more bottom-up organisational culture [2-3]. Along with the increasing adoption in companies, ESN have become an important topic in Information Systems (IS) research [4]. However, there remain unsolved challenges, one of which is how to implement and manage ESN successfully (e.g., [5]). ESN research generally agrees on the significance of top management commitment in the implementation of ESN and identifies a lack thereof as one of the most common reasons for ESN initiatives to fail (e.g., [6]). In addition, research by Chelmis and Prasanna [7] as well as Majumdar et al. [8] point to a major influence of middle managers and direct supervisors on employees to adopt social software. Only few studies, however, have focused on the ESN usage, perception, and role in ESN initiatives of middle and lower-

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

level managers [8-9]. Against this background, in this study, we focus on the following research question: *What are factors enabling or inhibiting, i.e. enablers or inhibitors of, ESN adoption and usage from the perspective of executives?* By exploring what leads executives and their teams to use or reject an ESN, our goal is to better understand their role in ESN adoption processes. To this end, we conduct an in-depth case study including interviews with executives at a multinational technology company. Drawing on prior work by Chin et al. [10], we classify the identified enablers and inhibitors according to the categories technological, organisational, social, and individual factors. The remainder of this paper is structured as follows: In the next section, we give an overview of related work on factors influencing ESN adoption and usage as well as the role of management in ESN initiatives. Section 3 introduces the case and provides details regarding the data collection and analysis. Next, we present the findings of the in-depth case study. The final sections summarise and discuss our findings in the light of prior research and outline the next steps of this project.

2 Literature Review

Based on prior work, ESN are defined as web-based intranet platforms that allow employees to send messages to specific co-workers or everyone in the organisation, to connect with others, to contribute content as well as to view the content and connections of others [11]. ESN integrate features such as profile pages, following, activity streams, search, group capabilities, discussion threads, and tagging [12]. In the following, we present related studies focusing on ESN adoption and usage as well as on the role of management in initiatives to introduce ESN.

2.1 Factors Influencing ESN Adoption and Usage

Prior scientific research as well as practice-oriented literature have dealt extensively with factors influencing the adoption and usage of ESN. As for adoption, previous work focuses on factors influencing the *decision to use* an ESN. For instance, the influence of privacy concerns [13], organisational climate [14] and promotional messages posted on an ESN [15] on employees' decision to engage in an ESN are investigated. Commonly identified aspects inhibiting adoption and usage include a lack of top management involvement and commitment, and an incompatible organisational culture, i.e., a culture that does not support social communication, collaboration, and knowledge sharing [16-17]. A comprehensive framework of factors enabling (+) and inhibiting (-) ESN usage is derived by Chin et al. [10]. As such, it differentiates between enablers and inhibitors of ESN usage in terms of *technological* factors (e.g., accessibility (+), competing technologies (-)), *organisational* factors (e.g., reward system (+), lack of well-defined purpose (-)), *social* factors (e.g., collaborative climate (+), lack of feedback (-)), and *individual* factors (e.g., sociable personality (+), lack of ESN skills (-)).
Besides studies focusing on the users' intention to use an ESN and their overall level of engagement, factors and motives influencing *how* users engage in ESN are investigated in prior work. For instance, Osch et al. [18] reveal a diverse set of motivations,

among them personal reputation building and resource sharing, underpinning user behaviour. Such motivations are closely related to users' perceived value from using an ESN. In this regard, utilitarian values, e.g., easier competence sharing [16], are derived from participating in an ESN. Enabling users to build relationships as well as to receive feedback and feel involved [19-20], ESN engagement also creates social values. Moreover, users may perceive a hedonic value in using an ESN, for instance, when using a platform for the purpose of entertainment or self-fulfilment [21].

2.2 The Role of Management in ESN Initiatives

As mentioned above, management plays a crucial role in implementing ESN and fostering platform engagement. In terms of implementation strategies, Richter and Stocker [22] differentiate between exploration, i.e. a bottom-up implementation driven by employees, and promotion, which refers to a top-down strategy facilitated by management. Especially when pursuing a promotion strategy, management needs to define the business case and clear goals of the ESN upfront as well as motivate employees to use the new tool [22]. While such strategic tasks are often in a company's top management's area of authority, executives in lower hierarchical positions, i.e. individuals with managerial authority over a group of employees at different levels [23], are suggested to significantly influence social software utilisation in their respective teams (e.g. [24]). In a study by Chelmis and Prasanna [7], middle managers are found to be more successful in promoting the adoption of a microblogging service than higher-level managers. As such, executives need to act as role models in order to drive ESN adoption. Yet, the introduction of social software entails changes in organisations which, in turn, have consequences for leadership [25]. While managers recognise benefits of social software, among them better access to information, improvements in communication, as well as the opportunity to establish personal contacts [8], in many cases, they are hesitant to engage in and take on responsibility for promoting the use of these platforms. According to a survey by van Dick et al. [26], this lack of engagement may partly be due to a lack of competencies and skills, e.g., skills in terms of using social media for work-related purposes, communicating on an equal footing with employees as well as fostering networking among employees.
In conclusion, prior work has dealt with factors determining ESN adoption and usage perceived by users in general. Due to their importance in driving ESN adoption, the factors influencing ESN engagement from the perspective of executives need to be better understood. As such, these insights indicate the extent to which executives act as digital leaders and help understand how they can be supported in becoming the same.

3 Research Method

Our study seeks to explore factors influencing ESN adoption and usage from the perspective of executives. Drawing on our related work analysis (section 2.1) and in particular the study by Chin et al. [10], we consider the influence of technological, organisational, social, as well as individual factors in shaping the perceived value of ESN use, which in turn, impacts adoption and usage (Figure 1). Within these categories, we differentiate between factors enabling (+) and inhibiting (-) adoption and usage, i.e. enablers and inhibitors of ESN adoption and usage. Due to the exploratory nature of this study, a qualitative research design was selected. Specifically, a case study [27, p. 46 ff.] including semi-structured interviews [28, p. 181] was conducted. The following sections introduce the case and provide details on the data collection and analysis.

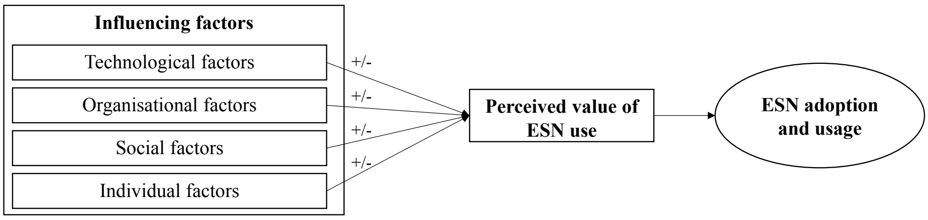


Figure 1. Factors influencing ESN adoption and usage

3.1 Case Organisation and ESN Platform

This case study is conducted with a globally operating technology company that belongs to a large corporation. The case company's ESN was launched in 2013 based on an initiative of the parent company. It allows users to create a profile, connect with other users, to join groups, and access news in a general or customised activity stream. Dealing with e.g. project work or common interests, groups have emerged as the most used communication channel on the ESN. Moreover, users may upload files, share links, start polls, highlight certain articles as new ideas, search for specific topics and keywords, and create own challenges to collect ideas on a topic area or find solutions for problems. Furthermore, the ESN features virtual meetings, which are similar to a live chat and facilitate moderated discussions. Including these features, the platform was launched to enable collaborative work and to promote the exchange of knowledge and ideas among employees across geographic and hierarchical boundaries. To date, the level of engagement on the ESN is generally low and varies strongly across different departments. Especially the company's top management shows a lack of engagement.[1]

3.2 Data Collection and Analysis

In preparation for the interviews, an interview guide [29, p. 43] was compiled to support the discussion with the participants and to enable comparability of the answers. Organised into different topic areas, the interview guide includes questions concerning the ESN implementation within the organisation, the interviewee's behaviour on the platform (e.g., "*Please describe your usage of the ESN during a typical working day.*"), as well as perceived enablers and inhibitors of ESN adoption and usage (e.g., "*What motivates you and your team to / prevents you and your team from using the ESN?*").

[1] For confidentiality reasons, detailed statistics regarding the usage of the ESN were not provided by the case company.

While the ESN is available to members of the case company worldwide, for this study, only participants located at one of the company's German offices were selected. The sampling strategy aimed at achieving variation regarding adoption and usage levels within the interview participants' departments, as well as in terms of their hierarchical level (e.g., department manager versus executive), and gender. Interview candidates were contacted via email and, in total, 12 managers agreed to take part in the study. As can be seen in Table 1, the 12 interviewees work in seven different departments and hold (management) positions in three different hierarchical levels. In this regard, nine interviewees are team leaders with a personnel responsibility of up to 25 employees. I3 and I9 are department managers and I10 is member of the company's top management. The usage levels within the different business units range between *medium-high* (BU1), *medium* (BU3, BU5), *low-medium* (BU2), and *low* (BU4, BU6, BU7).

All interviews were held in German language at one of the company's offices over a period of two months from mid of April to mid of June 2017. Eleven interviews were conducted face-to-face and one interview was done over the telephone. The interviews, which have an average length of about 33 minutes, were recorded using a digital voice recorder and transcribed verbatim. Using the software for qualitative content analysis "MAXQDA", the material was analysed following the steps of a structuring content analysis by Mayring [30]. Accordingly, the topics included in the interview agenda and elements of Figure 1 served as given categories, which were assigned deductively to applicable passages of the material. Next, relevant statements of the interviewees were filtered out, compiled in an Excel table, paraphrased, and summarised. For instance, statements regarding perceived advantages of ESN use due to the platform's features (e.g., "*It allows me to share information quickly*" (I1)) were classified as enablers in the category *technological factors*.

Table 1. Overview of interviewees

Interviewee	*Gender*	*Business unit*	*Personnel responsibility*	*Level of ESN usage*
I1	m	BU1	1-25	high
I2	m	BU2	1-25	low
I3	m	BU3	>200	medium
I4	m	BU4	1-25	medium
I5	m	BU5	1-25	high
I6	m	BU5	1-25	low
I7	m	BU2	1-25	low
I8	m	BU5	1-25	high
I9	m	BU1	>200	low
I10	m	BU6	>200	low
I11	f	BU2	1-25	medium
I12	f	BU7	1-25	low

4 Findings

This section presents findings regarding the ESN implementation at the case company, the usage behaviour of the interviewees as well as the identified enablers and inhibitors of platform adoption and usage.

4.1 Implementation of the ESN and Perceived Value

Concerning the implementation, a project group in the marketing communications department informed the executives about the ESN in an email, which included links to the platform itself as well as video tutorials and further information material (I2, I9). As neither upper management nor the project group in charge did provide any specific instructions on the ESN (I2, I6, I7), many of the interviewees arranged team meetings to discuss how the ESN could be integrated in their work (I3, I8).

In terms of the ESN's purpose, half of the interviewees (e.g., I3, I5, I8) consider it as a tool to share and acquire information and knowledge. In addition, interviewees emphasise the use of the ESN as a collaboration platform and additional communication channel with worldwide reach. Consequently, the perceived value of the platform relates to using the ESN as a tool to find information and keep oneself up-to-date within the various groups (I1, I5, I8). Also, the ESN facilitates easier and faster access to experts on specific topics and solutions for problems: *"[…] there are requests by people looking for a solution […] I have already replied to such requests […] because I knew someone who should know it.*" (I8). In the team of I11, using the ESN contributes to improved knowledge transparency and to the reduction of knowledge silos. It further enables informal learning (I4) as well as higher levels of participation and self-organisation in teams (I3, I8, I11): *"[...] using the ESN led to a much faster information flow and more intensive information exchange as compared to traditional methods*." (I3). In addition, I8 appreciates the possibility to collaborate with colleagues abroad using the ESN: "*[...] a group of people not working at the same location has access to a single knowledge base. […].* (I8). On the other hand, about half of the interviewees (e.g., I6, I9, I10) do not see an additional benefit in the ESN and rather rely on other existing applications.

4.2 Usage of the ESN

While the creation of an account on the ESN is not mandatory, all interviewed managers have a profile and are members of various groups. Yet, they use the ESN to different extents. As such, the group function is the most frequently used feature, followed by the activity stream. While more than half of the interviewees do not create posts in groups, they use them to keep themselves up-to-date about news in their community: "*I follow certain topics and am member of these groups, but I am not really active. Rather, I'm a listener.*" (I8). Others take a more proactive stance in terms of using the ESN: *"[...] I use the ESN for sharing information. When I am on business trips, I keep my employees in the loop about what is happening."* (I3).

On the other hand, half of the interview participants show very low levels of ESN usage (Table 1). Although they have a profile and are members of two to four groups, they mainly use other tools to perform their tasks (I2, I7, I9). Also, except for I1, none of the interviewees has ever initiated a virtual meeting either due to not being aware of this function or due to having no suitable use case for it.

4.3 Enablers of ESN Adoption and Usage

Enablers facilitating ESN adoption and usage are derived based on statements of those executives using the platform themselves and within their teams. As such, these factors relate to the perceived value of the ESN (section 4.1) and the positive outcomes achieved from using it. The identified aspects can be associated with the categories *technological* factors, *organisational* factors, and *individual* factors (Figure 1).

In terms of *technology*, enablers relate to the ESN's *usability and range of features*. As such, employing the ESN within the interviewees' day-to-day work leads to gains in efficiency and effectiveness, which motivate further use of the ESN. Specifically, using the ESN, e.g., to keep oneself up-to-date or finding solutions for problems, is perceived as timesaving and convenient: *"An advantage in comparison to email is that […] all the information is stored in one place and can be accessed by all members of a team […].*" (I11). Providing a uniform communication channel for everyone, the ESN also contributes to a more open corporate culture and improves knowledge work (I1, I8).

In terms of *organisational factors*, the distribution of work across many different locations, which is related to the company's *organisational structure*, entails a need to communicate efficiently across geographic boundaries. This need is met by the ESN: *[…] you can collaborate virtually around the clock. Which is very important […] as we have teams all over the world."* (I8). As such, the geographic distribution can be considered as an enabler of ESN adoption and usage.

As for *individual factors*, several *motivations underpinning usage* can be identified. For instance, experts contribute to the ESN due to their intrinsic motivation to help others (I1, I5). Also, the continuous contribution of relevant content may strengthen one's reputation in the company (I11). Finally, individuals are motivated to participate in the ESN due to seeking (informal) exchange and learning from each other (I3, I4).

4.4 Inhibitors of ESN Adoption and Usage

Inhibitors of ESN adoption and usage are determined based on statements by executives contributing to the ESN (I1, I6) as well as by users focusing on consuming content (I3, I5) or hardly using the platform at all (I3, I9, I10). From the viewpoint of the contributing users, the mitigation of these inhibiting aspects would likely incite them to become even more active on the ESN. The derived aspects can the assigned to the categories *technological*, *organisational*, *social*, and *individual* factors.

As for *technological factors*, inhibitors related to the platform's *usability and range of features* are suggested to affect adoption and usage. Besides being considered as too complex and to lack structure (I15), interviewees criticise the support of the ESN for mobile devices (I1, I9): *"[…] if I have to make three detours, […] before I can post something, my motivation is gone already.*" (I9). Moreover, important features, such as the uploading and sharing of larger files, a filtering function for the search engine, the possibility to directly export communications (e.g., to a text document) and a single sign-on, are currently missing (I4, I6, I7). Also, inhibitors related to *information security* affect platform adoption and usage. The fact that content once posted to the platform is persistent and can be associated with its authors years later as well as uncertainty with regards to who will have access to content in the future, e.g., in case parts of the business are split off and new subsidiaries are created, worries the executives (I2). As a result, many discussions still take place off rather than on the ESN. In terms of the ESN's *integration into the existing IT landscape*, the interviewees affirm the lack of integration with other tools as well as the existence of competing technologies and tools (e.g., chat) to affect ESN usage (I4, I10): "*[...] It would be helpful and more efficient if one could switch directly from the ESN to other tools [...].*" (I4).

Concerning *organisational factors*, inhibitors related to the company's *top management and ESN strategy* as well as *ESN management* are identified. As for the *top management and ESN strategy*, the interviewees report the lack of a clear vision as well as a lack of support and involvement on the part of the senior management (I1, I6). Due to the lack of clearly defined goals and role models, many of the interviewed executives neither perceive the platform as a priority nor do they see any value in using the tool (I2). Consequently, from the top management via the lower levels of management through to the employees, the lack of role models entails as lack of engagement. In terms *of ESN management*, a missing code of conduct concerning what may be shared and discussed on the ESN and how it should be used or not used is suggested to inhibit usage. In combination with the lack of role models, potential contributors are unsure what may be shared and thus, lack the confidence to post something on the ESN. Furthermore, there is currently no systematic way to keep information on the platform up-to-date, which leads to a lot of content being obsolete (I5) or irrelevant (I8). In addition, the fact that confidential information must not be shared on the platform prevents users from integrating the ESN efficiently in their day-to-day work and hampers in-depth conversations (I2, I10).

As for *social factors*, especially the lack of a *critical mass* of users inhibits adoption and prevents contributing executives from become more active. On the one hand, the general lack participation directly leads to a lack of feedback on posted content (I4, I7), which demotivates the contributing users. On the other hand, many co-workers, among them experts for certain topics, cannot be reached on the ESN as they are not registered (I8). Also, the company's *corporate culture* appears to be not (yet) compatible with the ESN. In this regard, the combination of a top-down management style and the lack of engagement on the part of the upper management (see above) to some extent prevents lower management levels from promoting the use of the ESN in their teams. Furthermore, the case company has only little understanding concerning a tolerance for failure, which hinders open conversations and the discussion of mistakes on the ESN (I11). Along these lines, the visibility of communications on the ESN is suggested to lead to feelings of pressure, for instance, to respond to posts with a short delay (I1, I10).

In terms of *individual factors*, inhibitors relate to the *attitude towards the ESN* as well as a *preference for other means of communication*. In this regard, many executives are sceptical towards the ESN and do not understand why they should invest time to deal with "yet another tool", the benefits of which are not clear (I2, I7). Indeed, participating in the ESN is considered as an additional burden instead of contributing to a reduction of an individual's workload. Furthermore, some executives perceive other means of communication, e.g., face-to-face interactions, as necessary for collaborating efficiently: *"[…] I share content on the ESN […] To get real feedback, I need to meet occasionally with the people who I interact with on the ESN […]*" (I1).

5 Discussion

Based on the presented findings, Table 2 shows the technological, organisational, social, and individual enablers (7) and inhibitors (20) of ESN adoption and usage. Most identified enablers, among them better task support, are based on the perceived (utilitarian) benefits from actual platform use. While the identified enablers are similar to the ones suggested in prior work (e.g. [10]), many previously found drivers, e.g., a sense of connectedness, do not apply to our case. Also, values obtained from a hedonic or social use [21] are less indicated in our analysis. Furthermore, the lack of certain aspects at the case company, which are suggested to drive engagement in prior work (e.g., tolerance for failure or policy [10]), in fact inhibits adoption and usage (e.g., little tolerance for failure, missing code of conduct). On the other hand, most inhibitors relate to technological and organisational factors. The identified inhibitors can be connected with factors suggested to negatively influence ESN usage (e.g. [16-17]) and to obstruct the integration into work from the perspective of managers [9].
In comparison to findings of prior work (e.g., [10]), our analysis points to strong imbalance between the number of enablers and inhibitors. Possibly due to the generally low level of platform engagement, many advantages which might drive further adoption have not (yet) been realised at the case company. In this regard, the overall ESN usage level in a department does not necessarily correlate with individuals' usage levels (section 3.2). Indeed, the few contributing interviewees engage with the ESN on their own initiative, rather than due to external motivators. In addition, the interviews include many and various statements regarding inhibitors (especially by the interviewed occasional and non-users of the ESN) as compared to few statements regarding enablers. As such, ESN adoption and usage at the case company appear to be particularly hindered by the co-occurrence of a lack of top management commitment, strict confidentiality requirements and a top-down management style. In fact, it is interesting to see that enablers and inhibitors of ESN usage suggested by the interviewed managers show a strong overlap with previously identified enablers and inhibitors perceived by employees in general (e.g. [10]). While some interviewees assume responsibility in terms of promoting the ESN, hence driving adoption, most of them consider the top management to be in charge for making the ESN a priority. As such, they do seem to be aware of their power and / or do not see a need to drive ESN engagement even though their management position would give them authority to do so. In this regard, findings of studies conducted at other organisations suggest that neither formal responsibility nor authority are essential for promoting an ESN [31]. Due to the hierarchical organisation of the case company, however, it appears that ESN adoption and usage can only be achieved through a common effort of the company's top management and executives

at different hierarchical levels. The engagement and support of leaders in recognising and embracing their role is a major task in the process of digital transformation [25].

Table 2. Enablers and inhibitors of ESN adoption and usage

Category	*Factors*	*Enablers (+) and inhibitors (-)*
Technological factors	ESN's usability and range of features	Increases in efficiency and effectiveness in day-to-day work (+); Better task support (+); High complexity of the user interface (-); Lack of support for mobile use (-); Missing features (-)
	Information security	Persistency of content (-); Uncertain future access (-)
	Integration into IT landscape	Lack of integration (-); Competing tools (-)
Organisational factors	Organisational structure	Geographic distribution (+)
	Top management and ESN strategy	Lack of a clear vision (-); Lack of top management commitment (-)
	ESN management	Missing code of conduct (-); Obsolete / irrelevant content (-); Missing content due to confidentiality requirements (-)
Social factors	Critical mass	Lack of critical mass (-); Lack of feedback (-)
	Corporate culture	Top-down management style (-); Little tolerance for failure (-); Social pressure (-)
Individual factors	Motivations underpinning usage	Enjoyment of helping others (+); Reputation management (+); Desire to learn (+)
	Attitude towards ESN	Scepticism (-); No / low perceived value-added (-)
	Personal preferences	Preference for face-to-face interactions (-)

6 Conclusion

In this paper, we have explored enablers and inhibitors of ESN adoption and usage from the perspective of executives at a technology company. Based on 12 semi-structured interviews, we determined 7 enablers and 20 inhibitors, which were aggregated into more abstract factors within the categories technological, organisational, social, and individual factors. While most identified enablers relate to positive experiences from actual platform use, ESN adoption and usage is strongly inhibited by technological and organisational factors.

With this paper, we contribute to the ESN literature by determining factors influencing ESN adoption and usage from the perspective of executives. In addition, our findings inform research on digital leadership and social software. In terms of managerial implications, the findings of this study highlight the importance of aligning strategies aiming at fostering ESN adoption and usage with an organisation's culture. Especially in hierarchical organisations, ESN implementation and management requires clearly defined goals, use cases, and upper management commitment to be successful.

The results of the study have to be seen in the light of its limitations. In terms of the sample of interviewees, men as well as executives in lower management levels (i.e. with a personnel responsibility of up to 25 employees), are overrepresented. Also, our findings are based on one case only and cannot be generalised. Yet, they may usefully inform ESN initiatives in companies in similar industries or companies with a comparable organisational structure and culture. To contribute to more generalised findings, we plan to conduct further case studies in other organisations. Also, in our future work, we will distinguish more precisely between factors influencing adoption as compared to factors influencing usage to understand better which factors should be considered during implementation and in the post-implementation stage.

References

1. Mäntymäki, M., Riemer, K.: Enterprise Social Networking: A Knowledge Management Perspective. Int. J. Inf. Manage. 36, 1042–1052 (2016)
2. Riemer, K., Stieglitz, S., Meske, C.: From Top to Bottom. Bus. Inf. Syst. Eng. 57, 197–212 (2015)
3. Dornis, A., Scheerer, B., Wagner, D., Vollmar, G., Kimmich, M., Peter, G., Dückert, S.: Wissensmanagement und Enterprise 2.0, http://www.gfwm.de/wp-content/uploads/2014/02/GfWM-Positionspapier_WM_und_E20.pdf (Accessed: 31.01.2017)
4. Viol, J., Hess, J.: Information Systems Research on Enterprise Social Networks – A State-of-the-Art Analysis. In: Nissen, V., Stelzer, D., Straßburger, S., and Fischer, D. (eds.) Multikonferenz Wirtschaftsinformatik (MKWI) 2016. pp. 351–362. Universitätsverlag Ilmenau, Ilmenau (2016)
5. Bala, H., Massey, A.P., Rajanayakam, J., Hsieh, C.J.: Challenges and Outcomes of Enterprise Social Media Implementation: Insights from Cummins, Inc. In: 48th Hawaii International Conference on System Sciences, pp. 1839–1848. IEEE Press, New York (2015)
6. Günther, O., Krasnova, H., Riehle, D., Schöndienst, V.: Modeling Microblogging Adoption in the Enterprise. In: Proceedings of the 15th Americas Conference on Information Systems, Paper 544. San Francisco, CA (2009)
7. Chelmis, C., Prasanna, V.K.: The role of organization hierarchy in technology adoption at the workplace. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 8–15. ACM Press, New York (2013)
8. Majumdar, A., Krishna, S., Bjorn, P.: Managers' Perceptions of Social Software Use in the Workplace: Identifying the Benefits of Social Software and Emerging Patterns of its Use. In: Proceedings of the 19th Americas Conference on Information Systems. Chicago, IL (2013)
9. Richter, A., Hetmank, C., Klier, J., Klier, M., Muller, M.: Enterprise Social Networks from a Manager's Perspective. In: 49th Hawaii International Conference on System Sciences, pp. 4242–4251. IEEE Press, New York (2016)
10. Chin, C.P.-Y., Evans, N., Choo, K.-K.R.: Exploring Factors Influencing the Use of Enterprise Social Networks in Multinational Professional Service Firms. J. Organ. Comput. Electron. Commer. 25, 289–315 (2015)
11. Leonardi, P.M., Huysman, M., Steinfield, C.: Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations. J. Comput. Commun. 19, 1–19 (2013)
12. Le Clair, C., Andrews, C., Schoeller, A., Sjoblom, S.: The Forrester Wave ™ : Enterprise Collaboration, Q4 2016, https://www.forrester.com/report/The+Forrester+Wave+Enterprise + Collaboration+Q4+2016/-/E-RES131184 (Accessed: 22.05.2017)

13. Buettner, R.: Analyzing the Problem of Employee Internal Social Network Site Avoidance: Are Users Resistant due to their Privacy Concerns? In: 48th Hawaii International Conference on System Sciences, pp. 1819–1828. IEEE Press, New York (2015)
14. Kügler, M., Lübbert, C., Smolnik, S.: Organizational Climate's Role in Enterprise Social Software Usage: An Empirical Assessment. In: Proceedings of the 12th International Conference on Wirtschaftsinformatik, Paper 55. Münster (2015)
15. Alarifi, A., Sedera, D., Recker, J.: Posters versus Lurkers: Improving Participation in Enterprise Social Networks through Promotional Messages. In: Proceedings of the 36th International Conference on Information Systems. Forth Worth, TX (2015)
16. Han, S., Sörås, S., Schjodt-Osmo, O.: Governance of an Enterprise Social Intranet Implementation: The Statkraft Case. In: Proceedings of the 25th European Conference on Information Systems, Paper 70. Münster (2015)
17. Li, C.: Why No One Uses the Corporate Social Network. Harvard Bus. Rev. Digit. Artic. 2–5 (2015)
18. Osch, W. van, Bulgurcu, B., Kane, G.J.: Classifying Enterprise Social Media Users: A Mixed-Method Study of Organizational Social Media Use. In: Proceedings of the 37th International Conference on Information Systems. Dublin (2016)
19. Fulk, J., Yuan, Y.C.: Location, Motivation, and Social Capitalization via Enterprise Social Networking. J. Comput. Commun. 19, 20–37 (2013)
20. Riemer, K., Scifleet, P.: Enterprise Social Networking in Knowledge-Intensive Work Practices: A Case Study in a Professional Service Firm. In: Proceedings of the 23rd Australasian Conference on Information Systems. Geelong (2012)
21. Kügler, M., Smolnik, S.: Uncovering the Phenomenon of Employees' Enterprise Social Software Use in the Post-acceptance Stage – Proposing a Use Typology. In: Proceedings of the 22nd European Conference on Information Systems. Tel Aviv (2014)
22. Richter, A., Stocker, A.: Exploration & Promotion: Einführungsstrategien von Corporate Social Software. In: Proceedings of the 10th International Conference on Wirtschaftsinformatik, pp. 1114–1123. Zurich (2011)
23. DGFP: Führen im digitalisierten Unternehmen. DGFP-Praxispapiere, 03/2016, (2016)
24. Lu, B., Guo, X., Luo, N., Chen, G.: Corporate Blogging and Job Performance: Effects of Work-related and Nonwork-related Participation. J. Manag. Inf. Syst. 32, 285–314 (2015)
25. Richter, A., Wagner, D.: Leadership 2.0: Engaging and Supporting Leaders in the Transition towards a Networked Organization. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 574–583. IEEE Press, New York (2014)
26. Van Dick, R., Gross, M., Helfritz, K.H., Holz, F., Stickling, E.: Digital Leadership: Die Zukunft der Führung in Unternehmen, http://www.clbo-frankfurt.org/media/2016/08/Digital-Leadership-Studie-NEW.pdf (Accessed: 19.09.2017)
27. Yin, R.K.: Case Study Research: Design and Methods. Sage Publications, Los Angeles, CA (2009)
28. Creswell, J.W.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications, Los Angeles, CA (2009)
29. Mayer, H.O.: Interview und schriftliche Befragung: Grundlagen und Methoden empirischer Sozialforschung. Oldenbourg, München (2013)
30. Mayring, P.: Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173 (Accessed: 22.09.2017)
31. Hacker, J., Bodendorf, F.: Conceptualising and Understanding User Behaviour in Enterprise Social Networks: A Qualitative Analysis. In: 28th Australasian Conference on Information Systems. Hobart (2017)

How Experience with Private IS Affects Employees' Satisfaction with Organizational IS

Uwe Ostermann[1]

[1]Goethe University, Business Informatics and Information Economics, Frankfurt, Germany
{ostermann}@wiwi.uni-frankfurt.de

Abstract. Most traditional IS adoption and acceptance research assumes that a majority of employees are resistant to the adoption of new technologies for work. However, phenomena like IT consumerization, bring your own device, and shadow IT illustrate that some employees actively introduce private IS to their organizations and thereby shape their own workplace. In this study, we investigate what leads employees to use alternative (private) IS for work. We draw on the cybernetic negative feedback loop as our theoretical framework to develop our research model. Our analysis indicates that satisfaction with organizational IS drives behavioral change. Further, our results show that not only the perceived performance of organizational IS predicts the employees' satisfaction with organizational IS but also the perceived performance of private substitutes.

Keywords: File sharing, satisfaction, intention to use private IT, post-adoption

1 Introduction

More and more digital natives [1] are entering the workforce [2]. This new generation of employees has grown up in a digital and connected world [3] and is used to being surrounded by ubiquitous information systems (IS) [4]. This extensive experience with technology gives employees the knowledge and ability to compare organizational IS with alternatives/substitutes. Phenomena like consumerization, bring your own device (BYOD), or shadow IT [5, 6] describe that various employees have the desire to use private IS instead of business/organizational IS for work. On the one hand, these phenomena challenge an underlying assumption of traditional adoption and acceptance research. Traditionally, most IS adoption and acceptance research assumes that the majority of employees/users are resistant and inert towards the adoption of new technologies for work [2, 7, 8]. However, we see that some employees actively shape their workspace instead of being hesitant or resistant towards new technology. On the other hand, this behavior challenges the established top-down approach of technology diffusion in organizations and affects organizational structures to a large extent [9, 10]. Until recently, the IT department decided to purchase new technology and thereby introduced organizational change. However, currently employees are taking the initiative and are introducing IS into the organization, while the organization itself has

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

to react [5, 9]. With our study, we want to identify the underlying mechanism that leads employees to actively introduce new (private) technology into their organizations. Therefore, our research questions are:

RQ1: What leads employees to introduce new (private) IS into their workspace?

RQ2: How does knowledge and experience with private IS affect the attitude towards organizational IS?

Drawing on the cybernetic negative feedback loop [11] as our theoretical framework, we argue that employees are getting more critical and demanding towards their organizational information systems, since they use new and innovative technology in their private lives [6]. To empirically validate this relationship, we develop a research model linking *perceived performance of organizational* and *private IS* with *satisfaction with organizational IS*, and the *intention to use private IS for work*. We use satisfaction and intention as dependent variables of our research model as post-adoption research has shown that satisfaction is a good predictor for employees' (dis)continuance intention [12, 13]. To test our hypotheses, we conducted an online survey with 154 employees from different organizations. As a setting for our study, we chose document and file sharing, as we think that there are relatively mature consumer systems, which can substitute conservative and traditional ways of file sharing in organizations. Especially cloud-based services like Dropbox or Google Drive are widely spread and accepted among private users [14, 15]. To analyze our data, we use covariance-based structural equation modelling (CB-SEM). We contribute to business informatics and information systems by demonstrating that the familiarity with private alternative IS alters personal standards and decreases satisfaction with organizational IS.

The remainder of our paper is structured as follows. In Section 2, we introduce the related work and our theoretical framework. In Section 3, we develop our hypotheses and research model. In Section 4, we describe our research methodology. In Section 5, we present the data analysis and results of our evaluation. In Section 6, we conclude with a discussion, give implications, and outline limitations of our work.

2 Related work

2.1 Private IS in Organizations

Over the last years, phenomena describing employees that use private IS for work, gained much attention in IS research. These phenomena are called consumerization, BYOD, and shadow IT. The term consumerization was first mentioned by Moschella et al. (2004) and since then has been assigned several definitions. In our study, we adopt the characterization of Harris et al. (2012) who describe consumerization as "[…] the adoption of consumer devices and applications in the workforce […]" [5, p.99]. Bring your own system strategies, on the other hand, are a way for organizations to formally approve and regulate consumerization behavior [6, 17, 18]. Non-approved usage of (private) IT in organizations has been termed shadow IT [6, 19, 20]. Whether allowed or not, the usage of private IS challenges the established top-down approach of introducing IT to an organization [9, 10]. Traditionally, IT systems were chosen and implemented by the organization, whereas users/employees were expected to merely adopt these systems [2, 9, 21]. However, with employees using private IS for work the direction of the diffusion of IT systems into the organization shifts from a top-down to a bottom-up approach [9, 22–24]. Passively adopting employees turn into employees taking action and challenging their IT department [10, 25]. We see an increasing dissatisfaction with organizational IT as the major cause for this evolution [5, 6, 26] and as a precondition for behavioral change. To assess the emergence of dissatisfaction, we use the cybernetic negative feedback loop [11] as our theoretical framework and describe it in the following section**.**

2.2 Cybernetic Negative Feedback Loop

The cybernetic negative feedback loop [11] is applied as a high-level framework to establish the variables for our research model (Figure 1). The negative feedback loop has already been applied in IS research, for instance by Burton-Jones and Grange (2013) and Liang and Xue (2009), and enables us to explain behavioral change of employees. It consists of an input function (perception), a goal/standard/reference value (from now on referred to as reference value), a comparator, and an output function (action) [29]. According to Ashby (1956), difference is the fundamental concept in cybernetics. In the negative feedback loop, the comparator detects this difference. The comparator compares a sensed value (perception) against a reference value. In case of a difference between perception and reference value actions are carried out to reduce this difference. Usually the feedback loop is triggered by a disturbance in the environment which alters the perception.

Applying the negative feedback loop to the phenomenon that employees introduce private IS into their organizations, we assume that the emergence of innovative IS in private life corresponds with a disturbance affecting the environment of an employee. We theorize that this experience with superior alternative IS in private life increases employees' reference value (dashed line in Figure 1). This new experience and knowledge creates needs [31] and shifts the reference value to which employees compare their organizational IS to [6]. Assuming a constant perception of organizational IS an upshift of the reference value creates a discrepancy and triggers the feedback loop. Discrepancy in turn affects the satisfaction of employees [32, 33]. To reduce this discrepancy and dissatisfaction employees become active and change their behavior until their perception and standard are aligned again. Hence, discrepancy and satisfaction are preconditions for action, employees can change their behavior for instance by demanding new technology at work or using their private IS for work tasks even if the organization does not allow this practice. This employee driven change in turn affects the environment. As employees and their organization are mutually interdependent, a change in employees' behavior can have an impact on the organization and its structure [29, 34]. In our study, we do not empirically investigate the impact of behavioral change on the organization, but at this point, we want to emphasize the relevance.

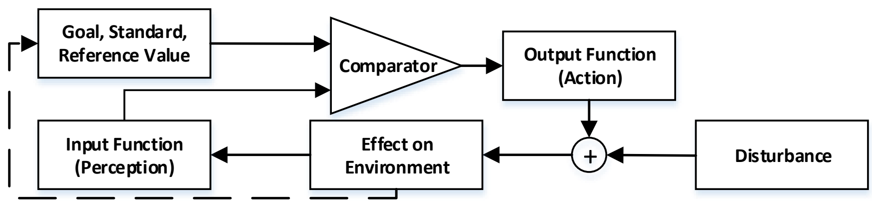


Figure 1. Negative feedback loop (adapted from Carver and Scheier (1998))

3 Hypothesis Development

3.1 Satisfaction

Oliver (1980) defines satisfaction as a function of expectation and disconfirmation. Accordingly, disconfirmation results from the comparison of expectations with actual experiences of a product [32]. In the IS context, Bhattacherjee (2001) uses this conceptualization of satisfaction to explain IS continuance with the IS continuance model [13]. Both assume that a consumer/user builds up expectations before the actual usage of a product/system. Subsequently, he/she compares his/her actual experience to prior expectations. When a consumer senses a discrepancy between expectation and experience he/she feels dissatisfied [12, 13, 32]. We translate this relation to the negative feedback loop and assume that a sensed discrepancy between a personal standard and actual experience leads to dissatisfaction which in turn leads to an action/change in behavior [29]. Existing research has shown that satisfaction is a good predictor for (dis)continuance, migration, and loyalty [12–14, 35–37]. In our research model, we operationalize action/behavioral change by measuring the intention to use private IS for work. Consequently, we argue that low satisfaction with organizational IS increases the intention to use private IS for work.

H1: Satisfaction with organizational IS will have a negative effect on the intention to use private IS for work.

3.2 Performance of Organizational IS

Following the logic of the cybernetic negative feedback loop, a perception is compared to a standard, reference value or goal. In our research model, perceived performance is operationalized by the established construct perceived usefulness (PU) of the technology acceptance model (TAM) [38]. In our research setting, this perception is a post-adoption belief in a mandatory setting based on actual experience. With employees perceiving their organizational IS as useful, a negative discrepancy between private and organizational IS is less likely. We further argue that employees will be dissatisfied if they have to use IS, which they believe does not enhance their work performance. In accordance with Bhattacherjee (2001), we propose a positive relationship between PU of organizational IS and satisfaction with organizational IS.

H2: PU of organizational IS will have a positive effect on satisfaction with organizational IS.

Moreover, existing post-adoption research shows a positive relationship between PU of the incumbent system and continuous intention [8, 12, 13]. With an increasing intention to continue organizational IS use employees' intention to use private IS for work will decrease. Thus, we hypothesize:

H3: PU of organizational IS will have a negative effect on the intention to use private IS for work

3.3 Performance of Private IS

Taking Oliver's (1980) definition, satisfaction is a function of expectation and confirmation. In our model, we replace confirmation with the belief about performance of private IS. Concerning the negative feedback loop, we theorize that this belief represents the reference value. Assuming the perceived performance of organizational IS remains unchanged and perceived performance of private IS increases, a negative discrepancy should occur. This in turn leads to a decrease of satisfaction with organizational IS. Consistent with the performance of organizational IS, we operationalize private IS using PU.

H4: PU of private IS will have a negative effect on satisfaction with organizational IS.

According to TAM one of the main antecedents of intention to use a technology is PU. Consistently, consumerization and BYOD literature show that beliefs about the performance of private IT in the work environment positively affect employees' intention to use private IS for work [39–42]. Therefore, we incorporate this relationship into our model and hypothesize:

H5: PU of private IS will have a positive effect on the intention to use private IS for work.

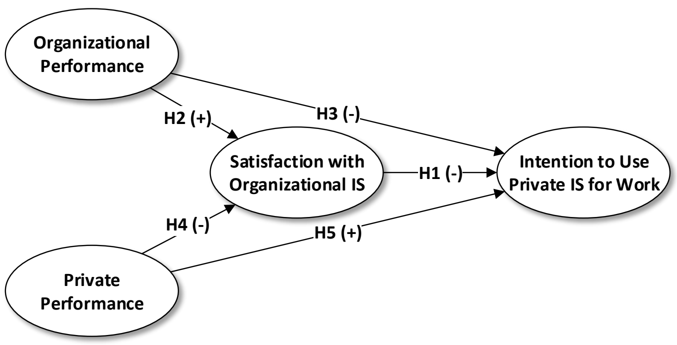


Figure 2. Research model

4 Research Methodology

4.1 Data Collection

To empirically test our hypotheses and capture the individual perceptions of our participants, we conducted a cross sectional online-survey. We distributed the survey during the third quarter of 2016 and received 154 responses. We dropped all responses of participants that do not have a job in which they have to share files with their colleagues/teammates. We further excluded responses of participants that do not use systems to share files in their private lives. Thereby, we ensured that a comparison of both systems is possible for each respondent. Table 1 displays the sociodemographic data of the 120 remaining respondents including gender, education, age, and whether participants are forbidden to use private IT for work.

Table 1. Sample characteristics

N = 120	*Count*	*%*		*Count*	*%*
Gender			**Private IT**		
Female	42	35.00	Forbidden	60	50.00
Male	78	65.00	Not forbidden	60	50.00
Education					
University Degree	98	81.67		**Mean**	**Std. Dev.**
No University Degree	22	18.33	**Age**	28.06	6.52

4.2 Study Design and Instrument Development

We selected document and file sharing as a setting for our survey, as we are convinced that document and file sharing is a common task for both work and private life with relatively mature and comparable systems. The online survey consisted of three parts. In the first part, we asked participants if they have to share files at work and which system their employer provides (e.g., SharePoint, fileserver, email or organizational cloud solution). Subsequently, we asked the participants if they use a system for document/file sharing in private life and which system they preferred (e.g., Dropbox, iCloud, Google Drive). Both systems for private and work life were automatically inserted into the PU, and intention items. In the second part, participants had to assess their perception towards the performance of their organizational IS and private IS, their satisfaction with their organizational IS, their intention to use their private IS for work. PU and intention to use were rated on a seven-point Likert scale using the established constructs of Davis (1989) and Venkatesh et al. (2003). To measure satisfaction with organizational IS, we used the four-point semantic differential also used by Bhattacherjee (2001). In the third part, we asked the participants for socio-demographic characteristics and included them as control variables into the estimation of our research model.

5 Data Analysis and Results

5.1 Measurement Model

To test for reliability of our measurement model we conducted a confirmatory factor analysis and calculated Cronbach's alpha (Cα) and Composite Reliability (CR) of our research constructs. Cα and CR both should exceed a threshold of 0.7 to indicate good reliability [44–46]. To check for convergent validity, we measured the Average Variance Extracted (AVE) of each construct. The AVE requires values of 0.5 or higher to indicate sufficient convergent validity [44, 47]. We further tested our constructs on discriminant validity using the Fornell-Larker-Criterion. Accordingly, discriminant validity of a construct is ensured if the square root of the AVE exceeds the correlations with any other construct [44, 47]. Our results, shown in Table 2, indicate that the latent constructs used in our measurement model have sufficient reliability, convergent validity, and discriminant validity.

Table 2. Evaluation of latent constructs

	Cα	*CR*	*AVE*	**OPU**	**PPU**	**Sat**	**Int**
OPU	.914	.916	.786	**.886**			
PPU	.944	.944	.848	.052	**.921**		
Sat	.890	.891	.731	.467	-.330	**.855**	
Int	.972	.972	.921	-.181	.632	-.514	**.959**

OPU = PU of organizational IS; PPU = PU of private IS; Sat = Satisfaction with organizational IS; Int = Intention to use private system for work; The diagonal represents the squared AVE values. Off diagonal elements are the correlations among latent constructs.

5.2 Common Method Bias

As all responses of our survey are self-reported, we conducted Harman's single factor test [48, 49] to check for common method bias. We performed a factor analysis with all items of the four latent variables using principal component analysis and no rotation. Based on the Kaiser criterion (Eigenvalues > 1), three factors were extracted. The first factor accounted for about 45.86 % of the total variance. Thus, as no single factor could be derived and no general factor explains the majority of variance in the variables, the chances for common method bias are unlikely [48, 49].

5.3 Structural Model and Hypothesis Testing

After we confirmed the factor structure in our dataset in the CFA, we performed a CB-SEM. The absolute fit indices of our research model indicating an overall excellent model fit (*Chi²/DF*: 1.231; *CFI*: 0.986; *SRMR*: 0.061; *RMSEA*: 0.044; *PClose*: 0.618) [50].

The results of our model show a significant relationship between satisfaction with organizational IS and the intention to use private IS instead (-0.291; p: 0.003). Thereby, confirming H1. Further, our data confirms H2, hence a positive relationship between PU of organizational IS and satisfaction with organizational IS (0.487; p: 0.000). Our model could not verify a direct negative effect of PU of organizational IS on the intention to use private IS for work (-0.070; p: 0.387). Consequently, we are not able to confirm H3. However, our model reveals a significant negative relationship between PU of private IS and the satisfaction with organizational IS (-0.353; p: 0.000), thereby confirming hypothesis H4. Moreover, our results confirm a significant positive relationship between PU of private IS on the intention to use private IS for work, IS (0.536; p: 0.000), confirming H5.

In our analysis, we controlled for gender, age, education, income, and a dummy variable indicating whether the participants are not allowed to use private IS for work. None of the control variables have a significant effect on the satisfaction with organizational IS or the intention to use private IS for work. Our research model explains 35.5% of the variance of satisfaction with organizational IT, and 51.3% of the variance of intention to use private IS for work with control variables excluded. The R^2 values for the model including control variables are 41.5% for satisfaction with organizational IT, and 51.6% for intention to use private IS for work.

Table 3. Results of hypothesis testing

Hypothesis	*SC*	*P-value*	*Result*
H1	- .291**	.003	Supported
H2	.487***	.000	Supported
H3	- .070	.387	Not supported
H4	- .353***	.000	Supported
H5	.536***	.000	Supported
Indirect effects	*SC*	*P-value*	
OrgPU → Sat → Int	- .168*	.012	
PrivPU → Sat → Int	- .110**	.006	

SC: Standardized Coefficients; * Significant at a .05 level;
** Significant at a .01 level; *** Significant at a .001 level

In addition, we tested for indirect effects following the approach of Zhao et al. (2010). Applying bootstrapping with 2000 samples our data reveals a significant indirect effect of perceived performance of organizational IS on the intention to use private IS mediated by satisfaction. Since there is no direct effect of perceived business performance on intention, we consider this an indirect only mediation.

Second, our analysis supports an indirect effect of PU of private IS on intention mediated by satisfaction. As the product of the three path coefficients is positive, we consider this a complementary mediation (see Table 3).

6 Discussion

The objectives of this study were to examine how knowledge and experience with alternative private IS affects the attitude towards organizational IS and to uncover the mechanism that drives employees to introduce new (private) IS to their organizations.

We applied the cybernetic negative feedback loop as our theoretical framework to develop our research model. To answer RQ1 we proposed satisfaction with organizational IS as antecedent for behavioural change and the intention to use private IS for work. To answer RQ2 we proposed PU of both personal and organizational IS as antecedent for satisfaction with organizational IS.

First, our results support our assumption that satisfaction is one major driver of behavioral change. Since qualitative consumerization literature already proposed a relationship between satisfaction with organizational IS and the intention to use private IS for work [6, 26], we quantitatively verified this relationship. Second, we find that satisfaction mediates the effects of PU of private and organizational IS on the intention to use private IS for work. Since the relationship between PU of private IS on the intention to use private IS for work may not be surprising [23, 39–42], the mediation analysis also reveals that low degrees of PU of organizational IS indirectly increases the intention of employees to use a private alternative for work. Third, our results show that not only does inferior performance of organizational IS lead to decreased satisfaction with organizational IS, *but positive experience with alternative private IS does too*. This finding is further in line with Rogers (1962), who describes that an individual gets into an “uncomfortable state of mind” when he/she is willing to adopt an innovation but is not able to [52].

Imagine an employee who has privately switched from innovation I to innovation II (Figure 4) at a certain point in time, but is required to use innovation I at work. Hence, this mandatory use at work will make him/her feel uncomfortable respectively dissatisfied because he/she is willing to adopt innovation II. The longer it takes an organization to adopt innovation II, the more employees will migrate from innovation I to II in their private lives, leading to an increasing quantity of dissatisfied employees.

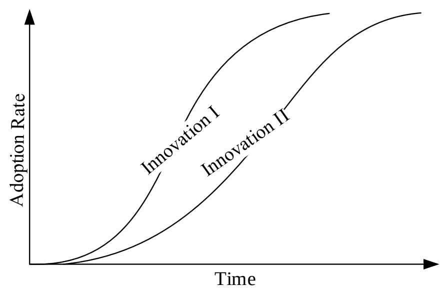


Figure 3. S-curves (adapted from Rogers (1995)).

6.1 Practical Implications

Our results imply that it is important for organizations to monitor developments and trends outside the organization in the consumer market. As employees get to know and used to superior IS in their private lives, their standards for IS will increase in general. Consequently, satisfaction with organizational IS will decrease if the organization cannot keep up with the speed of innovation outside the organization. This may happen, even if employees were satisfied with the work IS at the time of initial adoption. As satisfaction is a precondition for change in behavior, possible consequences are employees complaining or using alternative IS even without approval (shadow IT) [53]. In the setting of file sharing in an organization, the dissatisfaction with the traditional way of file sharing might lead to employees using private cloud solutions instead and thereby threatening the organizations data security. Thus, organizations must be sensitive concerning innovations in the consumer market and the variation of employees' satisfaction. One way to address these issues is to allow the use of individual and private IS as long as it does not harm the organization. Another feasible way is to mimic consumer systems to provide employees with the standard they are accustomed to from their private lives.

6.2 Limitations and Future Research

This study is subject to several limitations. First, participants of our study were faced with a hypothetical scenario while reporting their intention to use a private IS for work. The actual situation in participants' work life might influence the reported intention. However, our study shows that the effects of perceived performance of private and organizational IS remain the same as we controlled for the actual situation at work using a control variable. Second, our study focusses on a specific kind of system – a system for file sharing. Although, we think that this system is very suitable, as there are comparable solutions in both work and private life, future research could extend this study by investigating different systems or by differentiating for instance between physical and non-physical IS. Furthermore, only employees who had experience with an alternative IS in their private lives participated in our study. Drawing on the cybernetic negative feedback loop, we propose that knowledge and familiarity with a superior system shifts the reference value of users/employees. This implies that employees who do not know alternative IS would be more satisfied with their organizational IS. Subsequent studies could further investigate this causal relationship.

References

1. Prensky, M.: Digital Natives, Digital Immigrants Part 1. Horiz. 9, 1–6 (2001).
2. Vodanovich, S., Sundaram, D., Myers, M.: Digital Natives and Ubiquitous Information Systems. Inf. Syst. Res. 21, 711–723 (2010).
3. Palfrey, J., Gasser, U., Simun, M., Barnes, R.F.: Youth, Creativity, and Copyright in the Digital Age. Int. J. Learn. Media. 1, 79–97 (2009).
4. Yoo, Y.: Computing in Everyday Life: A Call for Research on Experiential Computing. MIS Q. 34, 213–231 (2010).
5. Harris, J., Ives, B., Junglas, I.: IT Consumerization: When Gadgets Turn Into Enterprise IT Tools. MIS Q. Exec. 11, 99–112 (2012).
6. Köffer, S., Ortbach, K.C., Niehaves, B.: Exploring the Relationship between IT Consumerization and Job Performance. Commun. Assoc. Inf. Syst. 35, 261–284 (2014).
7. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User Acceptance of Information Technology: Toward A Unified View. MIS Q. 27, 425–478 (2003).
8. Polites, G.L., Karahanna, E.: Shackled to the Status Quo: The Inhibiting Effects of

Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. MIS Q. 36, 21–42 (2012).
9. Leclercq-Vandelannoitte, A.: Managing BYOD: how do organizations incorporate user-driven IT innovations? Inf. Technol. People. 28, 2–33 (2015).
10. Junglas, I., Way, A., Ives, B., Hall, M., Harris, J.G.: Consumer IT at Work: Development and Test of an IT Empowerment Model. Proc. Int. Conf. Inf. Syst. (2014).
11. Wiener, N.: Cybernetics: Control and Communication in the Animal and the Machine. MIT Press, Cambridge, MA (1948).
12. Bhattacherjee, A., Premkumar, G.: Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. MIS Q. 28, 229–254 (2004).
13. Bhattacherjee, A.: Understanding Information Systems Continuance An Expectation-Confirmation Model. MIS Q. 25, 351–370 (2001).
14. Bhattacherjee, A., Park, S.C.: Why end-users move to the cloud: a migration-theoretic analysis. Eur. J. Inf. Syst. 23, 357–372 (2014).
15. Legner, C., Eymann, T., Hess, T., Matt, C., Böhmann, T., Drews, P., Mädche, A., Urbach, N., Ahlemann, F.: Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. Bus. Inf. Syst. Eng. 59, 301–308 (2017).
16. Moschella, D., Neal, D., Opperman, P., Taylor, J.: The "Consumerization" of Information Technology Position Paper. CSC, El Segundo, California,USA. (2004).
17. Baskerville, R., Lee, A.: Individual - Organizational bindpoints: A Design Theory for Bring-Your-Own-System. Proc. - Pacific Asia Conf. Inf. Syst. PACIS 2013. (2013).
18. French, A., Guo, C., Shim, J.P.: Current Status, Issues and future of bring your own device (BYOD). Commun. Assoc. Inf. Syst. 35, 10 (2014).
19. Györy, A., Cleven, A., Uebernickel, F., Brenner, W.: Exploring the Shadows : It Governance Approaches To User- Driven Innovation. ECIS 2012 Proc. (2012).
20. Haag, S., Eckhardt, A.: Justifying Shadow IT Usage. PACIS 2015 Proc. (2015).
21. Nan, N.: Capturing Bottom-up Information Technology Use Processes: A Complex Adaptive Systems Model. MIS Q. 35, 505–532 (2011).
22. Ortbach, K., Walter, N., Ayten, Ö.: Are You Ready to Lose Control? A Theory on the Role of Trust and Risk Perception on Bring-Your-Own-Device Policy and Information System Service Quality. ECIS 2015 Res. Pap. (2015).
23. Ortbach, K., Bode, M., Niehaves, B.: What Influences Technological Individualization? – An Analysis of Antecedents to IT Consumerization Behavior. AMCIS 2013 Proc. (2013).
24. Weiß, F., Leimeister, J.M.: Consumerization: IT innovations from the consumer market as a challenge for corporate IT. Bus. Inf. Syst. Eng. 4, 363–366 (2012).
25. Koch, H., Zhang, S., Giddens, L., Milic, N., Yan, K., Curry, P.: Consumerization and IT Department Conflict. Proc. Int. Conf. Inf. Syst. 1–15 (2014).
26. Ostermann, U., Wiewiorra, L.: Bring it On(e)! Personal Preferences and Traits as Influencing Factors to Participate in BYOD Programs. Proc. Eur. Conf. Inf. Syst. (2016).
27. Burton-Jones, A., Grange, C.: From Use to Effective Use: A Representation Theory Perspective. Inf. Syst. Res. 24, 632–658 (2013).
28. Liang, H., Xue, Y.: Avoidance of Information Technology Threats: A Theoretical Perspective. MIS Q. 33, 71–90 (2009).
29. Carver, C.S., Scheier, M.F.: On the Self-Regulation of Behaviour. Cambridge University Press, Cambridge, UK (1998).
30. Ashby, W.R.: An Introduction to Cybernetics. (1956).
31. Rogers, E.M.: Diffusion of innovations. Newyork Free Press, New York, NY (1995).
32. Oliver, R.L.: A cognitive model of the antecedents and consequences of satisfaction

decisions. J. Mark. Res. 17, 460–470 (1980).
33. Festinger, L.: A theory of cognitive dissonance. (1957).
34. Orlikowski, W.J.: The Duality of Technology : Rethinking the Concept of Technology in Organizations. Organ. Sci. 3, 398–427 (1992).
35. Kim, S.S., Son, J.-Y.: Out of Dedication or Constraint? A Dual Model of Post-Adoption Phenomena and Its Empirical Test in the Context of Online Services. MIS Q. 33, 49–70 (2009).
36. Fang, Y., Quershi, I., Sun, H., McCole, P., Ramsey, E., Lim, K.H.: Trust, Satisfaction, and Online Repurchase Intention: The Moderating Role of Perceived Effectiveness of E-Commerce Institutional Mechanisms. MIS Q. 38, 407–427 (2014).
37. Ray, S., Kim, S.S., Morris, J.G.: Online users' switching costs: Their nature and formation. Inf. Syst. Res. 23, 197–213 (2012).
38. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User Acceptance of Computer Technology: a Comparison of Two Theoretical Models. Manage. Sci. 35, 982–1003 (1989).
39. Loose, M., Weeger, A., Gewald, H.: BYOD–The Next Big Thing in Recruiting? Examining the Determinants of BYOD Service Adoption Behavior from the Perspective of Future Employees. AMCIS 2013 Proc. (2013).
40. Weeger, A., Wang, X., Gewald, H.: It Consumerization: Byod-Program Acceptance and Its Impact on Employer Attractiveness. J. Comput. Inf. Syst. 56, 1–10 (2015).
41. Hopkins, N., Sylvester, A., Tate, M.: Motivations For BYOD: An Investigation Of The Contents Of A 21st Century School Bag. Proc. 21st Eur. Conf. Inf. Syst. 1–12 (2013).
42. Ortbach, K.: Unraveling the Effect of Personal Innovativeness on Bring-Your-Own-Device (BYOD) Intention – The Role of Perceptions Towards Enterprise-Provided and Privately Owned Technologies. ECIS 2015 Complet. Res. Pap. (2015).
43. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of. MIS Q. 13, 319–340. (1989).
44. Hair, J.F., Ringle, C.M., Sarstedt, M.: PLS-SEM: Indeed a Silver Bullet. J. Mark. Theory Pract. 19, 139–152 (2011).
45. Peterson, R.A.: A Meta-analysis of Cronbach´s Coefficient Alpha. J. Consum. Res. 21, 381–191 (1994).
46. Bagozzi, R.P., Yi, Y.: On the Evaluation of Structural Equation Models. J. Acad. Mark. Sci. 16, 74–94 (1988).
47. Fornell, C., Larcker, D.F.: Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. J. Mark. Res. 18, 39–50 (1981).
48. Malhotra, N.K., Kim, S.S., Patil, A.: Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research. Manage. Sci. 52, 1865–1883 (2006).
49. Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P.: Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88, 879–903 (2003).
50. Hu, L., Bentler, P.M.: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct. Equ. Model. A Multidiscip. J. 6, 1–55 (1999).
51. Zhao, X., Lynch Jr., J.G., Chen, Q.: Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. J. Consum. Res. 37, 197–206 (2010).
52. Rogers, E.M.: Diffusion of innovations. Free Press of Glencoe, New York (1962).
53. Silic, M., Back, A.: Shadow IT - A view from behind the curtain. Comput. Secur. 45, 274–283 (2014).

DO ONLINE COMMUNITIES BENEFIT FROM APPOINTING VOLUNTEER MODERATORS? EVIDENCE FROM A REGRESSION DISCONTINUITY DESIGN

Jens Foerderer and Armin Heinzl[1]

[1] Business School, University of Mannheim, Germany.
{foerderer,heinzl}@uni-mannheim.de

Abstract. We study whether online communities can benefit from appointing users to "volunteer moderators", in terms of community members or leaders who are supposed to encourage participation and plant the seeds of community. We exploit a quasi-experiment on Stack Exchange, a network of more than 160 online communities, over the period from 2010 to 2017. These communities regularly hold democratic elections of moderators. By focusing on elections decided by a narrow vote margin, we can exploit a regression discontinuity that yields a quasi-random assignment of moderatorship. We find that online communities can significantly benefit from volunteer moderatorship: closely elected moderators contribute not only more content, they also provide more extensive content to the community. Furthermore, the direct outcomes of moderatorship appear to have positive spillover effects on community discourse: moderators spark more extensive discussions, which are rated higher and bookmarked more often.

Keywords: Online communities, community governance, moderators, quasi-experiment, regression discontinuity

1 Introduction

The "1% rule" describes a fundamental problem of user participation in online communities [1]: only 1% of an online community's users account for most contributions ("key users"), 9% contribute from time to time ("commenters"), and the majority of 90% of users free-ride ("lurkers"). For example, of the more than 20 million registered users on Wikipedia, less than .05% of them are estimated to create and edit the majority of articles. Similar estimations have been documented for participation on the question-and-answer site Stack Exchange, YouTube, and SAP's developer forums. The "1% rule" illustrates the fundamental challenge of online community governance [e.g., 2–5]: what governance instruments are effective in stimulating user contributions? While scholars have identified various economic, psychological, and sociological instruments that may engender contributions of the "lurkers" and "commentators" [e.g., 6–8], few studies systematically assess instruments that allow community owners to keep the 1%—in terms of the "core that contributes content and protects the boundary of the community" [9]—engaged. This issue is particularly critical given that scholars found that many online communities are sustained by these "1%" of members [10, 11] and how easily communities can tip toward failure if key users turn over.

One governance instrument that has found widespread but controversial resonance concerns promoting key users to volunteer "moderators", "community leaders", or "advocates". Moderators shall "facilitate discussion on a voluntary basis […] to engender trust, encourage participation, and plant the seeds of community" [8]. By appointing moderators, community owners hope to boost key users to participate even more actively in a community, while having positive spillover effects on other users and reducing a community owner's costs of managing a community [8, 12]. To achieve this goal, community owners grant moderators extensive control to create and enforce policies, promote moderatorship as the highest social status within a community, and manage relations to them. However, the effectiveness of moderatorship as a governance instrument is controversial. Instead of promoting contributive behavior, much anecdotal evidence suggests moderatorship to discourage contributions [e.g., 13, 14]. For example, Other anecdotes described that moderators shifted their efforts on policing the contributions of other users rather than contributing. Other examples describe how moderatorship triggered even deceptive and collusive behaviors of key users such as Wikipedia moderators who revised articles in the favor of third-parties or the various communities in the Reddit network that have become notorious for their moderators' misbehavior [12, 14]. Of course, these examples may reflect isolated instances of the "wrong people" being appointed moderators. Nevertheless, an evaluation of the consequences of moderatorship is so far missing.

In this paper, we address two important yet unresolved questions about the implications of appointing volunteer moderators. The first question we study is whether moderatorship indeed discourages key users to contribute to a community. The debate on this question is characterized by two opposing theoretical perspectives that suggest governance instruments like moderatorship as a catalyst or inhibitor of contributive user behavior, respectively. The inhibitor argument—put forward by proponents of negative effects of devolving control to users and awarding community members with social status—is that moderatorship has no or even a detrimental effect on a user's contributions [15]. Quite the opposite, once moderatorship is granted to users, they lose the incentives to contribute [15]. Following this logic, moderatorship is merely an effective instrument and may withdraw key users from the discourse, which is a reason why community owners should avoid its widespread adoption. By contrast, the catalyst argument says that moderatorship might motivate greater investments in the community by instrumental, sociologically-based, and psychologically-motivated mechanisms. Instrumentally, moderatorship provides a user with greater control from which a user may draw to contribute [15, 16]. Sociologically, moderatorship is a status that may motivate greater contributions to the community [e.g., 6]. Psychologically, moderatorship may trigger greater attachment to and identification with a community [e.g., 17]. If moderatorship indeed catalyzes key users to contribute more by one of these mechanisms, then community owners may substantially benefit from implementing volunteer moderators. By establishing the causal effects of moderatorship and untangling its underlying mechanisms we seek to shed light on this debate. The second question we study is whether volunteer moderator behavior has spillover effects on community discourse, in terms of leading to more extensive interactions and contributions that are also perceived as more valuable by community members. Despite the prevalence of this question, previous studies provide little insight on this question.

The key empirical problem here is that moderatorship is usually not randomly assigned but results from a two-stage selection process. Users first strive for moderatorship based on unobserved characteristics such as motivation or information, and then community owners screen and appoint candidates based on another set of unobserved characteristics such as skills or personal impression [12]. Because of this two-stage selection, naïve comparisons of regular users and moderators are strongly biased. Unless differences in omitted variables can plausibly be removed, attributing user behavior to moderatorship is not possible. To overcome the selection problem, we exploit a regression discontinuity design [18, 19] in the context of Stack Exchange, a network of online communities that comprises more than 30 million posts and 7 million users. Many of these communities decide on their moderators by holding secret ballot elections, which are remarkably similar to real-world democratic election processes. Comparing moderator candidates who closely won an election with candidates who closely lost provides us with a quasi-random assignment of moderatorship [19, 20]. Consistent with our identification assumption, we find that candidates close to the cutoff behave similarly in the period prior to the election. Our dataset combines ballot data of all elections in Stack Exchange communities with user-level data on 1,012 moderatorship candidates across more than 70 communities over the period from 2010 to 2017.

2 Theoretical Background

Online communities have become an important aspect of digital environments [21–23]. We refer to online communities as “virtual space[s] where people come together with others to converse, exchange information or other resources, learn, play, or just be with each other” [3]. Communities emerged as an important tool of firms to benefit from the ideas of outsiders [2, 24], foster brand building [25], or provide user-to-user support [26, 27]. Communities are also relevant outside the business world. Various interest groups and associations run online communities for fostering exchanges among like-minded people [5]. Communities are also the backbone of the open source software movement as well as of the various user-generated content sites like Wikipedia or Stack Exchange [28]. Regardless of their purpose, all communities have in common that they draw on the participation and contribution of their members.

The governance of online communities has proven to be a complex undertaking due to the inherent public goods problem: users can consume the contributions of others without having to contribute themselves [3, 7, 29, 30]. For example, Wikipedia users can consume the articles contributed by others without having to contribute back. Similarly, users of Stack Overflow can receive programming help without having to give advice themselves. If the public goods problem is not addressed by proper governance mechanisms, communities will suffer from an undersupply of content due to insufficient incentives for users to contribute, and eventually fail [29]. Traditional economic models are less informing about the design of such mechanisms as they mostly turn toward pricing schemes, which cannot directly be applied in online communities as online communities rely on voluntary participation and contribution of time and effort rather than the exchange of money [31, 32]. Thus, the key challenge to community owners is the design of mechanisms that create incentives for users to contribute and that motivate core participants to sustain and improve their contribution [5, 22, 23, 32, 33].

Volunteer moderators represent a widespread instrument of online community governance [12, 34]. There exists no coherent definition of the role of a moderator but for the purpose of this study, we refer to moderators as "members who facilitate discussion on a voluntary basis […] to engender trust, encourage participation, and plant the seeds of community" [8]. While the tasks of a moderator differ from community to community, they usually encompass the posting of new expert content, supporting the flow of discussions, and keeping a high quality of content contributions going, if necessary, by revising and deleting of content, the warning, suspending, and banning of users [12].

Our study of volunteer moderators relates to a broader discussion on three theoretical mechanisms that are suggested to be key to influence user participation, namely devolving control to users [8, 33], sociologically-driven status incentives [7, 13, 35], and psychologically-based community identification [17, 36]. First, moderatorship, in terms of the position or "office" of a moderator, grants a user substantial control, in terms of decision-making rights, to create and enforce community policies. As such, moderatorship relates to a broader array of literature that studies how online collectives govern, organize, and coordinate the actions of individuals to achieve collective outcomes [11, 33, 37]. Second, moderatorship represents a social status. Social status refers to an actor's standing in a group when standing is based on prestige, honor, or deference [38]. Work in this stream is focused on understanding whether community owners can foster user contributions by awarding user contributions with trivial symbols of status and graduating contributors in the status hierarchy [e.g., 6, 7, 35, 39]. For example, in the Stack Exchange communities, contributors can earn symbolic bronze, silver, or gold badges along with fictitious titles like "Copy Editor", "Explainer", or "Teacher". Finally, moderatorship may stimulate contributive behavior by triggering greater identification, affiliation, and attachment to a community in terms of a psychological motive. As such, moderatorship relates to work that outlined psychological identification with a community as a driver of contributive behavior [17, 36, 40].

3 Empirical Strategy and Data

Our empirical context is Stack Exchange, a network of online communities in a question-and-answer format. Communities cover various interests including computer programming, personal finance, or graphic design. According to Stack Exchange statements, moderators are expected to be "patient and fair, led by example, show respect for their fellow community members in their actions and words, are open to some light but firm moderation to keep the community on track and resolve (hopefully) uncommon disputes and exceptions". Moderators are also expected to contribute their expert knowledge to the community. On Stack Exchange, moderatorship grants users both status and control. Moderators can exert control by revising or deleting content as well as the warning or banning of users. Moderator status is the highest status granted to users, which is publicly displayed next to a user's name in the form of a diamond symbol (♦). With their status and control, moderators can directly influence the success of a community.

Stack Exchange moderatorship is decided by a community election, which closely mimics real-world democratic elections. Elections are held within individual communities and usually take place once a year. A typical election consists of three phases, nomination, primary, and final election. During the seven-day nomination phase, users can propose candidates, also themselves. If a user gets nominated by other users, the nominated user must accept the nomination before proceeding to the subsequent phase. Nominees are required to write a short, freely editable summary of why they might make a good community moderator. Most communities hold public Q&As with nominees, during which candidates must publicly stand questions of other users. Typical questions demand more details on the candidate's availability per day, reaction time, and prior experience, but also seek to challenge the user by showing potential misbehaviors in the past. Examples of such questions include, "[…] you have posted 5 questions on Meta (that still survive today), plus another 20 on Meta.SE. Do you plan to become more active?", "You recently hammer-closed a question as a duplicate […]. However, a long comment thread and disputatious edits from OP followed your action, and the question is now reopened. How else might you have handled the situation, in retrospect?", "Based on my short encounter with you, I doubt you'd be a good moderator".

During primary, the top 30 nominees by reputation score advance to preliminary community voting. Primary candidates are displayed in random order and votes are public. Voters have one up or down vote per candidate. After 4 days, the top 10 candidates based on primary vote score proceed to the final election. The election usually takes between two to three weeks. The election is based on a single transferable vote system. Under this voting system, a voter has a single vote that is initially allocated to their most preferred candidate. Once a candidate has more votes than the quota, the candidate is elected. Throughout the counting procedure and as candidates are elected, the vote is transferred according to the voter's preferences, proportional to any surplus or discarded votes. During the election, candidates are displayed in random order on the ballot to avoid "donkey voting" or strategic gaming.

Intermediate voting results are not displayed. After the election, moderatorship is granted for life.

The election setup has two features that allow us to overcome the outlined two-stage selection bias and other inferential challenges. The first feature of our research design is that it allows for causal inference by exploiting very close election outcomes in a regression discontinuity design [18, 41]. In such a design, we compare the behavior of users that closely won such an election with users that closely lost the election. Such “close call” elections are akin to a random assignment of moderatorship to users, the smaller the margin of votes. The identifying assumption of a regression discontinuity (RD) design is that, in addition to candidate, community, or time factors, there is some randomness that determines the outcome of a close election. Intuitively, the difference between a user that is granted status with 50.1% of the votes and a user that loses the election with 49.9% of the votes seems negligible [18, 41]. Close call elections have been used as an identification strategy in various studies in political science and economics [20, 42].

Stack Exchange held the first elections in December 2010. We identified all elections held on Stack Exchange since then until January 2017 and collected the published ballot data. The ballot data contained identifiers of the candidates and the vote preferences they received, as well as descriptive information on the time, seats, and quota of each election. In total, our sample includes 154 elections across 70 communities. Voter turnout in these elections ranged from 3% to 42%, with the median election having a turnout of 13%. On average, 10,984 voters were eligible, with 972 votes casted on average. On average, 6.54 candidates competed for 2.44 seats. Aside from one election in 2010, elections are almost uniformly distributed over the years we consider. When inspecting the density function of vote shares, we see that many observations are close to the cutoff (1), supporting the use of a RD design in our context.

We matched ballot data with candidate-specific information provided directly by Stack Exchange. This data included static demographics of candidates as well as their posting, reputation, and editing behavior. To assess changes in candidate behavior, we consider a post-election window of six months after the election. We stop observing candidate behavior six months after the election to provide enough variance for our estimations but also to avoid confounding events like preceding or succeeding elections. Our final dataset contains 1,012 candidates. To assess pre-election differences between candidates, we collect candidate-level data also until six months before the election. Our sample covers elections in 70 communities. The communities cover a wide range of topics, including various computer-related topics like operating systems, databases, and programming but also mundane topics like cooking, home improvement, photography, or music-making. On average, communities in our sample listed more than 267,000 questions and 3.5 Mio answers contributed by 177,000 users. The average community in our sample has existed for 4.7 years and held 2.1 elections over the observation period.

We measure candidate contributions in terms of the number of posts they contribute (POSTS) and the word count of the posts (POSTLENGTH), the number of upvotes casted (UPVOTES), and the number of downvotes casted (DOWNVOTES).

We measure community members' perceptions of moderators by their rating behavior, in terms of the rating they give to a candidate's questions (QUESTION RATING), the rating of a candidate's answers (ANSWER RATING), the number of views for a candidate's questions (QUESTION VIEWS), and the number of times a candidate's question is bookmarked (i.e., marked as favorite) by others (BOOKMARKS). Table 1 describes the sample. Overall, the sample includes 1,012 candidates over the period from 2010 to 2017.

Table 1. Summary Statistics

		Mean	*S.D.*	*Min.*	*Max.*	*P25*	*Median*	*P75*
1	Posts	8.88	19.46	0	164	0.00	2.00	8.00
2	Post length	895.89	897.25	0	12186	0.00	787.38	1140.15
3	Log(Answers)	0.10	0.37	0	8	0.00	0.01	0.06
4	Log(Comments)	265.45	799.93	0	10354	0.00	0.00	192.25
5	Log(Bookmarks)	0.03	0.15	0	4	0.00	0.00	0.02
6	Post rating	0.14	0.52	0	10	0.00	0.01	0.08
7	Badges	46.54	19.34	0	87	17.00	50.00	62.00
8	Months since joined	18.91	30.03	0	429	0.00	10.00	23.00

For each candidate i, $\mathrm{Y}_{\mathrm{i},\,t+1}$ denotes the outcome of interest in t+1, while $share_{i,t}$ is the "running variable" that determines treatment assignment. Candidates are assigned to the treatment group (1) if their vote share in t exceeds the defined quota and otherwise are assigned to the control group (0):

$$\mathrm{W_i} = f(share_{i,t}) = \begin{cases} 1 & if\ share_i > quota \\ 0 & if\ share_i < quota \end{cases} \tag{1}$$

Gelman and Imbens [43] outline higher order polynomials as a potential source of bias to RD estimates. Therefore, we fit second order polynomials. Following the suggestions of Imbens and Lemieux [41] and Lee and Lemieux [19], we use a triangular kernel. The robustness section shows that the choice of these parameters matters little for our dataset, as the estimates remain comparable across various parameter choices. Accordingly, we rely on the standard RD in the following form:

$$Y_{i,t+1} = \alpha + \beta\ W_i + f\ \left(share_{i,t}\right) + community_i + \epsilon_i \tag{2}$$

In this equation, $f(\cdot)$ is the function of the vote share outlined above, $share_{i,t}$ is the continuous vote share, $\mathrm{W_i}$ is an indicator of whether the candidate was elected, $\mathrm{community_i}$ controls for community-specific effects, and ϵ_i denotes the error term. The coefficient of interest is β, which gives the effect of being treated with moderator status. In the model, we assume that $\mathrm{W_i}$ is randomly distributed over candidates near the cutoff c.

4 Results

Direct effects. To assess the effects of moderatorship on the focal user, we compare candidates who were closely elected as moderators (winners) to candidates who closely lost at the ballots (losers). We therefore estimate equation (2) for increasingly small bandwidths around the vote share cutoff. Table 2 reports estimates of POSTS for close winners and losers for small bandwidths around the cutoff. Columns (1) - (4) indicate that, for increasingly small bandwidths around the cutoff, a consistently significant discontinuity in POSTS. The regressions estimate the increase between 10.5 and 15.3 posts. Columns (5) – (8) report the discontinuity in POST LENGTH. The models estimate an increase of POST LENGTH between 8.42 and 8.87 words per post. Taken together, both findings confirm what the above graphical analysis indicated: moderatorship triggers a strong increase in content contributed to the community, which stands in stark contrast to the suggestion that moderatorship discourages contributions.

Spillover effects. The findings so far indicate that moderators contribute more content. The natural follow-up question is whether these effects translate into greater discourse and exchange among other community members. To assess such spillover effects, we compare how community members react to a moderator's posts and how they participate in discussions initiated by a moderator. We would conclude in favor of positive effects on community discourse, if we were to observe that other community members contribute more posts as a response to a candidate's content (RESPONSES), if other members comment more actively the postings of moderators (COMMENTS), and if other members perceive a candidate's postings as more valuable, in terms of bookmarking it more often (BOOKMARKS) or rating it higher (RATING).

Table 2. Regression Discontinuity Estimates of Direct Moderatorship Effects

	Posts				*Post length*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth	15%	12.5%	10%	7.5%	15%	12.5%	10%	7.5%
Elected	10.46^{**}	12.37^{***}	12.63^{**}	15.32^{**}	8.42^{*}	8.54^{*}	7.81	8.87^{*}
	(3.76)	(3.38)	(4.21)	(4.84)	(3.61)	(3.80)	(4.15)	(4.42)
N	340	313	283	254	340	313	283	254

Note: *, **, *** indicate significance at the 5%, 1%, and .1% levels, respectively.

Data in Table 3 provides strong evidence in favor of such a spillover effect. Columns (1) - (4) estimate that community members respond more actively to a moderator's postings. Columns (5) to (8) estimate that community members comment a moderator's posts more intensively, which appears less marked than the discontinuity in RESPONSES. Columns (9) to (12) and columns (13) to (16) show strong discontinuities in ratings by other community members. Users bookmark a moderator's postings more often and rate it more positively. In sum, the regressions provide strong evidence for a spillover effect: moderatorship triggers a more desirable behavior of moderators, which has ripple effects on the contributions of other community members.

Table 3. Regression Discontinuity Estimates of Spillover Moderatorship Effects

	Responses				*Comments*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandw.	15%	12.5%	10%	7.5%	15%	12.5%	10%	7.5%
Elected	9.05**	11.86***	13.23***	11.96***	5.50*	6.88**	7.42***	6.61**
	(3.46)	(3.55)	(3.57)	(3.30)	(2.19)	(2.29)	(2.25)	(2.04)
N	340	313	283	254	340	313	283	254

	Bookmarks (weighted)				*Post rating (weighted)*			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Bandw.	15%	12.5%	10%	7.5%	15%	12.5%	10%	7.5%
Elected	13.53**	13.94***	12.79**	10.06**	29.68**	31.99***	32.25***	28.36***
	(4.12)	(4.16)	(4.04)	(3.73)	(9.27)	(9.34)	(9.12)	(8.60)
N	340	313	283	254	340	313	283	254

Note: *, **, *** indicate significance at the 5%, 1%, and .1% levels, respectively.

Robustness and rival explanations. We run four substantial robustness checks, which we can detail only shortly due to the space limitations. The RD approach in this paper makes two identifying assumptions [18, 44]. First, candidates are balanced on all covariates prior the election. Second, we do not find evidence that candidates are able to manipulate election outcomes. Third, we assure that the effects are not driven by losing candidates becoming alienated from a community as a result of the election. Finally, we run placebo regressions in which we randomly distribute moderator status—we do not observe significant effects, which supports the robustness of our findings.

5 Discussion and Conclusion

This paper contributes to the ongoing debate on the design and effectiveness of online community governance by providing new causal evidence about the effect of granting community members decision-making rights over a community on their behavior. To isolate the effects, we exploit the quasi-random assignment of moderatorship that occurs in close-call elections of moderators on Stack Exchange, one of the largest networks of online communities. We find that moderatorship catalyzes behavior desirable for online communities. They contribute more content, and their contributions are perceived as more helpful by community members. These lasting effects are likely to result from the psychological effects of holding moderator status within a community rather than resulting from newly granted instruments or resources. Whom should a community appoint as moderators? The evidence here identifies users with a below-median standing in a community to react strongest to moderatorship. In this sense, moderatorship should not necessarily be seen as a "crown" that is bestowed as a finishing touch of a long-standing and well-established community membership but rather as an "office" that should also be open to less-established community members.

These findings have important implications for theory and policy design. First, the finding runs counter to the paradigm that moderatorship inhibits or even demolishes desirable user behavior. Instead, our findings support those who argue that moderatorship acts as an important catalyst for even further, desirable contributions to a community. Second, the findings demonstrate that moderatorship catalyzes user behavior through psychological mechanisms rather than due to instruments and resources that accompany such an appointment. This suggests that fostering the psychological mechanisms of moderatorship—for example by praising the position among community members, signaling moderatorship by badges or other means, and creating high-awareness processes for moderatorship applications—might be beneficial to realize the full gains from this form of self-governance. Third, the fact that the effects of moderatorship are rather heterogeneous suggests that more work is needed to better understand how the effects of moderatorship vary across users. We hope that our study stimulates further research examining governance forms and decision-makers in digital environments.

References

1. Arthur, C.: What Is the 1% Rule? The Guardian. (2006).
2. Jeppesen, L.B., Frederiksen, L.: Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer-Controlled Music Instruments. Organization Science. 17, 45–63 (2006).
3. Kraut, R.E., Resnick, P.: Building Successful Online Communities: Evidence-Based Social Design. MIT Press, Cambridge, MA (2012).
4. Lakhani, K.R., Von Hippel, E.: How Open Source Software Works: "Free" User-to-User Assistance. Research Policy. 32, 923–943 (2003).
5. Wasko, M.M., Faraj, S.: Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. MIS Quarterly. 29, 35–57 (2005).
6. Chen, Y., Harper, F.M., Konstan, J., Xin Li, S.: Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens. The American Economic Review. 100, 1358–1398 (2010).
7. Goes, P.B., Guo, C., Lin, M.: Do Incentive Hierarchies Induce User Effort? Evidence from an Online Knowledge Exchange. Information Systems Research. 27, 497–516 (2016).
8. Williams, R.L., Cothrel, J.: Four Smart Ways to Run Online Communities. MIT Sloan Management Review. 41, 81–91 (2000).
9. Faraj, S., Lakani, K., Monteiro, E., von Krogh, G.: Online Community as Space for Knowledge Flows. Information Systems Research. 27, 668–684 (2016).
10. Butler, B., Sproull, L.K., Kiesler, S.: Community Effort in Online Groups: Who Does the Work and Why. In: Leadership at a Distance: Research in Technologically-supported Work. pp. 171–194. Lawrence Erlbaum Associates, Mahwah, NJ (2007).

11. Von Krogh, G., Spaeth, S., Lakhani, K.R.: Community, Joining, and Specialization in Open Source Software Innovation: A Case Study. Research Policy. 32, 1217–1241 (2003).
12. Matias, N.J.: Civic Labor of Online Moderators. In: Oxford Internet, Policy, and Politics Conference. , Oxford, UK (2015).
13. Chen, A.: When the Internet's 'Moderators' Are Anything But, https://www.nytimes.com/2015/07/26/magazine/when-the-internets-moderators-are-anything-but.html.
14. Peckham, M.: Reddit Demotes Technology Section to Punish Lazy Moderators, http://time.com/71644/reddit-demotes-technology-section-to-punish-lazy-moderators/.
15. Sauder, M., Lynn, F., Podolny, J.M.: Status: Insights from Organizational Sociology. Annual Review of Sociology. 38, 267–283 (2012).
16. Azoulay, P., Stuart, T., Wang, Y.: Matthew: Effect or Fable? Management Science. 60, 92–109 (2014).
17. Ma, M., Agarwal, R.: Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. Information Systems Research. 18, 42–67 (2007).
18. Angrist, J.D., Pischke, J.-S.: Mostly Harmless Econometrics: An Empiricist's Companion. Princeton Press, Princeton, NJ (2009).
19. Lee, D.S., Lemieux, T.: Regression Discontinuity Designs in Economics. Journal of Economic Literature. 48, 281–355 (2010).
20. Lee, D.S., Moretti, E., Butler, M.J.: Do Voters Affect or Elect Policies? Evidence from the U. S. House. The Quarterly Journal of Economics. 119, 807–859 (2004).
21. Armstrong, A., Hagel, J.: The Real Value of Online Communities. Knowledge and Communities. 74, 85–95 (2000).
22. Butler, B.S.: Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. Information Systems Research. 12, 346–362 (2001).
23. Faraj, S., Jarvenpaa, S.L., Majchrzak, A.: Knowledge Collaboration in Online Communities. Organization Science. 22, 1224–1239 (2011).
24. Bayus, B.L.: Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community. Management Science. 59, 226–244 (2013).
25. Porter, C.E., Donthu, N.: Cultivating Trust and Harvesting Value in Virtual Communities. Management Science. 54, 113–128 (2008).
26. Huang, P., Mithas, S., Tafti, A.: Platform Sponsor's Investments and User Contributions in Knowledge Communities: The Role of Knowledge Seeding," accepted for publication. MIS Quarterly. (2017).
27. Kane, G.C., Fichman, R.G., Gallaugher, J., Glaser, J.: Community Relations 2.0. Harvard Business Review. 87, 45–50 (2009).
28. Aaltonen, A., Seiler, S.: Cumulative Growth in User-Generated Content Production: Evidence from Wikipedia. Management Science. 62, 2054–2069 (2016).
29. Varian, H.R.: Microeconomic Analysis. Norton & Company, New York, NY (1992).
30. Wasko, M.M., Teigland, R., Faraj, S.: The Provision of Online Public Goods: Examining Social Structure in an Electronic Network of Practice. Decision Support Systems. 47, 254–265 (2009).

31. Lerner, J., Tirole, J.: Some Simple Economics of Open Source. The Journal of Industrial Economics. 50, 197–234 (2002).
32. Von Krogh, G., Von Hippel, E.: The Promise of Research on Open Source Software. Management Science. 52, 975–983 (2006).
33. O'Mahony, S., Ferraro, F.: The Emergence of Governance in an Open Source Community. Academy of Management Journal. 50, 1079–1106 (2007).
34. Ren, Y., Kraut, R.E.: A Simulation for Designing Online Community: Member Motivation, Contribution, and Discussion Moderation. Working Paper. (2011).
35. Hann, I.-H., Roberts, J.A., Slaughter, S.A.: All Are Not Equal: An Examination of the Economic Returns to Different Forms of Participation in Open Source Software Communities. Information Systems Research. 24, 520–538 (2013).
36. Bateman, P.J., Gray, P.H., Butler, B.S.: Research Note—The Impact of Community Commitment on Participation in Online Communities. Information Systems Research. 22, 841–854 (2011).
37. McWilliam, G.: Building Stronger Brands Through Online Communities. Sloan Management Review. 41, 43 (2000).
38. Berger, J., Cohen, B.P., Zelditch Jr, M.: Status Characteristics and Social Interaction. American Sociological Review. 37, 241–255 (1972).
39. Roberts, J.A., Hann, I.-H., Slaughter, S.A.: Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects. Management Science. 52, 984–999 (2006).
40. Spaeth, S., von Krogh, G., He, F.: Research Note—Perceived Firm Attributes and Intrinsic Motivation in Sponsored Open Source Software Projects. Information Systems Research. 26, 224–237 (2014).
41. Imbens, G.W., Lemieux, T.: Regression Discontinuity Designs: A Guide to Practice. Journal of Econometrics. 142, 615–635 (2008).
42. Eggers, A.C., Fowler, A., Hainmueller, J., Hall, A.B., Snyder, J.M.: On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races. American Journal of Political Science. 59, 259–274 (2015).
43. Gelman, A., Imbens, G.: Why High-order Polynomials Should not be Used in Regression Discontinuity Designs. National Bureau of Economic Research (2014).
44. McCrary, J.: Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. Journal of Econometrics. 142, 698–714 (2008).

Understanding Job Satisfaction of Crowd Workers: An Empirical Analysis of Its Determinants and Effects

David Durward[1], Ivo Blohm[2]

[1] University of Kassel, Information Systems, Kassel, Germany
david.durward@uni-kassel.de
[2] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
ivo.blohm@unisg.ch

Abstract. Crowd work has emerged as new pattern of digitally mediated collaboration. In this paper, we focus on the determinants and effects of crowd workers' job satisfaction – a perspective that has been largely neglected by current crowdsourcing research. We report results from a survey of 161 crowd workers participating in crowdsourced software testing. Our research shows that job satisfaction mediates the effects of monetary rewards, hedonic value, and cognitive stimulation on the intention to participate in future testing tasks. By contrast, factors of work context (i.e., flexibility and provided information) have no effects. We contribute to the literature by unraveling job satisfaction as causal mechanism influencing future participation. For practice, our results help to design more effective tasks in crowd work.

Keywords: Crowdsourcing, Crowd Work, Job Satisfaction, Motivation.

1 Introduction

The rise of digitalization provides a shared new communication and collaboration infrastructure. Further it enables crowdsourcing as an alternative system of organizing. As one consequence, a novel form of digital work has emerged; i.e. *crowd work*. According to Durward et al. [1], crowd work reflects a digital form of gainful employment based on the crowdsourcing idea in which an undefined mass of people creates digital goods via an open call on IT-facilitated platforms. The potentials of crowd work for an individual include the opportunity to generate an additional income [2] on a full- or part-time-basis. Crowd work is gaining importance since the number of platforms and crowd workers has been growing continuously. Thus, it is not surprising that the World Bank estimates the total crowd work market to increase from $4.8 billion in 2016 up to $25 billion by 2020 [3].

Despite this rather growing importance, research on crowd work is still in its inception, in particular regarding the ones who perform the work, i.e. the crowd workers. Most of prior research focused only on the business perspective and analyzed underlying strategies, potential risks and benefits [e.g., 4] or incentive mechanisms [e.g., 5] in crowd work projects. However, the perceptions of individual

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

crowd workers are equally important and have been largely neglected in existing research [6]. Since a balanced evaluation of crowd work requires multiple perspectives of the involved stakeholders, we consider this to be a serious shortcoming. Current research lacks sufficient insights particularly with respect to issues that affect the behavior of crowd workers. One facet of experiencing work – which has long been analyzed especially within organizational behavior, psychology or ergonomics – has hitherto been completely neglected in crowd work contexts: i.e., *job satisfaction*. Hence, we intend to fill the outlined research gaps regarding the individual perspective by addressing the following research question:

How do different motivations affect the job satisfaction of crowd workers?

Therefore, we address two issues with our study: First, we dissect factors that might affect crowd workers' job satisfaction (i.e. *information provision; flexibility; hedonic value; cognitive stimulation; monetary reward; career-related benefit*) and second, we analyze the relationship between job satisfaction and continuance intention.

2 Theoretical Background

2.1 Crowdsourcing and Crowd Work

According to Blohm et al. [7], the fundamental idea of crowdsourcing is that a crowdsourcer (a company, an institution or a non-profit organization) proposes to an undefined group of contributors or crowd workers (individuals, formal or informal teams, other companies) the voluntary undertaking of a task presented in an open call. In addition, the ensuing interaction process unfolds over IT-based platforms (i.e. crowdsourcing intermediaries). Prior research has found important differences between the notions of crowdsourcing and crowd work [e.g., 1, 2]. According to Durward et al. [1] crowd work resembles a distinct type of labor that is located at the intersection of digital work and gainful employment. While crowd work is always paid, participation in crowdsourcing initiatives may have different motives and does not necessarily require financial remuneration. Thus, from an individual's perspective, crowd work reflects a kind of digital gainful employment that is based on crowdsourcing as organization principle.

In general, there are much less insights into individual crowd workers' perspectives. In recent studies, some promising approaches focused on demographic data [8] or environmental aspects [e.g., 6, 9]. In their qualitative study, Deng and Joshi [6] found task characteristics and a digitally enabled environment to shape crowd workers' continued participation. We intend to join this stream and expand this research by analyzing the perception of crowd work based on motivations and examining its effects on the individual crowd worker in a quantitative study.

2.2 Job Satisfaction

Within the organizational behavior and psychology literature, job satisfaction plays a significant role and is defined by various researchers, [e.g., 10] as an attitude of employees towards their work or as a positive emotional state that is caused during the experiences at work. In IS context, researchers propose that satisfaction correlates with attributes such as flexibility, empowerment or involvement [11]. Several studies find empirical support for these attribute models of satisfaction in IS [12], and address the essential questions of the elements that need to be focused on. Since these models depend on which attributes are examined, they must be developed for every new kind of IS artifact or aspect: i.e., crowd work.

Motivations are predominant constructs in analyzing satisfaction, since they act as drivers, which impel individuals toward action [13]. On this basis, IS researchers propose models that posit satisfaction as a response to judgments about goals and their level of achievement [14]. This offers a practical basis for the motivation concept, whose nature is not yet fully understood, and has been empirically supported by studies in IS research [14].

Against this background, we address certain motivational dimensions to be considered as antecedents of crowd workers` satisfaction. Further, according to Deng and Joshi [6], we also include the work context as affecting environmental dimension. Since crowd work represents a new form of digital gainful employment, we propose an aggregate view by drawing on well-established perspectives in IS [e.g., 13, 14] to analyze job satisfaction of crowd workers.

3 Research Model and Hypotheses Development

Drawing on motivation theories, we develop our research model to discover the black box of certain effects on job satisfaction in crowd work initiatives. Figure 1 depicts our research model, in which intrinsic and extrinsic motivation as well as the work context positively affect crowd workers` satisfaction and indirectly continuance intention.

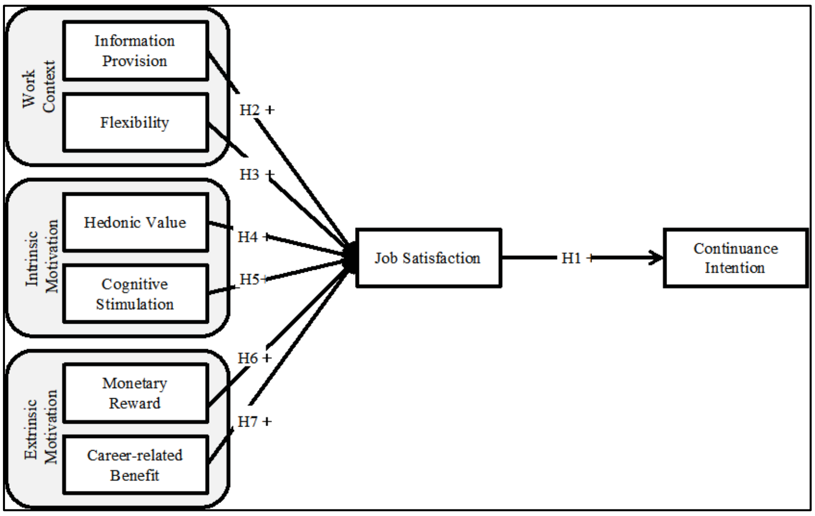


Figure 1. Research Model

3.1 Continuance Intention

Satisfaction is a positive emotion, which arises from an evaluation of an object or activity and represents, particularly in IS research, a critical factor for understanding users' IS continuance or post-adoption behavior [15]. Besides, Sun et al. [16] notes that satisfaction with prior task participation experience, directly effects the continuous task participation of people in online working-context. Hence, we assume that job satisfaction is positively associated with the crowd workers` intention to work for the crowdsourcing intermediary in the future.
- *H1: Job satisfaction positively affects crowd workers' continuance intention.*

3.2 Work Context

A certain aspect of work environment includes the concept of user information satisfaction that is defined as the extent to which users believe the information system available to them meets their information requirements [17]. A study of Ang & Soh [18] indicated that job satisfaction and user information satisfaction were highly correlated. Due to the fact, that crowd work tasks are often processed independently, this comprises informational requirements essential for task fulfilment. Hence, we assume that contentment with the information provided has an influence on a crowd workers´ job satisfaction.
- *H2: Information provision positively affects crowd workers' job satisfaction.*

Since the rise of new IT applications and communication networks the perceptions of work have changed and made it possible to work at any location at any time [19]. Researchers found that this emerging flexibility or the access to flexible work arrangements is, inter alia, associated with higher levels of job satisfaction and engagement [e.g., 20]. Concerning crowd workers, this flexibility addresses the opportunity of self-selection and freedoms in task processing. Therefore, we assume that flexibility in this new work context is a relevant antecedent of satisfaction.
- *H3: Flexibility positively affects crowd workers' job satisfaction.*

3.3 Intrinsic Motivation

Intrinsic motivation refers to the pleasure associated with the activity itself so that in IS an intrinsically motivated user is driven by benefits that derive from the interaction with the system per se [21]. In this regard, hedonic value in the field of IS specifies the extent to which fun can be derived from using the system as such [22]. Due to the freelance and voluntary nature of online work such as crowd work, particularly the hedonic value by working on interesting tasks can satisfy individuals [23] and give important insights regarding participation. Hence, in crowd work initiatives, it is essential to analyze perceived enjoyment that satisfies crowd workers in a hedonic intrinsic manner.
- *H4: Hedonic value positively affects crowd workers' job satisfaction.*

With their study on motivations in open source software projects, Lakhani and Wolf [24] have identified, inter alia, the intellectual stimulation derived from writing code as a main goal for participation. By processing splitted subtasks via the Internet and perceiving cognitive challenges, particularly in collaboration-based initiatives, crowd workers can also attain this intrinsic goal by participation. Hence, we assume that cognitive stimulation of an individuals` intellect is an essential goal of crowd workers and can be an antecedent of their satisfaction.
-*H5: Cognitive stimulation positively affects crowd workers' job satisfaction.*

3.4 Extrinsic motivation

Extrinsic motivation refers to the value an individual places on the ends of an action and the probability of reaching these ends [25]. Here, the underlying motives result directly from external stimuli that are perceived from the situational context. Particularly financial rewards affect the perceived fairness and job satisfaction [26], in almost any type of work arrangement. Although, the amount of compensation varies greatly in crowd work, it tends to be a key motive of participating and needs to be analyzed regarding satisfaction. Hence, we assume monetary reward to be a major factor in crowd work context with significant influence on the crowd workers.
- *H6: Monetary reward positively affects crowd workers' job satisfaction.*

Career decisions are related to specific jobs, and the attributes of the specific jobs can exert strong influences on these decisions [27]. Within open source software development, academic research has posited that external motivational factors in the form of better job opportunities or career advancement are the main goals of effort [24]. Career theories collectively provide a valuable framework for examining crowd workers' career pathway since they identify the critical work factors that drive the individuals to become a crowd worker [27]. Thus, we examine whether potential career advantages in crowd work initiatives are goals to be attained and further satisfy the testers.
- *H7: Career-related benefit positively affects crowd workers' job satisfaction.*

4 Research Method

4.1 Research Context and Data Collection

To empirically test our research model, we conducted a survey to collect data from crowd workers on a crowdsourcing platform in Germany. For our study, we chose a German start-up intermediary called testIO (https://test.io/de/) – former testcloud – that offers software testing services for companies intending to partly or fully outsource their testing activities to a certain crowd (in sum over 20,000 testers). testIO's crowd workers were informed about our survey via personal mails. For motivating testOI's crowd workers to take part in our study, we communicated to the crowd that we would donate two euros at "betterplace.org" for a charitable project for each complete survey. A total of 161 crowd workers provided a fully completed questionnaire.

In the following, we outline only some key characteristics of the sample: In the sample of the 161 crowd workers, 69% are males, whereas 86% of the participants are from Germany, 7% from India and 7% from other countries (Switzerland, Austria, etc.). 66.7% of the interviewees have a higher education degree. The “average crowd worker” of our sample is 34 years old, has about four and a half years of experience with software testing and has been registered at testIO for about 7 months.

4.2 Measures

All scales used in our study were adapted from previous studies and modified to fit these specific testing tasks in crowd work. We therefore had to modify the wordings of the existing items and adjust them to the study context. We used a seven-point Likert scale for all items. To measure hedonic value, we adapted the scale used by Sun et al. [16]. Cognitive stimulation was measured using the scale provided by Ke and Zhang [28]. Items to measure monetary reward were adapted from Spector [10], whereas career-related benefits was measured with the scale used by Clary et al. [29]. Measures for information provision were adapted from Borg [30] and for flexibility we adapted the scale used by Rimann and Udris [31]. Finally, for measuring job satisfaction and continuance intention, we adapted the items provided by Lim [32].

5 Data Analysis and Results

5.1 Construct Validation

In order to confirm validity and reliability of our measures, we applied exploratory and confirmatory factor analysis using SPSS 19 and SmartPLS 2.0 (cf. Table 1). The Measure of Sampling Adequacy was 0.91, indicating excellent applicability of exploratory factor analysis [33]. We extracted 8 factors that could be clearly interpreted. Alphas of at least 0.831 suggest good reliability of factors. However, we eliminated one item from our hedonic value scale as it did not load unambiguously on one factor. In confirmatory factor analysis, Composite Reliabilities (CR) exceeded values of 0.5, and the Average Variance Extracted (AVE) for each factor surpassed 0.5. Thus, convergent validity could be assumed [34]. According to Kline [35] factor loading values higher than the threshold value of 0.5 are adequate. The discriminant validity was checked by using the Fornell-Larcker criterion, which claims that the square root of one factor’s AVE should be higher than its correlation with every other factor [36]. Thus, discriminant validity could be assumed (cf. Table 2).

Table 1. Exploratory and Confirmatory Factor Analysis

Label	Factors								Factor Loading	CR	AVE	α
	MR	CS	CB	IP	FL	CI	HV	JS				
MR1	**.888**	.077	.110	.171	.133	.156	.132	.114	.931**	.948	.821	.851
MR2	**.872**	.141	.142	.224	.108	.121	.103	.172	.944**			
MR3	**.748**	.203	.163	.181	.226	.245	.074	.115	.844*			
MR4	**.683**	.045	.170	.241	.136	.292	.166	.222	.862**			
CS2	.170	**.877**	.215	.107	.068	.098	.113	.183	.935**	.956	.878	.931
CS1	.138	**.867**	.162	.077	.164	.119	.171	.107	.954**			
CS3	.054	**.818**	.234	.221	.162	.159	.090	.145	.924**			
CB1	.123	.210	**.878**	.136	.165	.037	.013	.024	.909**	.942	.844	.908
CB2	.112	.282	**.862**	.135	.135	.011	.055	.138	.935**			
CB3	.221	.096	**.787**	.164	.144	.070	.134	.288	.911**			
IP3	.197	.166	.182	**.816**	.201	.161	.059	.148	.909**	.942	.843	.907
IP1	.276	.132	.137	**.807**	.235	.130	.053	.087	.918**			
IP2	.229	.125	.153	**.796**	.175	.152	.204	.142	.928**			
FL1	.103	.074	.192	.147	**.812**	.133	.226	.074	.854*	.899	.747	.831
FL2	.178	.308	.173	.222	**.737**	.036	.123	.105	.872**			
FL3	.259	.083	.140	.306	**.690**	.214	.077	.171	.867**			
CI2	.338	.115	.129	.185	.047	**.738**	-.022	.270	.915**	.910	.772	.851
CI1	.267	.163	-.052	.246	.251	**.715**	.347	.080	.824*			
CI3	.222	.212	.036	.115	.172	**.714**	.425	.057	.895**			
HV3	.116	.188	.130	.123	.160	.172	**.856**	.146	.911**	.922	.855	.832
HV1	.293	.179	.045	.153	.315	.309	**.655**	.203	.939**			
JS3	.262	.245	.263	.268	.128	.191	.145	**.695**	.904**	.929	.813	.885
JS2	.288	.289	.375	.145	.256	.141	.163	**.635**	.914**			
JS1	.383	.293	.127	.168	.163	.245	.325	**.594**	.904**			
Eigen-value	11.44	2.29	1.69	1.43	1.13	.900	.686	.611				

n = 161; ** $p < 0.01$; * $p < 0.05$

[a]MSA = 0.91; Bartlett-test of sphericity: $\chi^2 = 1422.74$, $p = 0.000$; principal component analysis; varimax-rotation; The bold values indicate the attribution of the variables to one of the four factors.

MR = Monetary Reward, CS= Cognitive Stimulation; CB = Career-related Benefit, IP = Information Provision, FL = Flexibility, CI = Continuance Intention, HV = Hedonic Value, JS = Job Satisfaction, CR = Composite Reliability, AVE = Average Variance Extracted

As both the dependent and the independent variables were measured by means of a survey, we carefully checked the extent of common method variance. First, we applied Harman's single factor test applying principal component analysis. This approach suggests the occurrence of common method bias in case one factor accounts or the majority of variance of all included items [37]. In our case, one factor emerged reflecting a considerable part of the covariance. However, as this factor did not account for the majority of covariance, there is no indication of substantial common method bias. Second, we inspected the correlation matrix (cf. Table 2). The highest correlation is 0.674. The evidence of common method bias would have resulted in extremely high correlations (>0.9) [38]. Finally, we followed the suggestions of Lindell and Whitney [39] who propose to test for common method bias by assessing the correlation between a theoretical unrelated construct (i.e. a marker variable) with the principal study constructs. Common method bias can be assumed if these correlations are substantial. We applied a modified test in which we assessed the correlation between our study variables and a weakly related construct as we did not measure completely unrelated constructs in order to economize on survey items [38]. In our case, we measured agreeableness as central personality trait of crowd workers as these might reflect a central prerequisite for online collaboration. The highest correlation between agreeableness and the other study constructs is -0.1 (n.s.) indicating that our data does not suffer from substantial common method variance.

5.2 Hypotheses Test

We use Partial Least Square (PLS) analysis using SmartPLS 2.0 to test our research model. We have chosen PLS as it does not require assumptions of normally distributed data and works well for complex models even with smaller sample sizes [40]. Table 2 contains correlations, means, and standard deviation of our central study variables. We tested our hypotheses with 500 bootstrapping resamples and 161 cases.

Table 2. Descriptives and Correlations

Variable	Min	Max	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) MR	**1.0**	7.o	4.751	1.367	**.906**							
(2) CS	**1.0**	7.o	4.983	1.313	.399*	**.937**						
(3) CB	**1.0**	7.o	3.824	1.497	.429**	.501**	**.919**					
(4) IP	**1.0**	7.o	4.990	1.372	.581**	.424**	.443**	**.918**				
(5) FL	**1.0**	7.o	5.217	1.245	.517**	.452**	.470**	.591**	**.864**			
(6) CI	**1.0**	7.o	5.694	1.145	.637**	.441**	.278*	.526**	.500**	**.879**		
(7) HV	**1.0**	7.o	5.643	1.166	.669**	.606**	.594**	.579**	.569**	.612**	**.925**	
(8) JS	**1.0**	7.o	4.787	1.206	.507**	.462**	.330*	.445**	.560**	.646**	.606**	**.902**

n = 161; ** p < 0.01; * p < 0.05
The bold values indicate the squared AVE for each factor for assessing discriminant validity. MR = Monetary Reward, CS= Cognitive Stimulation; CB = Career-related Benefit, IP = Information Provision, FL = Flexibility, CI = Continuance Intention, HV = Hedonic Value, JS = Job Satisfaction, SD = Standard Deviation

Applying PLS analysis, we found a positive and significant effect of job satisfaction on continuance intention (β = .617, p < 0.01) supporting H1. Our work context variables, information provision and flexibility had no statistically significant influence on job satisfaction such that we reject H2 and H3. Our data indicates that hedonic value (β = .233, $p < .01$), cognitive stimulation (β = .217, $p < .01$), monetary reward (β = .303, $p < .01$) und career-related benefits (β = .225, $p < .01$) are positively and statistically significant associated with job satisfaction. Therefore, H4, H5, H6 and H7 are supported by our study as Figure 3 illustrates.

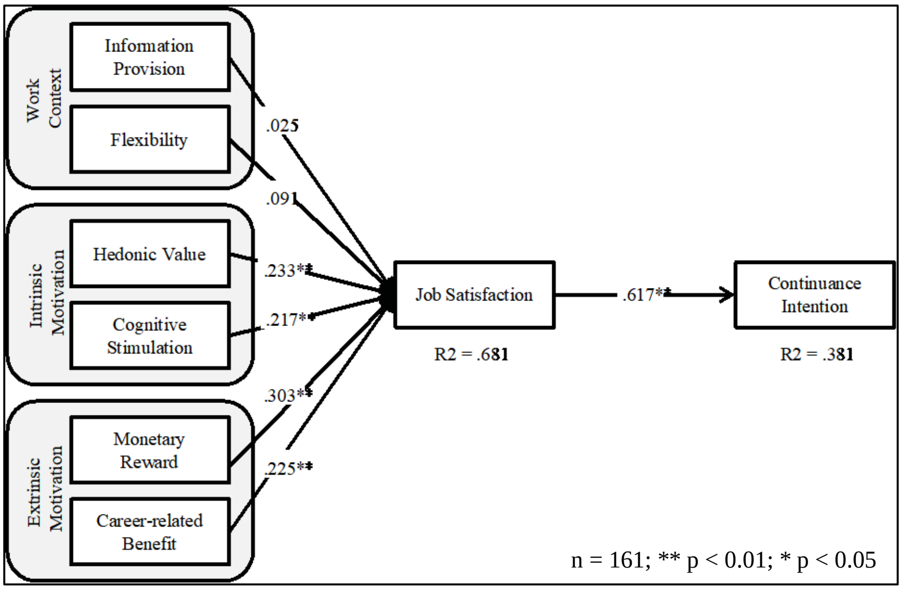


Figure 2. PLS Analysis

6 Discussion

6.1 Theoretical Implications

In line with existing research, our study shows that crowd workers are not only motivated by financial compensation, but also by non-financial motives [e.g., 8]. These results have important theoretical implications. First, our results show that besides financial compensation other internal as well as external motivations seem to be relevant to crowd workers. In line with Deng and Joshi [6] we found antecedences of crowd workers satisfaction as well as continued participation. Second, besides external motivation, we particularly extend prior research on internal motivation like hedonic value [23] and cognitive stimulation [24], since we found these to satisfy crowd workers in testing contexts as well. Thus, satisfaction and continuance intention are invoked by hedonic value and cognitive stimulation crowd workers experience while solving testing tasks. However, since crowd work resembles a financially remunerated form of digital employment on a full- or part-time basis [1], monetary rewards seems to have a major effect on crowd workers` job satisfaction and intention to take over tasks in future.

Finally, our study contributes to the systematic design of tasks – a pivotal success factor for the application of crowd work [7]. Neither, the systematic provision of information about the task and its context, nor the flexibility of the work context had a positive impact on job satisfaction or continuance intention. As a consequence, the flexibility may rather reflect a hygiene factor mitigating dissatisfaction and central prerequisite for participation, than an antecedent of job satisfaction. Similarly, our results indicate that crowd workers` needs of information in testing projects play a tangential role. As a consequence, crowdsourcers may reduce the provision of information to a minimum that is required for solving a given task without having to fear negative consequences of unwanted information spillovers.

6.2 Practical Implications

The present study has significant practical implications. First, monetary reward tends to be the most important factor on job satisfaction and continuance intention. Although, compensation varies greatly the intermediaries should consider about adequate payment models to their crowd to ensure further participation. There are various ways to pay the crowd, inter alia, fixed fee or revenue participation in the subsequent product. In addition, the career prospects tend to be an important extrinsic motivation that affects the crowd workers' job satisfaction and continuance intention. Hence, it could be essential to the intermediaries to invest in the professional future of their crowd. Further education and career opportunities can be realized, inter alia, by training programs or recruitment options. There is practical evidence by platforms, which already offer their crowd career opportunities. After they have worked successfully on several projects, the crowd workers can qualify for positions in quality control or as a consultant on the platform. Hence, a sort of crowd-recruitment arises that deserves future attention in practice and should be institutionalized.

7 Conclusion

Given the lack of empirical research on crowd workers' perspective, our primary objective was to achieve a better understanding of the factors that affect crowd workers' job satisfaction. Based on prior literature, we develop and test a research model that has both theoretical and practical contributions. Theoretically, this study introduces motivational constructs and their impact on job satisfaction and further continuance intention. Practically, our study contributes several guidelines for crowdsourcing intermediaries, especially in the field of software testing. However, we also recognize that our research has some limitations since the survey exclusively focused only on the testIO platform, mainly on German crowd workers and include only one survey. Thus, these facts are limiting the generalizability of our results with regard to other nationalities, cultures and platforms.

In conclusion, our study considered crowd work to be a new facet of digital gainful employment that deserves future theoretical and practical attention.

References

1. Durward, D., Blohm, I., Leimeister, J.M.: Crowd Work. Business & Information Systems Engineering 58, 281-286 (2016)
2. Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. Proceedings of the 2013 conference on Computer supported cooperative work, pp. 1301-1318. ACM (2013)
3. Kuek, S.C., Paradi-Guilford, C., Fayomi, T., Imaizumi, S., Ipeirotis, P., Pina, P., Singh, M.: The Global Opportunity in Online Outsourcing. (2015)
4. Schenk, E., Guittard, C.: Towards a Characterization of Crowdsourcing Practices Journal of Innovation Economics, 93-107 (2011)
5. Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H.: Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition. Journal of Management Information Systems 26, 197-224 (2009)
6. Deng, X., Joshi, K.: Why Individuals Participate in Micro-task Crowdsourcing Work Environment: Revealing Crowdworkers' Perceptions. Journal of the Association for Information Systems 17, 3 (2016)
7. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Quarterly Executive 4, 199-211 (2013)
8. Kaufmann, N., Schulze, T., Veit, D.: More than fun and money. worker motivation in crowdsourcing–a study on mechanical turk. (2011)
9. Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., Vukovic, M.: An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. ICWSM 11, 17-21 (2011)
10. Spector, P.E.: Measurement of human service staff satisfaction: Development of the job satisfaction survey. American Journal of Community Psychology 13, 693-713 (1985)
11. McKeen, J.D., Guimaraes, T., Wetherbe, J.C.: The relationship between user participation and user satisfaction: an investigation of four contingency factors. Mis Quarterly 427-451 (1994)
12. DeLone, W.H., McLean, E.R.: Information systems success: the quest for the dependent variable. Information systems research 3, 60-95 (1992)
13. Elliott, E.S., Dweck, C.S.: Goals: an approach to motivation and achievement. Journal of personality and social psychology 54, 5 (1988)
14. Reinig, B.A.: Toward an understanding of satisfaction with the process and outcomes of teamwork. Journal of Management Information Systems 19, 65-84 (2003)
15. Bhattacherjee, A.: Understanding information systems continuance: An expectation–confirmation model. Mis Quart 25, 351–370 (2001)
16. Sun, Y., Wang, N., Peng, Z.: Working for one penny: Understanding why people would like to participate in online tasks with low payment. Computers in Human Behavior 27, 1033-1041 (2011)
17. lves, B., Olson, M., Baroudi, J.J.: The measurement of user information satisfaction. Communications of the ACM 26, 785-793 (1983)
18. Ang, J., Soh, R.H.: User information satisfaction, job satisfaction and computer background: An exploratory study. Information & Management 32, 255–266 (1997)
19. Lee, S.Y., Brand, J.L.: Effects of control over office workspace on perceptions of the work environment and work outcomes. Journal of Environmental Psychology 25, 323–333 (2005)
20. Galinsky, E., Bond, J.T., Hill, E.J.: When work works: A status report on workplace flexibility. Families and Work Institute, New York (2004)

21. Brief, A.P., Aldag, R.J.: The Intrinsic Extrinsic Dichotomy: Toward Conceptual Clarity. Academy of Management Review 2, 496-500 (1977)
22. Van der Heijden, H.: User Acceptance of Hedonic Information Systems. Mis Quart 28, 695-704 (2004)
23. Sun, Y., Wang, N., Yin, C., Che, T.: Investigating the Non-Linear Relationships in the Expectancy Theory: The Case of Crowdsourcing Marketplace. AMCIS 2012 Proceedings. Paper 6., (2012)
24. Lakhani, K.R., Wolf, B.: Why Hackers Do What They Do. Understanding Motivation and Effort in Free/Open Source Software Projects. In: Feller, J., Fitzgerald, B., Hissam, S., Lakhani, K.R. (eds.) Perspectives on Free and Open Source Software. The MIT Press, Cambridge, MA (2005)
25. Staw, B.M.: Intrinsic and extrinsic motivation. General learning Press, Morristown, N.J. (1976)
26. Spector, P.E.: Job satisfaction: application, assessment, causes and consequences. Sage, Thousand Oaks (1997)
27. Deng, X.N., Joshi, K.D.: Is Crowdsourcing a Source of Worker Empowerment or Exploitation? Understanding Crowd Workers' Perceptions of Crowdsourcing Career. Thirty Fourth International Conference on Information Systems, Milan 2013, (2013)
28. Ke, W., Zhang, P.: The Effects of Extrinsic Motivations and Satisfaction in Open Source Software Development. Journal of the Association for Information Systems 11, 784-808 (2011)
29. Clary, E.G., Snyder, M., Ridge, R.D., Copeland, J., Stukas, A.A., Haugen, J., Miene, P.: Understanding and assessing the motivations of volunteers: A functional approach. Journal of Personality and Social Psychology 74, 1516-1530 (1998)
30. Borg, I.: Führungsinstrument Mitarbeiterbefragung: Theorien, Tools und Praxiserfahrungen. Hogrefe, Göttingen (2003)
31. Rimann, M., Udris, I.: Fragebogen – Salutogenetische Subjektive Arbeitsanalyse (SALSA): Teil B und D von: Fragebogen zur persönlichen Situation im Beruf und Betrieb. Eidgenössische Technische Hochschule Zürich, Zürich (1997)
32. Lim, S.: Job satisfaction of information technology workers in academic libraries. Library & Information Science Research 115-121, 2 (2008)
33. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate data analysis. Prentice hall Upper Saddle River, NJ (1998)
34. Bagozzi, R.P., Yi, Y.: On the Evaluation of Structural Equatation Models. J. Acad. Marketing Sci. 16, 74-94 (1988)
35. Kline, P.: An easy guide to factor analysis. Routledge (2014)
36. Fornell, C., Larcker, D.F.: Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. J. Marketing Res. 18, 39-50 (1981)
37. Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P.: Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. J. Appl. Psych. 88, 879-903 (2003)
38. Pavlou, P.A., Huigang, L., Yajiong, X.: Understanding and mitigatin uncertainty in online exchange relationships: a pricipal-agent perspective. MIS Quart. 31, 105-136 (2007)
39. Lindell, M.K., Whitney, D.J.: Accounting for common method variance in cross-sectional research designs. J. Appl. Psych. 86, 114 (2001)
40. Gefen, D., Rigdon, E.E., Straub, D.: An Update and Extension to SEM Guidelines for Administrative and Social Science Research. MIS Quart. 35, iii-A7 (2011)

Supporting Creative Processes with IT: A literature review

Kai Klinker, Manuel Wiesche and Helmut Krcmar

Technical University of Munich, Chair for Information Systems, Munich, Germany
{kai.klinker ,manuel.wiesche, krcmar}@in.tum.de

Abstract. Over the last decades, advances in IT have changed markets and businesses. Companies that were unable to adopt have vanished. In most businesses, the best strategy for keeping a competitive advantage is to innovate at a faster pace than rival organizations. This requires a systematic approach to innovation that is tailored to the individual situation of a company. In order to improve the speed and quality of their innovation process on an organizational level, companies can incorporate IT solutions such as electronic brainstorming that support the creativity of their employees on a group and individual level. The aim of this study is to investigate the current state of research in the IS literature regarding technology supported creativity. We have conceptualized the topic into individual research strands. For each research strand, we present theoretical models, their alignment with empirical findings and an agenda for future research.

Keywords: creativity, support systems, literature review, creativity techniques

1 Introduction

The starting point of every innovation is an idea. However, initial ideas rarely emerge fully formed from an innovative employee's head. Usually it needs to be shaped and modified significantly in order to be fleshed out into a successful business plan [1]. Therefore, working in a team on challenges with no obvious solution is very common in organizations that innovate. Ideas about products, practices, services, or procedures are considered creative when they are (a) novel or original and (b) potentially useful to the organization [2].

Many innovation methods and frameworks such as Design Thinking or Lean Startup promote interdisciplinary team collaboration and have become popular over the last years [3]. They make use of creativity techniques such as brainstorming or De Bono's six thinking hats. However, most of these approaches rely on hand written interactions techniques such as post-its and collaborating in face to face (f2f) sessions and make little use of technology for communication or stimulating group creativity.

Group work on topics with high uncertainty can be classified into divergent phases, in which ideas are generated and convergent phases that are focused on making

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

decisions. Typically, divergent and convergent phases are used iteratively to first generate and then select the best ideas [4].

Many information systems exist that hold the potential for supporting groups in these phases, such as group support systems (GSS), creativity support systems (CSS), decision support systems (DSS) or knowledge management systems (KMS) [5–8].

Although many such tools are already available on the market, a significant gap still exists between those products and the creative process that they are supposed to support [9]. Therefore, Gabriel et al call for a an re-examination of the entire creative process to understand and fulfill real needs [10].

In order to start this endeavor we have conducted a literature review focused on theoretical and empirical findings within this field. Our research contributes to theory by providing an overview of existing theoretical knowledge on supporting creativity with technology in form of a conceptual framework. Furthermore, we identify several research gaps and lay out a research agenda for future research.

Our literature review is structured in the following manner: Within the next section, we describe the design of our literature review, focusing on selection of appropriate journals, conferences and articles. Then we describe our procedure of coding the selected articles and identification of main research themes. Next, we provide an overview of each theme and develop a conceptual framework integrating the themes. Finally, our review gives a short conclusion and implications for future work.

2 Design of the literature review

The intention of this study is to investigate the current state of research in the IS literature regarding technology for support of creative work and to identify research gaps.

Table 1. Review Scope

Characteristic		*Categories*			
1	focus	research outcomes	research methods	theories	applications
2	goal	integration	criticism	central issue	
3	organization	historical	conceptual	methodological	
4	perspective	neutral representation	espousal of position		
5	audience	specialized scholars	general scholars	practictioners	general public
6	coverage	exhaustive	exhaustive & selective	representative	central/pivotal

We therefore conducted a systematic literature review following the five steps of vom Brocke et al. [11]: Definition of the Review Scope (1), Conceptualization of the Topic

(2), Literature Search (3), Literature Analysis and Synthesis (4) and Derivation of a Research Agenda (5). To increase the comprehensiveness of our literature review, we discuss the Research Agenda directly within each result section as described by Müller-Bloch et al. [12].

In the first step, we clarified the scope (1) of our research according to Cooper's [13] taxonomy (depicted in Table 1): We decided to carry out a conceptual literature review, taking a neutral perspective, in order to integrate research outcomes and theories of how technology can support creativity by reviewing journals from high quality IS journals. The paper is intended for specialized scholars and may be considered representative for the IS domain.

Secondly, we conceptualized the topic (2) following Rowley and Slack's method [14]. We started by conducting a brief search for relevant literature in leading journals and books. Therefore, we read several books about creativity techniques and related topics such as design thinking, lean startup and innovation management to get a general understanding of the topic and to derive relevant search terms. Next, we started conceptualizing the topic and decided which concepts from it were relevant to our research topic. The result of the conceptualization is our conceptual framework, which is described in-depth in the next section. Throughout the process of forming the conceptual framework, we iteratively reviewed literature and developed a list of search terms we found relevant to our topic.

Thirdly, we conducted the literature review (3) as described by vom Brocke [11]. We chose to include only articles from peer-reviewed high-quality IS-journals. Therefore we searched in all journals ranked B or higher in the vhb-jourqual3. In order to find relevant articles we used the Scopus database. We limited the search to the ISSNs of the selected journals and included all articles that contained one of following search terms in the title or abstract:

"triz", "brainstorming", "creativity technique", "creativity techniques", "disney method", "six thinking hats", "creativity support system", "brain writing", "creative problem solving", "group decision support system", "collaborative creativity", "mind mapping", "quality function development", "axiomatic design", "lateral thinking", ”kansei engineering", "open innovation", "group support system”.

In order to find current research gaps, we limited the search to articles published from the 1st of January 2006 until the 2nd of April 2017. Our search yielded 140 hits. Next we determined which articles were relevant to our review. We only included articles that reported research on a) how a group of people could be supported in the process of generating, refining or documenting ideas by the use of information technology. Moreover, the reported research needed to be b) focused on the individual or group level and not on an organizational level. Finally, c) we excluded articles that were limited to an application domain and thus not appropriate to contribute to theory. We screened title, abstract and if necessary full text of the articles in order to identify all articles relevant to our topic. As a result of the screening for relevance 36 articles remained in the further process of the literature review.

3 Findings

In this section we describe our conceptual framework (Table 2) and discuss existing research, theoretical knowledge and research gaps within each theme.

Table 2. Conceptual Framework

	Concept					***Type***	
Citation	***Idea Evaluation***	***Creativity Interventions***	***Communica-tion***	***Knowledge***	***Anonymity***	***Converging***	***Diverging***
(Ackermann et al. 2016) [15]					x	x	
(Althuizen/Reichel 2016) [8]		x		x			x
(Wong et al. 2016) [16]			x			x	x
(Tegarden et al. 2016) [17]			x		x	x	x
(Figl/Recker 2016) [18]				x			x
(Álvarez Carrillo et al. 2015) [6]			x			x	
(Ducassé/Cellier 2014) [19]				x		x	
(Jongsawat/Premchaiswadi 2014)			x				x
(Chen et al. 2014) [22]			x		x	x	
(Althuizen/Wierenga 2014) [4]		x		x		x	x
(Eden/Ackermann 2014) [23]					x	x	x
(Marett/George 2013) [24]			x			x	
(Dennis et al. 2013) [25]		x					x
(Bartelt et al. 2013) [26]		x				x	
(Javadi et al. 2013) [27]				x			x
(Kolfschoten/Brazier 2013) [28]				x		x	
(Reinig/Briggs 2013) [29]	x			x	x		x
(Voigt et al. 2013) [30]		x				x	x
(Müller-Wienbergen et al. 2011) [31]				x		x	x
(Ackermann/Eden 2011) [32]			x		x	x	x
(Ferreira et al. 2011) [33]				x			x
(Kuo/Yin 2011) [7]			x		x	x	x
(Barkhi/Kao 2011) [5]		x				x	
(Haines/Cheney Mann 2011) [34]			x		x	x	
(Alnuaimi et al. 2010) [35]			x				x
(Barkhi/Kao 2010) [36]			x			x	
(Briggs/Reinig 2010) [37]	x	x	x		x		x
(Paul/Nazareth 2010) [38]			x	x		x	
(Hahn/Wang 2009) [39]				x		x	x
(Cooper/Haines 2008) [40]			x		x	x	
(Lim/Guo 2008) [41]					x	x	
(Reinig/Briggs 2008) [42]	x						x
(Chen et al. 2007) [43]		x				x	x
(Reinig et al. 2007) [44]	x						x
(Shirani 2006) [45]			x	x		x	
(Heninger et al. 2006) [46]			x				
Σ	4	8	16	12	11	24	21

3.1 Evaluation of ideas

Since our study focuses on creative problem solving contexts with high uncertainty, the main unit of analysis for group and individual performance is the ideas that emerge from a creative problem solving session. Several metrics for idea evaluation are in use, such as idea-count, sum-of-quality, average-quality, and good-idea-count [29]. Researchers use these metrics to evaluate the efficacy of interventions targeted at improving the creativity of proposed solutions to a given problem. Despite the central role idea evaluation methods pose for creativity research, we identified very little discussion on the validity of the mentioned metrics. Only four studies evaluated the quality of idea evaluation methods and were all written by Reinig & Briggs. Moreover, their research is focused on ideas that stem from diverging phases.

Throughout their research, Reinig & Briggs theorize that only idea quality is a reliable measure to evaluate ideas. Moreover, they find that idea quantity is a poor surrogate for idea quality [29]. They also present theoretical reasoning which questions Osborn's conjecture that quantity of ideas will lead to high quality ideas [37]. They argue, that there may only be a limited amount of good solutions to a problem (limited solution space) and that humans are subject to cognitive inertia. Cognitive inertia occurs when people focus on a subset of concepts and have difficulties activating additional concepts from their long-term memory, making new ideas more and more similar to previous ideas.

Discussion and future research

So far, no standard metric has emerged to evaluate the outcome of creative problem solving sessions. The studies in our literature review used a variety of dependent variables, making it difficult to compare their findings. Quantity of ideas is far easier to measure than idea quality. Therefore, many studies include it as a dependent variable. Reinig & Briggs and Osborn provide contradictory theoretical arguments on the validity of idea quantity as a quality measure, whereas idea quality is an accepted measure within the research community. Empirical future research is needed to answer, whether quantity leads to quality. Another area for future research is to develop more efficient and robust methods to evaluate idea quality and decision quality.

3.2 Creativity improving interventions

In total, we found eight studies that tested interventions, which influenced the creativity of groups or individuals. The main contributors to theory on this topic are Reinig & Briggs and Althuizen et al. [4, 8, 29, 37, 42, 44, 47].

On a group level, Reinig & Briggs propose their own theory called Bounded Ideation Theory (BIT) [37]. BIT states that ideation team member's ability, understanding of the task, scarcity of attention resources, mental and physical exhaustion, goal congruence and the openness of the solution space predict the ratio of good ideas to total ideas. Findings from Barkhi et al.'s support BIT's proposition that understanding of the task improves group decision-making performance. On the other hand, they found that psychological safety leads to better performance, which is not modeled by BIT [5].

Althuizen et al. focus more on how the creativity of individuals can be increased [4, 8]. They found stimuli providers helped individuals create more novel and useful ideas than mind mappers or process guides [8]. To explicate their findings they draw upon the dual pathway to creativity theory [8]. This theory posits, that in order to find more novel solutions to a problem, individuals need to be stimulated either to be more persistent in the exploration of their knowledge base (persistence) or to search within a broader range of categories (flexibility).

However, creativity support systems are only effective if they take the creative ability of the individual into account and provide a sufficiently large and diverse set of cases that are closely linked to the problem at hand [8]. Moreover, creativity support systems are not likely to help highly creative individuals [4]. Achievement priming and creative problem solving training was successfully employed to generate a higher quantity and quality of ideas of groups [26, 43], whereas priming techniques which were known to be effective for idea generation were not effective in the same form for decision making [26].

Discussion and future research

Important topics for future research would be to explicate what factors influence the creative performance on a group level. BIT proposes several moderating constructs. An empirical measurement of their effect size could help to direct future research. Moreover, research is needed to explain how priming, creativity techniques and stimuli providers on the individual level effect outcomes on a team level. Further investigation of the optimal level of remoteness of stimuli to the target problem could help to effectively stimulate creativity. Finally, future research could develop and test other priming techniques to achieve better outcomes in decision processes.

3.3 Group communication

We differentiate three types of communication that can occur in the context of technology supported creativity: Face to face (f2f), distributed-synchronous and asynchronous communication [22]. In total, sixteen studies investigated the influence of communication on the creative process.

Asynchronous communication helps users to respond in a more measured and thoughtful way if they have time to process statements [15]. Users typically perform the same tasks in asynchronous communication as in f2f meetings. E.g. people that usually summarized contributions of others continue to do so in Computer Mediated Communication (CMC) [32]. Furthermore, individuals that usually dominated f2f meetings through rhetoric and charisma have less influence in CMC.

Text-based asynchronous interaction benefits those that have high levels of computer literacy [21, 32]. Furthermore, participants in synchronous text-based meetings typically wish to express their own point of view before focusing on the contributions of other group members [32]. Moreover, individuals often fail to process information they receive in text communication while concentrating on contributing to the discussion, due to an effect called dual-task interference [7, 46]. The resulting text from asynchronous communication can provide a group memory that can be revisited at a later point in time [21]. The presence of such documentation can lead to enhanced decision making, project awareness, collaboration and decision consent [22].

F2f groups tend to discuss information that is known to all members of the group before focusing on unshared information, whereas unshared information is discussed earlier in asynchronous text-based communication [45]. F2f communication via video conferencing or co-location is more efficient than distributed asynchronous communication for decision making [6, 21, 22, 36, 48, 49]. Moreover, individual decision-making performance depends on individual group member's perceived level of psychological safety, psychological meaningfulness and their understanding of the decision goal [5].

The motives underlying individual group member behaviors (insight awareness) are key to increasing decision quality and consensus. Greater insight awareness is obtained when individuals are able to track and characterize other individual's behaviors (behavior awareness). Behavior awareness depends on an individual's ability to identify and distinguish among the different individuals within the group [34, 40]. On the other hand, behavior awareness increases individual's conformity and perception of dispensability. This in turn decreases participation and perceived consensus in teams [34]. Interestingly, these findings seem to align with Alnunaimi's theoretical model of social loafing [35]. Furthermore, Jongsawat & Premchaiswadi found that group awareness information had a positive influence on the work effort on a given task and the quality of collaborative work [21].

Lies and deceptive statements occur more often in (CMC) than in f2f meetings. However, lies and deceptions have significantly more success in influencing group decisions in f2f meetings [24]. Finally, social loafing and free riding occur more often in CMC than in f2f meetings and hinder team productivity [35, 37].

Several theories have been used to explicate findings on group communications. Focus Theory assumes that communication, information access and deliberation are needed for groups to be productive. If individuals need to allocate more attention resources to one of the three processes, the other two are neglected [22]. This theory would explain the dual task-interference described by Heninger et al. Alnuaimi et al. build on the theory of moral disengagement to explain social loafing and its effects in on team productivity [35]. They found that social loafing occurs more often in big and dispersed teams. In bigger teams members feel less responsible for outcomes, attribute blame on other team members more often and dehumanize team members. Team dispersion also strongly contributes to the dehumanization of team members. The result is increased social loafing and less team productivity.

Discussion and future research

Future research could develop frameworks that help decide under which circumstances which type of communication is optimal for group creativity. Moreover, research on how dehumanization, diffusion of responsibility and attribution of blame in large distributed teams can be prevented is needed. Since dual-task interference during group communication negatively influences decision quality, future research could develop and test counteractive measures and research if dual-task interference also occurs in diverging phases. Many articles mention the advantages the use of mobile ICT could have for group communication [16]. However, none of the articles in our study report scenarios in which mobile ICT has been tested for creative problem solving.

There has been very little research focused on the behavior and speech of participants in technology supported creative processes. Interesting topics for research could include the analysis of cognitive load, stress level or body language. Knowing under which circumstances individuals experience cognitive load could help to empirically evaluate theories such as BIT, Focus Theory or idea integration.

Finally, since the findings of Haines et al. on behavioral awareness resemble those of Alnuaimi's findings on moral disengagement and social loafing future research could investigate whether those two theories are taking two perspectives on the same phenomenon.

3.4 The role of knowledge

We found twelve studies that discussed the role or representation of knowledge in technology supported creative processes. According to the dual pathway to creativity theory, creativity is tightly intertwined with knowledge, since ideas stem from the knowledge base of individuals [8, 31]. Therefore, measures that support individuals to search their knowledge base can be helpful. However, providing knowledge that is too remote to the problem at hand may be counterproductive [4]. Furthermore, knowledge should be presented differently in convergent and divergent phases. In convergent phases Knowledge Management Systems (KMS) that connect information seekers and specific answer providers are more efficient, whereas in divergent phases it is more important to bring together individuals with similar interests [39]. Moreover, when large amounts of information need to be processed, the team perceives time pressure, the information is complex or multiple criteria need to be considered, information overload can occur during decision making [19, 28, 38]. Regarding the representation of knowledge, Figl & Recker found that visual models helped participants to generate more appropriate ideas than text [18].

Information overload can also occur during idea integration in diverging phases when ideas need to be clustered or combined. Idea integration and existing knowledge can also limit the original creativity due to cognitive bias [27, 29, 31, 33]. On the other hand idea integration is a central mechanism for the inclusion of knowledge from other team members [27].

Discussion and future research

Future research could investigate methods of reducing cognitive load while searching the own knowledge base or integrating knowledge from other sources. Moreover, since too much knowledge can also have a detrimental effect on creativity, further investigation of the right amount and scope of knowledge necessary for sparking creative solutions is important.

3.5 Role of Anonymity

We found eleven studies that theorized on the use of anonymity in creative processes supported by ICT. Anonymity generally separates a person from their contribution, such that contributions can be evaluated by their merit and not by who voiced them [15, 32, 41]. Moreover, anonymity allows things to be said that could not be said otherwise [15], leads to blunter statements [23] and more controversial ideas like "tattoo AIDS victims to ID them" [29]. Anonymous brainstorming groups generate a larger quantity of unique ideas but of the same quality as non-anonymous ones [29].

Furthermore, Briggs & Reinig found that anonymity decreases evaluation apprehension [37]. Evaluation apprehension is the fear of receiving retribution from peers or superiors for submitting an unpopular idea. Anonymity also reduces the uncertainty of minorities and helps them voice unpopular opinions [41]. Therefore, it can be used to flatten hierarchies and to develop ideas in a non-power based manner, which also works for a more objective evaluation of ideas [17].

Finally, anonymity can be provided with or without the use of nicknames. Not being able to differentiate which individual voiced a comment can be confusing for users [22, 34] and can lead to coordination problems during consensus tasks, because users do not know who holds which opinion. This effect decreases group influence, whereas when nicknames are provided and users can identify which individual holds which opinion, the group influence is increased [34].

Discussion and future research

The advantages and limitations of anonymity and the different forms of group communication have been researched to a certain degree. However, no theory has been tested or developed to explain the findings. A framework that explains under which circumstances anonymity is beneficial would be helpful.

4 Conclusion

In this article, we investigated the current state of the scientific literature about technology for supporting creative processes in the IS literature. We find that the IS field has focused on supporting creativity via group support systems, creativity support systems, decision support systems and knowledge management systems.

We have identified several research strands in the literature about these systems: Quick and correct assessment of ideas, measures that improve the creativity of groups and individuals, the influence of the different types of group communication, the role of knowledge and how it is presented to groups and individuals, the behavior of team members and the effects of anonymous communication. Future work in the field of technology supported creativity could target the research gaps we have identified within the individual research strands.

Even though an abundance creativity techniques such as de bono's six thinking hats, mind mapping or TRIZ have been used for more than 20 years, very few of the articles in our literature review reported scenarios in which technology supported versions of such techniques were employed. Brainstorming seems to be the only creativity technique of major interest in the IS literature.

Finally, research on creativity is vast and scattered among many disciplines such as neuroscience, cognitive science, sociology or economics. Our study is only limited to what has been researched in the field of Information Systems. Comparing the findings from other disciplines with ours could lead to a fruitful discussion.

5 Acknowledgements

This research and development project is / was funded by the German Federal Ministry of Education and Research (BMBF) within the Program "Innovations for Tomorrow's Production, Services, and Work" (02K14A080) and managed by the Project Management Agency Karlsruhe (PTKA). The author is responsible for the contents of this publication.

References

1. Christensen, C.M., Johnson, M.W., Rigby, D.K.: Foundations for Growth: How To Identify and Build Disruptive New Businesses. MIT Sloan Manag. Rev. 43, 22–31 (2002).
2. Oldham, G.R., Da Silva, N.: The impact of digital technology on the generation and implementation of creative ideas in the workplace. Comput. Human Behav. 42, 5–11 (2015).
3. On, S., Evolution, T.H.E., Design, O.F.: Design Thinking Comes of Age Design Thinking Comes of Age. (2015).
4. Althuizen, N., Wierenga, B.: Supporting Creative Problem Solving with a Case-Based Reasoning System. J. Manag. Inf. Syst. 31, 309–340 (2014).
5. Barkhi, R., Kao, Y.-C.: Psychological climate and decision-making performance in a GDSS context. Inf. Manag. 48, 125–134 (2011).
6. Álvarez Carrillo, P.A., Leyva López, J.C., Sánchez Castañeda, M.D.L.D.: An Empirical Study of the Consequences of Coordination Modes on Supporting Multicriteria Group Decision Aid Methodologies. J. Decis. Syst. 24, 383–405 (2015).
7. Kuo, F.-Y., Yin, C.-P.: A linguistic analysis of group support systems interactions for uncovering social realities of organizations. ACM Trans. Manag. Inf. Syst. 2, (2011).
8. Althuizen, N., Reichel, A.: The Effects of IT-Enabled Cognitive Stimulation Tools on Creative Problem Solving: A Dual Pathway to Creativity. J. Manag. Inf. Syst. 33, 11–44 (2016).
9. Fries, V., Pflügler, C., Wiesche, M., Krcmar, H.: The hateful six – factors hindering adoption at small and medium sized enterprises. Twenty-second Am. Conf. Inf. Syst. (2016).
10. Gabriel, A., Monticolo, D., Camargo, M., Bourgault, M.: Creativity support systems: A systematic mapping study. Think. Ski. Creat. 21, 109–122 (2016).
11. von Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., Brocke, J. Von, Reimer, K.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. 17th Eur. Conf. Inf. Syst. 9, 2206–2217 (2009).
12. Müller-Bloch, C., Kranz, J.: A Framework for Rigorously Identifying Research Gaps in Qualitative Literature Reviews. Thirty Sixth Int. Conf. Inf. Syst. (ICIS 2015). 1–19

(2015).
13. Cooper, H.M.: Organizing knowledge synthesis: a taxonomy of literature reviews. Knowl. Soc. 1, 104–126 (1988).
14. Rowley, J., Stack, F.: Conducting a Literature Review. Manag. Res. News. 27, 31–39 (2004).
15. Ackermann, F., Eden, C., Pyrko, I.: Accelerated Multi-Organization Conflict Resolution. Gr. Decis. Negot. 25, 901–922 (2016).
16. Wong, T.Y.T., Peko, G., Sundaram, D., Piramuthu, S.: Mobile environments and innovation co-creation processes {&} ecosystems. Inf. Manag. 53, 336–344 (2016).
17. Tegarden, D., Tegarden, L., Smith, W., Sheetz, S.: De-Fusing Organizational Power Using Anonymity and Cognitive Factions in a Participative Strategic Planning Setting. Gr. Decis. Negot. 25, 1–29 (2016).
18. Figl, K., Recker, J.: Process innovation as creative problem solving: An experimental study of textual descriptions and diagrams. Inf. Manag. 53, 767–786 (2016).
19. Ducassé, M., Cellier, P.: Fair and Fast Convergence on Islands of Agreement in Multicriteria Group Decision Making by Logical Navigation. Gr. Decis. Negot. 23, 673–694 (2014).
20. Jongsawat, N., Premchaiswadi, W.: A Study Towards Improving Web-Based Collaboration Through Availability of Group Awareness Information. Gr. Decis. Negot. 23, 819–845 (2014).
21. Jongsawat, N., Premchaiswadi, W.: A Study Towards Improving Web-Based Collaboration Through Availability of Group Awareness Information. Gr. Decis. Negot. 23, 819–845 (2014).
22. Chen, F., Nunamaker Jr., J.F., Briggs, R.O., Corbitt, G., Sager, J., Gardiner, S.C.: An Application of Focus Theory to Project Management Processes. Gr. Decis. Negot. 23, 961–978 (2014).
23. Eden, C., Ackermann, F.: "Joined-Up" Policy-Making: Group Decision and Negotiation Practice. Gr. Decis. Negot. 23, 1385–1401 (2014).
24. Marett, K., George, J.F.: Barriers to Deceiving Other Group Members in Virtual Settings. Gr. Decis. Negot. 22, 89–115 (2013).
25. Lee, A.S., Dennis, A.R.: A hermeneutic interpretation of a controlled laboratory experiment: A case study of decision-making with a group support system. Inf. Syst. J. 22, 3–27 (2012).
26. Bartelt, V.L., Dennis, A.R., Yuan, L., Barlow, J.B.: Individual Priming in Virtual Team Decision-Making. Gr. Decis. Negot. 22, 873–896 (2013).
27. Javadi, E., Gebauer, J., Mahoney, J.: The impact of user interface design on idea integration in electronic brainstorming: An attention-based view. J. Assoc. Inf. Syst. 14, 1–21 (2013).
28. Kolfschoten, G.L., Brazier, F.M.T.: Cognitive Load in Collaboration: Convergence. Gr. Decis. Negot. 22, 975–996 (2013).
29. Reinig, B.A., Briggs, R.O.: Putting Quality First in Ideation Research. Gr. Decis. Negot. 22, 943–973 (2013).
30. Voigt, M., Bergener, K., Becker, J.: Comprehensive Support for Creativity-Intensive Processes - An Explanatory Information System Design Theory [Ganzheitliche Unterstützung für kreativitätsintensive Prozesse - Eine Informationssystem-

Designtheorie]. Wirtschaftsinformatik. 1–18 (2013).
31. Müller-Wienbergen, F., Müller, O., Seidel, S., Becker, J.: Leaving the Beaten Tracks in Creative Work – A Design Theory for Systems that Support Convergent and Divergent Thinking. J. Assoc. Inf. Syst. 12, 714–740 (2011).
32. Ackermann, F., Eden, C.: Negotiation in Strategy Making Teams: Group Support Systems and the Process of Cognitive Change. Gr. Decis. Negot. 20, 293–314 (2011).
33. Ferreira, A., Antunes, P., Herskovic, V.: Improving Group Attention: An Experiment with Synchronous Brainstorming. Gr. Decis. Negot. 20, 643–666 (2011).
34. Haines, R., Cheney Mann, J.E.: A new perspective on de-individuation via computer-mediated communication. Eur. J. Inf. Syst. 20, 156–167 (2011).
35. Alnuaimi, O.A., Robert Jr., L.P., Maruping, L.M.: Team size, dispersion, and social loafing in technology-supported teams: A perspective on the theory of moral disengagement. J. Manag. Inf. Syst. 27, 203–230 (2010).
36. Barkhi, R., Kao, Y.-C.: Evaluating decision making performance in the GDSS environment using data envelopment analysis. Decis. Support Syst. 49, 162–174 (2010).
37. Briggs, R.O., Reinig, B. a.: Bounded Ideation Theory. J. Manag. Inf. Syst. 27, 123–144 (2010).
38. Paul, S., Nazareth, D.L.: Input information complexity, perceived time pressure, and information processing in GSS-based work groups: An experimental investigation using a decision schema to alleviate information overload conditions. Decis. Support Syst. 49, 31–40 (2010).
39. Hahn, J., Wang, T.: Knowledge management systems and organizational knowledge processing challenges: A field experiment. Decis. Support Syst. 47, 332–342 (2009).
40. Cooper, R.B., Haines, R.: The Influence of workspace awareness on group intellective decision effectiveness. Eur. J. Inf. Syst. 17, 631–648 (2008).
41. Lim, J., Guo, X.: A study of group support systems and the intergroup setting. Decis. Support Syst. 45, 452–460 (2008).
42. Reinig, B.A., Briggs, R.O.: On the relationship between idea-quantity and idea-quality during ideation. Gr. Decis. Negot. 17, 403–420 (2008).
43. Chen, M., Liou, Y., Wang, C.-W., Fan, Y.-W., Chi, Y.-P.J.: TeamSpirit: Design, implementation, and evaluation of a Web-based group decision support system. Decis. Support Syst. 43, 1186–1202 (2007).
44. Reinig, B.A., Briggs, R.O., Nunamaker Jr., J.F.: On the measurement of ideation quality. J. Manag. Inf. Syst. 23, 143–161 (2007).
45. Shirani, A.I.: Sampling and pooling of decision-relevant information: Comparing the efficiency of face-to-face and GSS supported groups. Inf. Manag. 43, 521–529 (2006).
46. Heninger, W.G., Dennis, A.R., Hilmer, K.M.: Individual cognition and dual-task interference in group support systems. Inf. Syst. Res. 17, 415–424 (2006).
47. Reinig, B.A., Briggs, R.O., Nunamaker, J.F.: On the Measurement of Ideation Quality. J. Manag. Inf. Syst. 23, 143–161 (2014).
48. Wiesche, M., Keskinov, H., Schermann, M., Krcmar, H.: Classifying Information Systems Risks : What Have We Learned So Far ? 46th Hawaii Int. Conf. Syst. Sci. (HICSS 2013). (2013).
49. Pflügler, C., Wiesche, M., Krcmar, H.: Subgroups in Agile and Traditional IT Project Teams. 51st Hawaii Int. Conf. Syst. Sci. (2018).

Social Collaboration Analytics for Enterprise Social Software: A Literature Review

Florian Schwade[1], Petra Schubert[1]

[1] University of Koblenz-Landau, Institute for IS Research, Koblenz, Germany
{fschwade,petra.schubert}@uni-koblenz.de

Abstract. Over the last years, research on benefits and success measurement of Enterprise Social Software (ESS) has gained momentum. Literature reviews in this topic area discuss methods for measuring ESS success and benefits and demonstrate that Social Collaboration Analytics is gaining increasing importance in the context of benefits research for ESS. This paper provides an overview of the literature on Social Collaboration Analytics for ESS. The selected literature can be categorised in seven key themes: *measuring system usage*, *identification of usage patterns*, *identification of types of users*, *analysis of groups*, *identification of expertise*, *network analysis* and *organisational impacts*. The paper concludes with a discussion of the importance of Social Collaboration Analytics for other relevant research streams.

Keywords: Enterprise Social Software, Enterprise Social Network, Data Analytics, Social Collaboration Analytics.

1 Introduction

Over the last decade, we have witnessed the emergence of a new type of collaboration software, the so-called "Enterprise Social Software" (ESS). The use of Social Media in private life has changed the way people communicate and exchange information. Consequently, companies have been experimenting with the "social features" (e.g. social profiles or tagging/commenting) of this software type [1], which has stimulated a renewed interest in Enterprise Collaboration Systems (ECS) both in research and practice [2, 3].

The existing literature on Enterprise Social Software is extensive and focuses particularly on the benefits of this kind of software, e.g. [2, 4]. Existing studies suggest that measuring the usage of ESS is an important aspect of benefits measurement [5, 6]. Consequently, we observe an increasing number of studies relying on ESS data. Despite of many attempts to measure *usage* of ESS, research papers providing an overview on possible approaches to Social Collaboration Analytics are scarce [2]. In order to address the perceived lack of clear definition in the terminology, we are providing a description of our basic terminology on *Social Collaboration Analytics* (SCA) [7] in chapter 2.

This paper reports on the results of a structured literature review on SCA. The literature review aims to identify, synthesize and conceptualise existing research related to SCA. The main objective is to (1) identify and document the *status quo of research*

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

in this topic area. Further objectives are (2) to identify the aspects of the analysis of *system usage* and (3) to identify the *measurement approaches* (e.g. applied methods and metrics). Thus, this literature review complements existing literature reviews in the field of success measurement, e.g. [4].

The paper is organised as follows: we begin with a definition of the term *Social Collaboration Analytics*. This is followed by a description of the research design of the literature review. In chapter 4, we present the findings from our literature analysis. The paper concludes with a discussion on the current status quo in research on SCA and possible future research directions. The discussion is complemented with a note on the relevance of SCA in the context of research on Enterprise Social Software

2 Terminology

Schwade and Schubert [7] define Social Collaboration Analytics (SCA) as "the approach for analysing and displaying collaboration activity of users in socially-enabled collaboration systems". The relevant data for SCA are *organisational data* (e.g. organisational structure), *transactional data* (e.g. event records stored in databases or log files) and *content data* (user-generated content).

Throughout this paper, the term Enterprise Social Software is referred to as a collaboration software containing multiple social media features (e.g. social profile, recommend, comment, blogs, Wikis, etc.). In contrast to this, a single social software is a standalone application that is not (necessarily) integrated with other applications (e.g. an installation of WordPress). The term Enterprise Collaboration System (ECS) refers to "integrated systems which provide multiple applications (modules) under a uniform user interface" [7]. While Enterprise Social Networks (ESN) are structures between users in ESS, ESN *platforms* are Social Networking platforms deployed in an enterprise context [7].

3 Design of the Literature Review

The design of the literature review is guided by multiple academic sources [8–10]. Following these recommendations, five sequential and intertwined phases were developed. Figure 1 depicts the research design of the literature review.

The *definition of the review scope* is guided by the research objectives and thus the paper focusses on publications on *measuring ESS system usage*. The scope of the search process itself is described in a later section. The *conceptualisation of the topic* is derived from the objectives. Accordingly, a concept matrix as proposed by Webster and Watson [8] was developed. This concept matrix contains concepts and sub-themes that are relevant for the literature review and metadata of publications, the type of system that was investigated (e.g. ESS, ECS, ESN) and a preliminary classification of measurement concepts (e.g. applied methods and metrics, criteria such as user analysis, community analysis, network analysis).

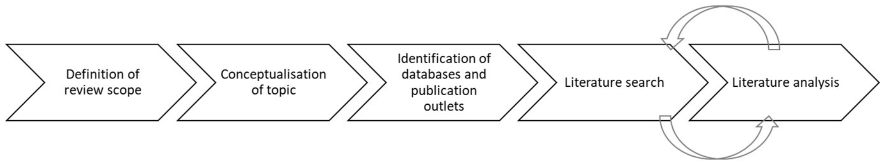


Figure 1. Research design of the literature review

Because the concept matrix was developed as an annotated bibliography in MS Excel, it allows a rigorous documentation of the whole literature review process. Thus, the concept matrix addresses the issue of recording and documenting the literature review. Following this, a keyword set was conceptualised which is shown in the following table.

Table 1. Overview of keywords and publication outlets

Keywords (technology)	ECS, Enterprise Collaboration System, Enterprise Social Network, ESN, {social, enterprise} collaboration, {collaboration, collaborative shared} workspace, enterprise 2.0, social software. Enterprise Social Software, Corporate Social Media, Corporate Social Software, groupware, cooperative system, BSCW, Groove, Notes, SharePoint, Connections
Keywords (analytics)	Metrics, measuring, measurement, measure, KPI, analytics, analysis, data, user {behaviour, interaction}, usage, usage analysis, activity, log, log file, database
Databases	Google Scholar, IEEE Explore, SpringerLink, ACM Digital Library, AIS eLibrary
Journals	EJIS, Information & Management, ISJ, Information Systems Management, ISR, IJeC, IJIM, JIS, JIT, MISQ, BISE, Journal of AIS, Journal of MIS, JSIS, Measuring Business Excellence, Computer Networks, Journal of Computer Mediated Communication
Authors	Nunamaker, Briggs, Grudin, Muller, Behrendt, Hacker/Viol, Richter
Date of publication	No restriction

Each search term consists of two parts. The first part refers to the technology and the second part refers to analytics. Each keyword from the technology branch was combined with all keywords from the analytics branch and Boolean operators were used where possible. This step is followed by the *identification of relevant databases and publication outlets*, which are listed in Table 1. The topic domain of SCA is an emerging topic. Therefore, the search process included conference proceedings as commonly suggested [8, 10].

The literature search and the literature analysis were designed as iterative activities. The *literature analysis* was conducted in three phases per publication. (1) A publication was added to the concept matrix if the title, abstract or keywords indicated a possible relevance. (2) The abstract was analysed in the second phase. The main criterion for classifying a publication as relevant for the literature review is that the abstract clearly points out that the publication addresses metrics, measurements or an analysis of ESS. (3) In the final phase, the publication was reviewed in detail according to the concept matrix. The final selection only contains publications, which describe the application of SCA, conceptual papers as well as papers performing content analysis of messages. Finally, a snowball search on the most relevant publications and an analysis of publication records of key authors in the field concluded the literature review.

4 Seven Key Themes for Social Collaboration Analytics

The following sections present an overview of the retrieved literature and the final conceptualisation of the identified topics. In total, 217 publications were retrieved of which 83 were included in a detailed analysis. The table below shows the number of publications per key theme. Adding up the individual counts does not lead to 83 because in some cases publications were assigned to multiple key themes.

Table 2. Number of publications per key theme

Key theme	*Number of publications*
Measuring system usage	30
Identification of usage patterns	7
Identification of types of users	10
Analysis of communities	7
Identification of expertise	6
Network analyses	14
Organisational impacts	13
Barriers to SCA	26

Besides the key themes, the literature review also identified several barriers to SCA. The identification of barriers to SCA was no intended objective for the literature review. However, as a considerable amount of publications mentions different kinds of barriers, barriers to SCA were also included as a review criterion. The iterative design of the literature review facilitated the inclusion of barriers to SCA as a further review criterion.

The measurement and analytics approaches in the literature were classified into seven key themes. The following sections contain the discussion of selected studies from each theme. Due to space limitations, this paper reports on the findings of selected key publications from the seven key themes.

4.1 Measuring system usage

There is a large volume of literature concerned with the *measurement of system usage*. What stands out from the literature on analysis of system usage is that a large proportion of studies is concerned with general usage measures and action specific usage measures. *General usage measures* provide a high-level overview on system usage. Shami et al. [11] provide usage statistics for ESS based on counting the number of events. Similarly, Behrendt et al. analyse the usage of communication features by counting the number of events for different communication features [12]. Herzog et al. [6] present an overview on methods and metrics (e.g. number of blog posts, and unique visitors), which can be used to measure ESS success. Following a similar idea, Steinhueser et al. [13] propose measuring ESS use by considering usage statistics, the degree of crosslinking, the number of hyperlinks, direct messages and comments.

In contrast to general usage measures, *action specific measures* contain some degree of aggregation and thus provide deeper indications on *how* a platform is used. Appelt [14] classifies user actions into different groups such as (1) creation of information, (2) modification of information or (3) reading of information. For each of the identified groups, the number of user actions is counted. As an addition to this, Jeners et al. [15] suggest investigating the most frequent activities such as reading, creating and changing content on a platform. In contrast to this, Muller et al. [16] differentiate between "producing" and "consuming" content. A further approach for measuring ESS usage is based on use cases. Richter et al. [5] identify seven collaborative actions (search, edit, rate, label, clarify, notify, share). For each of these collaborative actions, metrics are suggested. Similar metrics are proposed by Hacker et al. [17]. A more specialised aspect of measuring system usage is the investigation of *dynamics of ideas and innovations generation* [18]. An important observation can be made regarding the measurement of ESS usage *over time*. Only few authors include the aspect of time (e.g. measurements at multiple points in time) which would allow the tracking of ESS usage over time [5, 16, 17, 19]. The literature on measuring system usage also reveals that there are two basic intentions for conducting system analyses. In the more traditional research stream, analytics are conducted for deriving *design implications* [11, 16] (improving or changing the system). In recent years, a new research stream concerned with evaluation and success measurement [5, 6, 13] has emerged, with the objective of providing *decision makers* with information on platform usage and success.

4.2 Identification of usage patterns

Some of the publications included in our literature review discuss the concept of "usage patterns". The literature shows a variety of different approaches for the identification of usage patterns. Whereas Millen et al. [20] associate usage patterns with clickstreams in the context of *exploratory search patterns*, Ferron et al. [21] focus on patterns related to *communication and networking*. Muller et al. [22], on the other hand, refer to usage patterns as the *intensity and frequency of module use* in different communities. Boving and Simonsen [23] propose a more sequence-oriented understanding of usage patterns by arguing that usage patterns have to be viewed as a sequence of different (inter)actions. Chaves and Córdoba [24] and Naderipour [25] follow a more sophisticated approach by conducting "pattern analysis" based on *process minin*g. This approach is based on a process-based understanding of usage patterns.

We believe that research on ESS would benefit from a harmonised understanding of the term "usage patterns" and might be guided by process-oriented concepts such as *collaboration scenarios* proposed by Schubert and Glitsch [26] or an application of *process mining* as described by Chaves and Córdoba [24] and Naderipour [25]. The transfer of process mining to ESS could give rise to a new research area, which we termed *Social Process Mining.*

4.3 Identification of types of users

In recent years, there has been an increasing amount of literature concerned with the *identification and measurement of different types of users in ESS*. Studies in this area typically discuss (structural) characteristics of user types. Two types of approaches can be observed. User types are identified either based on intensity of system usage or based on the nature of usage. In early approaches, user activity is distinguished by platform access [14] or the frequency of activities [15] (*intensity of usage*). In such studies users are considered as active when they access the collaboration system on average at least on two days per week [14]. Frequently, lurkers are an object of study in ESS. *Lurkers* are typically defined as "a community member who has made zero visible contributions to the community" [22]. Consequently, lurkers are identified by analysing consuming activities. Complementary to lurkers, the user types *contributors* and *uploaders* are defined. The behaviour of contributors is characterised by creating metadata and commenting on existing content. Uploaders typically create new content and can also engage in contributing and consuming activities [27]. Accordingly, in analyses, these user types are distinguished by counting the amount of creating, contributing and consuming activities. In recent publications, researchers suggest identifying user types based on the *nature of system usage*. Based on a literature review on knowledge worker roles Hacker et al. [28] propose a knowledge worker role typology that is adapted for the context of ESS. The authors identify different knowledge worker roles such as Helper, Sharer or Seeker. Each of these roles is characterised by performing a combination of distinct "knowledge actions". Following the proposed typology, each of these roles can be determined by quantifying the knowledge actions. An alternative approach also suggested by Hacker et al. [29] focusses on *dimensions of user behaviour* (e.g. social dispersion, engagement, focus, information sharing, discussing).

4.4 Analysis of communities

A small body of literature pays particular attention to the analysis of communities or workspace groups. Jeners and Prinz [19] introduce *metrics for measuring productivity* (average items created per member), *activity* (average events per member) and *cooperativity* (average edits per member) of workspaces. Further metrics are suggested for the evaluation of shared workspaces. The *degree of labour division* describes the distribution of activities among the members of a workspace. Additionally, the *responsiveness* reveals how fast information in a workspace is accessed and it also shows which proportion of information is never accessed.

Researchers consistently argue that community mangers require meaningful reports for assessing the status and health of their community. Consequently, topics of interest are the assessment of *health and activity of communities* [30], *measurement of community success* [31] and the *comparison of key characteristics of different communities*) [32]. Typically, community analyses rely on similar metrics. *Participation metrics* include the overall activity of the platform, the number of posts, views and the types of contributions on the platform. *People metrics* reflect membership-changes, top contributors and the geographical location of users. *Content metrics* display popular topics and the value of single posts.

4.5 Identification of expertise

A further topic of interest is the identification and location of expertise in ESS. Nasirifard and Peristeras [33] demonstrate a two phase approach for the identification of expertise. First, documents are assigned with topics based on keyphrase extraction. Authors are assigned with the topics of the documents created by them. The second phase is the log file analysis. The authors assume that if a user creates or revises a document, there is a higher expertise on a topic compared to a read event. Therefore, Nasirifard and Peristeras distinguish between *expert in* and *familiar with*. John and Seligmann [34] discuss the potential of collaborative tagging and propose a mechanism for ranking expertise based on tags called ExpertRank. Based on the number of created bookmarks and the tags of these bookmarks, an "expertise rank" for this topic is assigned to the authors.

Considering that many surveys on the introduction of ESS [35] suggest that these systems are introduced with the intention to ease access to knowledge and experts, it is surprising that the number of publications on the identification of expertise is comparably small.

4.6 Network analysis

A popular theme related to SCA is the *analysis of networks* in Enterprise Social Software. Several publications propose *measures and characteristics for networks* and different *types of networks*. In an early study, Smith et al. [36] provide a typology for analysing networks in ESS. As underlying metrics, SNA metrics such as betweenness centrality, closeness centrality, eigenvector centrality, clustering coefficient are applied. Smith et al. [36] argue that that the analysis of networks can reveal data about social interactions that were previously invisible. This information can also be used to identify and foster important content and contributors. In contrast to this, Behrendt et al. [37] focus on communication networks. They analyse how fast new information travels across a network and how actions of participants affect the exchange of information. The authors describe different types of communication networks such as 1:1 communication, 1:n communication and m:n communication.

Researchers have also shown an increasing interest in the investigation of *connectedness of users*. Wu et al. [38] analyse the closeness of relationships between employees based on interactions in ESS by computing behavioural metrics for a subject's behavioural factors, friends' behavioural factors, subject-friend interaction factors, mutual connection factors and company direction factors. Steinhueser et al. [13] and Hacker et al. [17] suggest analysing the connectedness of users as an indicator of networking behaviour and general platform usage.

While the before mentioned researchers focus on analysing *user-centric* networks, Nasirifard et al. [39] suggest analysing *document-centric* networks. The authors argue that document-centric networks, which are constructed based on activities on documents, provide better insights on actual collaboration behaviour than user-centric networks.

4.7 Organisational impacts

Recent years have shown an increasing interest in measuring organisational impacts of and on ESS. A frequently addressed issue is the *impact of hierarchies on networking behaviour*. In this research stream, researchers are investigating how a user's position in the organisation influences activities in ESS. Behrendt et al. [40] examine the impact of formal organisational hierarchy on users' network positions in ESS and the creation of ties in ESS. In a similar study, Stieglitz et al. [41] examine the impact of hierarchy on ESN behaviour. With the influence of a user's position in the hierarchy (formal influence) and the influence of contributions (informal influence) on the response rate/time of questions, Riemer et al. [42] explore a closely related aspect. The three studies have in common that researchers work with two datasets. One dataset contains actions in the ESS, which can be statistically analysed and the second dataset contains information on the position of an employee in the hierarchy of the organisation.

The literature on organisational impacts reveals particular interest on the *effects of geographical dispersion* on user activity in ESS. Warshaw et al. [43] display them by determining the spatial distance between users, the timezone differences and the isolation of actors. These measures are then related to user activity. Similarly, Recker and Lekse [18] investigate whether the geographical distance between users influences intensity of communication and knowledge sharing. In close relation to geographical dispersion, recent years have shown an increasing interest in measuring and identifying intra-organisational *boundary spanning* activities in ESS [44, 45].

5 Conclusion and Directions for Future Research

Our literature review aimed to identify and classify studies in the field of Social Collaboration Analytics (i.e. analysis of the use of Enterprise Social Software). We identified seven different key themes: (1) *measuring system usage* (2) *identification of usage patterns,* (3) *identification of types of users,* (4) *analysis of communities,* (5) *identification of expertise,* (6) *network analysis and* (7*) organisational impacts*. The following sections discuss further findings related to methodology, limitations of current research and propositions for future research. Finally, we discuss Social Collaboration Analytics in context of existing ESS research.

The classification of the studies provided in this literature review shows that most studies are about aspects related to networking or general system usage. We observed a lack of research on measuring collaboration activities and patterns. A reasonable explanation for this is that both concepts draw on methods and metrics that are already established and matured in other contexts (especially Web Analytics and Social Network Analysis). For a better understanding of the usage of collaboration software, there is a pressing need for measuring and understanding collaboration activities.

The majority of researchers suggest applying mixed-methods approaches for measuring system usage [12]. Consequently, the majority of studies makes use of additional surveys or interviews complementary to analysing system data. The literature review further shows, that most authors adapt a case study design. A possible explanation is provided by Behrendt et al. [40] who argue that case studies allow to investigate user behaviour in a specific context without interfering the ESS. It can further be observed that studies on analytics in ESS often lack a structure and guidance of the analytics process. In most of the studies, the steps for collecting and preparing data are not described. The most likely cause of this is that researchers often do not have the possibility to extract system data themselves. Instead, the authors are frequently provided with a prepared data set in MS Excel or CSV format [40, 42] and thus have no influence on the extraction and data preparation.

Recent research suggests making use of the CRISP-DM as a guidance for SCA [7, 17]. Behrendt et al. [12] provide descriptions of data dimensions in ESS and a conceptualisation of digital traces [46], which are refined by Schwade and Schubert for the purpose of Social Collaboration Analytics (SCA) [7]. The Social Collaboration Analytics Framework [7] provides additional guidance for developing questions that can be posed to an existing data set and for developing the respective data queries. Consequently, we identified the publication by Behrendt et al. [12], Hacker et al. [17] as well as Schwade and Schubert [7] as key publications in the field, which could provide guidance for future research in this topic area.

In a recent literature review on Enterprise Social Networks, Viol and Hess [2] identified the meta-topics (1) *implementation*, (2) *motivation to adopt and use*, (3) *usage and behaviour*, (4) *impact on organisation*, (5) *success measurement* and (6) *data analytics* as the predominant research streams on ESS. The following concluding part of our paper discusses these meta-topics in relation with the key themes identified in our literature review to establish a view of SCA in context. There is no direct relation between the meta-topics *implementation of ESS* and *motivation to adopt and use* and the measurement concepts. However, there are some studies attempting the measurement of the adoption of systems. According to Viol and Hess [2], the meta-topic *usage and behaviour* publications aims to identify how employees use ESS and to identify usage patterns. This is also true for the measurement concept *usage patterns*. Therefore, this aspect of analytics has potential to support and to contribute to the biggest research stream on ESN according to Viol and Hess [2]. The same applies to the impact of ESN on organisations. This demonstrates that the meta-topic *data analytics* and thus the measurement concepts identified in this literature review are intertwined with their research meta-topics and that SCA can contribute to ESS research. This argument is strengthened by the observation that publications on success measurement increasingly make use of system measurements.

In this paper, we identified key themes for SCA from the literature. However, we have not yet confirmed whether these key themes are relevant in practice. In the current phase of our research on SCA, we are working with practitioners in order to identify and develop meaningful methods and metrics, which are also relevant in practice. Thus, the key themes provide the foundation for future research on SCA.

References

1. Williams, S.P., Hausmann, V., Hardy, C.A., Schubert, P.: Enterprise 2.0 Research: Meeting the Challenges of Practice. In: 26th International Bled Conference. pp. 251–263. Bled, Slovenia (2013).
2. Viol, J., Hess, J.: Information Systems Research on Enterprise Social Networks – A State-of-the-Art Analysis. In: Multikonferenz Wirtschaftsinformatik (MKWI 2016). pp. 351–362. Ilmenau, Germany (2016).
3. Wehner, B., Ritter, C., Leist, S.: Enterprise social networks: A literature review and research agenda. Comput. Networks. 114, 125–142 (2016).
4. Meske, C., Stieglitz, S., Middelbeck, D.: Mehrwerte von Intranet Social Soft- ware – Status quo in der Wissenschaft. In: Multikonferenz Wirtschaftsinformatik. pp. 1775–1785 (2014).
5. Richter, A., Heidemann, J., Klier, M., Behrendt, S.: Success Measurement of Enterprise Social Networks. In: Wirtschaftsinformatik 2013. pp. 1–15 (2013).
6. Herzog, C., Richter, A., Steinhueser, M.: Towards a Framework for the Evaluation Design of Enterprise Social Software. In: International Conference on Information Systems. pp. 1–20. Fort Worth, USA (2015).
7. Schwade, F., Schubert, P.: Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities. In: 50th Hawaii International Conference on System Sciences (HICSS). pp. 401–410. Hilton Waikoloa Village, Hawaii, USA (2017).
8. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii–xxiii (2002).
9. Torraco, R.J.: Writing integrative literature reviews: Guidelines and examples. Hum. Resour. Dev. Rev. 4, 356–367 (2005).
10. vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., Cleven, A.: Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. Commun. Assoc. Inf. Syst. 37, 205–224 (2015).
11. Shami, N.S., Muller, M., Millen, D.: Browse and discover: social file sharing in the enterprise. In: ACM Conference on Computer Supported Cooperative Work. pp. 295–304. ACM (2011).
12. Behrendt, S., Richter, A., Trier, M.: Mixed methods analysis of enterprise social networks. Comput. Networks. 75, 560–577 (2014).
13. Steinhueser, M., Herzog, C., Richter, A., Hoppe, U.: A process perspective on the evaluation of enterprise social software. In: 2nd European Conference on Social Media. p. 429436. Porto, Portugal (2015).

14. Appelt, W.: What groupware functionality do users really use? Analysis of the usage of the BSCW system. In: Euromicro Workshop on Parallel and Distributed Processing. pp. 337–341. Mantova, Italy (2001).
15. Jeners, N., Lobunets, O., Prinz, W.: What groupware functionality do users really use?: A study of collaboration within digital ecosystems. In: 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST). pp. 49–54 (2013).
16. Muller, M., Freyne, J., Dugan, C., Millen, D.R., Thom-Santelli, J.: Return On Contribution (ROC): A metric for enterprise social software. In: 11th European Conference on Computer Supported Cooperative Work (ECSCW). pp. 143–150. Vienna, Austria (2009).
17. Hacker, J., Bodendorf, F., Lorenz, P.: A Framework to Analyze Enterprise Social Network Data. In: Atzmueller, M., Samia, O., and Roth-Berghöfer, T. (eds.) Enterprise Big Data Engineering, Analytics, and Management. pp. 84–107. IGI Global (2016).
18. Malsbender, A., Recker, J.C., Kohlborn, T., Beverungen, D., Tanwer, S.: Much Ado about Nothing? Tracing the Progress of Innovations Borne on Enterprise Social Network Sites. In: 34th International Conference on Information Systems. Milan, Italy (2013).
19. Jeners, N., Prinz, W.: Metrics for Cooperative Systems. In: GROUP'14 18th International Conference on Supporting Group Work. pp. 91–99. ACM (2014).
20. Millen, D.R., Yang, M., Whittaker, S., Feinberg, J.: Social bookmarking and exploratory search. In: 10th European Conference on Computer Supported Cooperative Work. pp. 21–40. Limerick, Ireland (2007).
21. Ferron, M., Massa, P., Odella, F.: Analyzing collaborative networks emerging in Enterprise 2.0: the Taolin Platform. In: 4th & 5th UK Social Networks Conference. pp. 68–78. Trento, Italy (2011).
22. Muller, M., Ehrlich, K., Matthews, T., Perer, A., Ronen, I., Guy, I.: Diversity among enterprise online communities: collaborating, teaming, and innovating through social media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2815–2824 (2012).
23. Bøving, K.B., Simonsen, J.: http Log Analysis - an Approach to Studying the Use of Web-Based Information Systems. Scand. J. Inf. Syst. 16, (2004).
24. Chaves, M.A., Cordoba, E.R.: Deciphering event logs in SharePoint Server: A methodology based on process mining. In: XL Latin American Computing Conference (CLEI). pp. 1–12. IEEE (2014).
25. Naderipour, F.: Mining of ad-hoc business processes using Microsoft Sharepoint, NITRO & PROM 6.0. In: 13th International Conference on Enterprise Information Systems. pp. 413–418. SciTePress - Science and and Technology Publications (2011).
26. Schubert, P., Glitsch, J.H.: Use Cases and Collaboration Scenarios: How employees use socially-enabled Enterprise Collaboration Systems (ECS). Int. J. Inf. Syst. Proj. Manag. 4, 41–62 (2016).
27. Muller, M., Shami, N.S., Millen, D.R., Feinberg, J.: We are all lurkers: consuming behaviors among authors and readers in an enterprise file-sharing service. In: 16th ACM international conference on Supporting group work. pp. 201–210. Sanibel Island, Florida, USA (2010).
28. Hacker, J., Bodendorf, F., Lorenz, P.: Helper, Sharer or Seeker?-A Concept to Determine Knowledge Worker Roles in Enterprise Social Networks. In: 13th International Conference on Wirtschaftsinformatik. pp. 668–682. St. Gallen, Switzerland (2017).
29. Hacker, J., Bernsmann, R., Riemer, K.: Dimensions of User Behavior in Enterprise Social Networks. In: Helms, R., Cranefield, J., and Reijsen, J. van (eds.) Social Knowledge Management in Action. pp. 125–146. Springer International Publishing, Heidelberg, Germany (2017).

30. Matthews, T., Whittaker, S., Badenes, H., Smith, B. a, Muller, M., Ehrlich, K., Zhou, M.X., Lau, T.: Community Insights : Helping Community Leaders Enhance the Value of Enterprise Online Communities. In: SIGCHI conference on human factors in computing systems. pp. 513–522. Paris, France (2013).
31. Damianos, L.E., Holtzblatt, L.J.: Measuring Community Success: One Size does not fit all. Bedford, MA, USA (2010).
32. Xu, A., Chen, J., Matthews, T., Muller, M., Badenes, H.: CommunityCompare: Visually comparing communities for online community leaders in the enterprise. In: SIGCHI conference on human factors in computing systems. pp. 523–532. Paris, France (2013).
33. Nasirifard, P., Peristeras, V.: Expertise extracting within online shared workspaces. In: Proceedings of the WebSci'09: Society On-Line (2009).
34. John, A., Seligmann, D.: Collaborative tagging and expertise in the enterprise. In: Collaborative Web Tagging Workshop. Edinburgh, UK (2006).
35. Williams, S.P., Schubert, P.: Social Business Readiness Survey 2014. University of Koblenz-Landau, Working Report of the Research Group Business Software, Koblenz (2015).
36. Smith, M., Hansen, D.L., Gleave, E.: Analyzing enterprise social media networks. In: 12th International Conference on Science and Engineering (CSE). pp. 705–710. Vancouver, BC, Canada (2009).
37. Behrendt, S., Richter, A., Schäfer, S., Trier, M.: Business Intelligence 2.0. In: Business Intelligence for New-Generation Managers. pp. 97–111. Springer International Publishing, Cham, Heidelberg, New York, Dordrecht, London (2015).
38. Wu, A., Dimicco, J.M., Millen, D.R.: Detecting professional versus personal closeness using an enterprise social network site. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 1955–1964 (2010).
39. Nasirifard, P., Peristeras, V., Hayes, C., Decker, S.: Extracting and Utilizing Social Networks from Log Files of Shared Workspaces. In: 10th IFIP Working Conference on Virtual Enterprises. pp. 3–11 (2009).
40. Behrendt, S., Klier, M., Klier, J., Richter, A., Wiesneth, K.: The Impact of Formal Hierarchies on Enterprise Social Networking Behavior. In: 36th International Conference on Information Systems. pp. 1–19 (2015).
41. Stieglitz, S., Riemer, K., Meske, C.: Hierarchy or Activity? The Role of Formal and Informal Influence in Eliciting Responses from Enterprise Social Networks. In: 22nd European Conference on Information Systems (ECIS). Tel Aviv, Israel (2014).
42. Riemer, K., Stieglitz, S., Meske, C.: From Top to Bottom: Investigating the Changing Role of Hierarchy in Enterprise Social Networks. Bus. Inf. Syst. Eng. 57, 197–212 (2015).
43. Warshaw, J., Whittaker, S., Matthews, T., Smith, B. a: When Distance Doesn't Really Matter: Effects of Geographic Dispersion on Participation in Online Enterprise Communities. In: 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. pp. 335–345. New York, NY, USA (2016).
44. van Osch, W., Steinfield, C.: Boundary Spanning through Enterprise Social Software: An External Stakeholder Perspective. In: 34th International Conference on Information Systems. Milan, Italy (2013).
45. Steinfield, C., van Osch, W., Zhao, Y.: Intra-Organizational Boundary Spanning: A Machine Learning Approach. In: 21st Americas Conference on Information Systems (AMCIS). Puerto Rico (2015).
46. Behrendt, S., Richter, A., Riemer, K.: Conceptualisation of Digital Traces for the Identification of Informal Networks in Enterprise Social Networks. In: 25th Australasian Conference on Information Systems. pp. 1–10 (2014).

Combining Humans and Machine Learning: A Novel Approach for Evaluating Crowdsourcing Contributions in Idea Contests

Dominik Dellermann[1], Nikolaus Lipusch[1], Mahei Manhai Li[1]

[1] University of Kassel, ITeG, Chair for Information Systems, Kassel, Germany
{dellermann,lipusch,mahei.li}@uni-kassel.de

Abstract. The creative potential from innovative contributions of the crowd constitutes some critical challenges. The quantity of contributions and the resource demands to identify valuable ideas is high and remains challenging for firms that apply open innovation initiatives. To solve these problems, research on algorithmic approaches proved to be a valuable way by identifying metrics to distinguish between high and low-quality ideas. However, such filtering approaches always risk missing promising ideas by classifying good ideas as bad ones. In response, organizations have turned to the crowd to not just for generating ideas but also to evaluate them to filter high quality contributions. However, such crowd-based filtering approaches tend to perform poorly in practice as they make unrealistic demands on the crowd. We, therefore, conduct a design science research project to provide prescriptive knowledge on how to combine machine learning techniques with crowd evaluation to adaptively assign humans to ideas.

Keywords: idea contests, design-science research, machine learning, idea filtering

1 Introduction

Firms increasingly engage in open innovation efforts to leverage the creative potential of a huge and diverse crowd of contributors [1-2]. Therefore, one popular approach is to solve innovative problems by starting an open call to a crowd with heterogeneous knowledge and diverse experience via a web-based innovation platform (e.g., BrightIdea, Salesforce, and Ideascale). Individual members of the crowd then contribute creative ideas to solve such problems and the firm rewards the best contribution in a contest approach [3]. This novel way to solicit ideas from online communities is a powerful mechanism to utilize open innovation.

However, the creative potential that arises from the innovative contributions of the crowd constitutes some critical challenges. The quantity of contributions and the demands on expertise to identify valuable ideas is high and remains challenging for firms that apply crowdsourcing [4]. Famous examples illustrate these novel phenomena. For instance, during the IBM “Innovation Jam” in 2006 more than 150,000 users from 104 countries generated 46,000 product ideas for the company [5]. Moreover, Google launched a crowd-innovation challenge in 2008 to ask the crowd ideas that have the potential to change the world in their “Project 10^100”. After

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

receiving over 150,000 submissions, thousands of Google employees reviewed the ideas to pick a winner, which took nearly two years and tens of thousands of dollars [6]. As previous research suggests only about 10–30% of the ideas from crowdsourcing engagements are considered valuable. Furthermore, screening this vast amount of contributions to identify the most promising ideas is one of the toughest challenge of crowdsourcing to date [7].

To solve these problems, different streams of research emerged that attempt to filter ideas [8]. First, expert evaluations, which use executives within the firm to screen ideas, were identified as costly and time consuming [9]. Second, research on algorithmic approaches proved to be a valuable way by identifying metrics to distinguish between high- and low-quality ideas [10-12]. However, such filtering approaches always risk missing promising ideas by identifying "false negatives" (classifying good ideas as bad ones) and are rather capable to cull low quality ideas than identifying valuable ones, which is a task that demands human decision makers. In response to this, the third approach to screen ideas is crowd-based evaluation [8-9, 13]. Organizations have turned to the crowd to not just for generating ideas but also to evaluate them to filter high quality contributions. This way has in fact shown to be of same accuracy such as expert ratings, if the members of the crowd have suitable domain knowledge [14]. However, this approach frequently fails in practice, when facing huge amounts of ideas. Crowd-based filtering approaches tend to perform poorly as they make unrealistic demands on the crowd regarding their expertise, time, and cognitive effort [8].

To address the aforementioned shortcomings, we propose the following research question: *How should idea filtering approaches be designed that improve the identification of valuable ideas in large scale crowdsourcing engagements?*

By combining algorithmic machine learning approaches with human evaluation to adaptively assign crowd members that have the required domain knowledge to ideas, we propose a semi-automatic approach that leverages the benefits of both approaches and overcomes limitations of previous research. We thus propose that a hybrid approach is superior to sole crowd-based and computational evaluation for two reasons: First, various research suggests that computational models (or machines) are better at tasks such as information processing and provide valid results [15], while human decision makers are cognitively constrained or biased [16]. Additionally, previous research shows the importance of human decision makers in the context of innovation [17]. In this highly uncertain and creative context, decision makers can rely on their intuition or gut feeling [18].

Following a design science approach [19-20], we so far identified awareness of real-world problems in the context of filtering crowdsourcing contributions and derived design principles for such systems, which we evaluated with experts on crowdsourcing and requirement engineering. We then explain our further progress of research and how we plan to implement and evaluate our proposed filtering approach ex post.

We, therefore, intend to extend previous research on idea filtering in crowdsourcing engagements through combining algorithmic and crowd-based evaluation. This research therefore will contribute to both descriptive and prescriptive knowledge [21-22], which may guide the development of similar solutions in the future.

2 Related Work

Idea Contests

In general, crowdsourcing denotes a mechanism that allows individuals or companies, who face a problem to openly call upon a mass of people over the web to provide potentially valuable solutions. One instantiation of crowdsourcing that seems to be particularly interesting from both a practical and a research perspective are idea contests [1,23-25]. Idea contests are usually conducted via platforms that allow companies to collect ideas from outside the organization. The output (i.e. the ideas) of such contests are usually artefact ideas that can take on different forms such as plain text, plans, designs and predictions from both experts and lay crowds [25]. The basic idea behind idea contests is thereby for companies to expand the solution space to a problem and thereby increasing the probability to obtain creative solutions to said problem [3,25]. The effectiveness of idea contests is also underpinned by research showing that only under certain conditions users are willing, as well as capable to come up with innovative ideas [24,26]. Thus, by providing various incentives such as monetary rewards, firms increase the number of contributions and the probability to receive a creative submission [27]. In simple terms attracting larger crowds leads to a more diverse set of solutions [28-29].

Previous Approaches to Identify Valuable Ideas

Such idea contests lead to a high number of ideas that cannot be efficiently processed by current approaches. Thus, successful idea contests often lead to a flood of contributions that must be screened and evaluated before they can be moved to the next stage and further developed [7]. To identify valuable contributions that are worth implementing, one important task is filtering the textual contributions in such idea contests. Existing filtering approaches to separate valuable from bad contributions in crowdsourcing mainly apply two content-based filtering approaches to evaluate the creative potential of ideas: computational, algorithmic evaluation approaches and crowd-based evaluation approaches [8].

Computational Evaluation Approaches

One current approach to evaluate textual contributions in the context of crowdsourcing is computational evaluation, wherein algorithms are used to filter ideas based on metrics for idea quality such as word frequency statistics [10]. Within the approaches for computational evaluation, two dominant approaches are emerging to support the decision making of the jury, which reviews the ideas to identify the most valuable ones.

First, clustering procedures examine how the vast amount of textual data from crowdsourcing contributions can be organized based on topics [10] or domain-independent taxonomy for idea annotation [11]. Second, machine learning approaches can be used to filter ideas based on rules that determine the value of the content [12, 30]. This approach is particularly useful if training data sets are available. Previous research in this context uses variables for contextual (e.g., length, specificity, completeness, writing style) or representational (e.g., readability, spelling mistakes) characteristics [12] as well as crowd activity (e.g., likes, page views, comments), and behavior of the contributor of the idea (e.g. date of submission, number of updates) [30] to determine the value of crowdsourcing contributions.

Crowd-based Evaluation Approaches

The second approach to evaluate crowdsourcing contributions is applying crowd-based evaluation approaches. In this context, members of the crowd evaluate contributions individually and the results are aggregated [8-9]. Such users might include other users of the contest, or even paid crowds on crowd work platforms [31] that are asked to evaluate ideas from the crowdsourcing engagement.

Previous research on crowd-based evaluation examined the applicability of one or multiple criteria [32] in voting mechanism (where users vote for valuable ideas) [33], ranking approaches (where members of the crowd rank submissions) [34], and rating mechanisms (where the crowd score ideas) [9]. Moreover, prediction markets can be used where users trade ideas by buying and selling stocks to identify the most valuable idea by aggregating this trades as a stock price [35]. Depending on the context of evaluation settings, these approaches proved to be equally accurate compared to the evaluation of experts [14].

3 Methodology

For resolving the above-mentioned limitations, we conduct a design science research (DSR) project [19-20, 22] to design a new and innovative artifact that helps to solve a real-world problem. To combine both relevance and rigor we use inputs from the practical problem domain and the existing body of knowledge (rigor) for our research project [36]. Abstract theoretical knowledge thus has a dual role. First, it guides the suggestions for a potential solution. Second, the abstract learnings from our design serve as prescriptive knowledge to develop other artefacts that address similar problems in the future [21]. To conduct our research, we followed the iterative design research cycle methodology interpretation of [37] as illustrated in figure 1.

So far, we analyzed the body of knowledge on collective intelligence, idea contests, and crowd-based evaluation as well as computational filtering approaches and identified five theory-driven problems of current idea filtering approaches that adversely affects evaluation accuracy. These problems represent the starting point for our solution design. Based on deductive reasoning, we derived five design principles for a potential solution that we evaluated in an ex-ante criteria-based evaluation with experts in the field of community- and service -engineering [38]. In the next steps, we will develop a prototype version of the novel filtering technique and implement it within the context of an idea contest. By conducting an A/B-test to compare the accuracy of our filtering approach against current filtering approaches [30], we intend to evaluate our proposed design. This also constitutes our summative design evaluation [38]. We will, therefore, use a consensual assessment of experts as baseline [9]. Finally, the abstract learning from our design will provide prescriptive knowledge in the form of principles of form and function for building similar artefacts in the future [21].

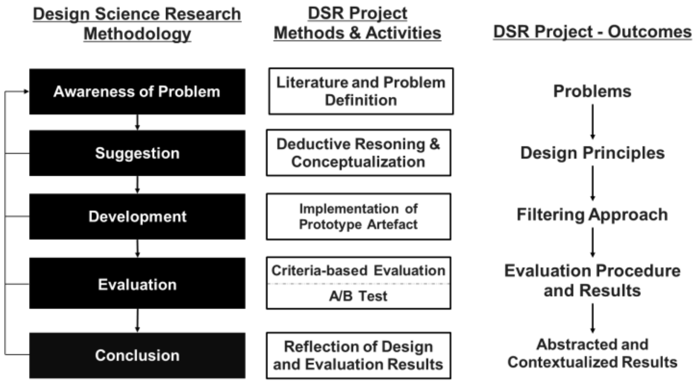


Figure 1. Research Design

4 Awareness of Limitations of Computational and Crowd Approaches

One solution that is currently employed in idea contests is shortlisting. Shortlisting can be considered as an algorithmic solution with the aim to shortlist the best ideas. In doing so shortlisting algorithms often face a tradeoff between specificity and sensitivity. Thus, if such algorithms are not balanced out (i.e. they are too specific, or they are too sensitive) this may lead to ideas being shortlisted that are not innovative (i.e. the algorithm might include false positives) or to promising ideas not being shortlisted (i.e. the algorithm might favor false negatives). In both cases this might lead to unfavorable results such as ideas that are labelled as innovative when in fact they are not truly innovative ideas (**Problem 1**).

One limitation of previous crowd-based evaluation approaches is the cognitive load associated with the volume and variety of idea contributions in crowdsourcing [8]. As cognitive load increases, users in the crowd may become frustrated [39] make low quality decisions [9] or simply deny evaluating ideas. Such load may arise due to the complexity of the evaluation mechanism itself (e.g., prediction markets) and the increasing time and cognitive complexity demands for the raters. Moreover, the information overload in which cognitive processing capacity is exceeded by the volume and diversity of the crowdsourcing contributions makes it difficult for the crowd to evaluate each idea especially when the proposals are complex, such as in the context of innovation problems. Thus, users need to judge manifold, diverse, maybe even paradox ideas with a high degree of novelty. This cognitive load renders previous approaches of crowd-based evaluation problematic for use in practice, where the number of contributions is large (**Problem 2**).

Furthermore, contributions will vary in their textual representation such as writing style, schema, or language which accelerates the cognitive demands on the crowd. Consequently, in practice only a small number of contributions are evaluated. These contributions and their (positive) evaluations then create an anchoring effect [40-41] and will socially influence other decision makers in the crowd [42]. Generally, the ones that are presented on the top of the page and have been positively evaluated by peers a priori, which creates (potentially negative) information cascades [8] (**Problem 3**).

Another major problem in crowd-based evaluation methods so far is that not all users in an idea contest are necessarily capable to evaluate ideas. Therefore, the crowd-based evaluation results might not be a proxy for expert ratings, if users do not have the required expertise for being a "judge" [14], [43-44]. This is particularly problematic when crowdsourcing contributions are complex and diverse. Although previous research highlighted the requirements on the crowd for evaluating ideas, the bottleneck of domain expertise is almost neglected in both theory and practice. To be appropriate for identifying valuable ideas and improving decision quality and predictions in idea filtering, a user should also be an expert in the field [25,45]. Therefore, the crowd should combine both problem knowledge as well as solution knowledge [4,46], which are crucial in the evaluation of innovation. While knowledge about the problem domain might be assumed for users that contribute an idea to a specific problem call, the variety of submitted solutions might be enormous as each diverse solver within the crowd deeply know different parts of the potential solution landscape [24]. Therefore, not every user in the crowd is equally appropriate to evaluate a certain idea due to limited domain knowledge of each part of the solution space submitted, which represents a major weakness of previous approaches in crowd-based evaluation (**Problem 4**).

5 Suggestion and Development of Design Principles for a Hybrid Filtering Approach

To overcome the limitations of previous approaches and to define objectives for a potential solution, we combine algorithmic approaches from machine learning with crowd-based evaluation approaches rather than treat them as substitutes. This approach enables our solution to support the human judge by using machine learning algorithms that identify the expertise of a crowd user, the expertise requirements for evaluating a specific crowdsourcing contribution, and match both to gather more reliable results in identifying valuable contributions. Our proposed design principles (DP)mainly focuses on improving the idea evaluation phase in innovation contests (see Figure 1).

First, the expertise requirements for each textual contribution needs to be identified to match it with suitable members of the crowd [44]. Therefore, the hybrid filtering approach should extract topical features (i.e. latent semantics) to identify the knowledge requirements for potential judges. Thus, we propose:

DP1: *Filtering crowdsourcing contributions should be supported by approaches that extract solution knowledge requirements from textual idea contributions within an idea contest by identifying relevant themes.*

In the next step, the hybrid filtering approach needs to consider the expertise of a crowd participant [47]. One source of such expertise description is the user profile, which includes the self-selected proficiency of a participant. Thus, we propose:

DP2: *Filtering crowdsourcing contributions should be supported by approaches that screen user profiles to extract expertise.*

Apart from the expertise description in the users´ profile (i.e. static), crowd participants gain ability through their activity (i.e. dynamic) in idea contests over time. Users constantly learn through their own contributions [48]. This needs to be additionally considered for the hybrid filtering approach. Moreover, this offers the possibility to ensure that users have really expertise in a topic as they proved it by making contributions. In contrast, expertise descriptions in user profiles might be biased due to overconfidence. Thus, we propose:

DP3: *Filtering crowdsourcing contributions should be supported by approaches that extract solution expertise from users´ prior textual idea contributions across idea contests by identifying relevant themes.*

Idea contest are highly dynamic [7]. To match crowd participants with suitable ideas for evaluation, the expertise profiles of each user need to be dynamic [49]. This means it should constantly update the expertise of a user through dynamically updating the abstract user profile based on the input and contributions of a crowd participant. Contributions include both past idea proposals, as well an idea quality indicator (i.e. the corresponding idea rating) Thus, we propose:

DP4: *Filtering crowdsourcing contributions should be supported by approaches that create adaptive user profiles containing expertise extracted from the user profile and prior contributions.*

As the evaluation quality of the crowd is highly dependent on the ability of each individual member of the crowd [44], in the last step the hybrid filtering approach needs to match crowdsourcing contributions with suitable users. Previous work on such select crowd strategies in the field of psychology suggests that approximately five to ten humans are required to benefit from the aggregated results of evaluation [43-44]. This sample size is most suitable for leveraging the error reduction of individual biases as well as the aggregation of diverse knowledge. Thus, we propose:

DP5: *Filtering crowdsourcing contributions should be supported by approaches that match solutions with users that have the required expertise and assign textual contributions to this user for evaluation.*

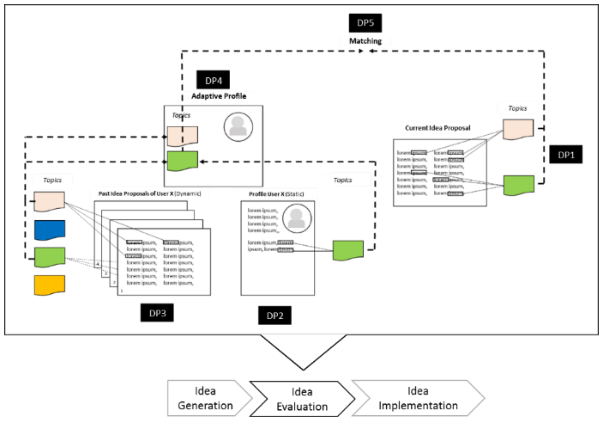


Figure 2. Proposed Hybrid Filtering Approach

6 Ex-Ante Evaluation of Design Principles

For the criteria-based ex-ante evaluation of our artifact (DP1 -DP5) we conducted an online-survey with experts on community and systems engineering. The criteria for our evaluation were derived from [38]. Specifically, we made use of the following criteria: completeness, understandability, fidelity with real world, applicability, level of detail, internal consistency, clarity. Hence, experts were asked to evaluate each of the proposed design principles based on the aforementioned criteria. Table 1 provides the results to our criteria-based ex-ante evaluation.

Table 1: Results of ex-ante design principles evaluation

criteria	N	DP1 Mean (SD)	p	DP2 Mean (SD)	p	DP3 Mean (SD)	p	DP4 Mean (SD)	p	DP5 Mean (SD)	p
completeness	13	4.85 (1.07)	.0017	5.38 (.96)	.000111	5.32 (.93)	.00022	5.23 (0.83)	.00009	4.69 (.75)	.003055
understand-ability	13	5.31 (.85)	.000066	5.15 (1.14)	.001694	4.92 (.76)	.000448	4.77 (1.17)	.017381	5.38 (1.12)	.000396
fidelity	13	5.23 (1.17)	.001242	4.15 (1.07)	.001074	5.54 (1.05)	.000097	4.77 (1.30)	.027277	5.15 (.80)	.000111
applicability	13	5.15 (1.21)	.002493	5.31 (0.58)	.000066	5.0 (1.0)	00179	5.00 (1.21)	.007045	4.85 (1.21)	.013704
level of detail	13	5.15 (1.14)	.001694	5.32 (1.17)	.001242	5.08 (.95)	.000777	5.08 (1.19)	.003352	5.23 (.93)	.00022
internal consistency	13	5.58 (.79)	.00001	5.46 (0.97)	.000074	5.46 (1.27)	.000661	4.85 (0.69)	.000411	5.23 (1.09)	.000791
clarity	13	5.15 (.90)	.00029	5.46 (1.27)	.000661	5.0 (1.08)	.002944	5.15 (1.34)	.004681	5.33 (.98)	.000262

Our results suggest that the majority of our design principles score relatively high in terms of the proposed criteria (i.e. the means of our criteria to evaluate DP1-DP5 are found on the upper bound of a seven-point Likert-scale). This is also supported by the p-values indicating that the scores of our criteria are all significantly different from the mean. Based on these results, we conclude that our design principles are clear and concise to warrant further development and refinement of our idea filtering approach.

7 Further Work and Summative Evaluation

As we proceed, we will develop and implement our hybrid filtering approach within the context of an idea contest on an existing crowdsourcing platform. To identify required solution knowledge and users´ expertise, we will design a machine learning algorithm based on probabilistic topic modelling [51]. Topic modeling is a text mining approach that uses Latent Dirichlet Allocation (LDA) [50] as unsupervised statistical learning method to discover abstract “topics” in text documents.

We then automatically match users to current idea proposals based on the proximity of topics extracted from the static user profiles and their past idea contributions [51-52]. The developed filtering approach will then be evaluated. We will therefore conduct two A/B-tests to compare our filtering approach against current filtering approaches. Our first comparison will include our hybrid filtering approach (A) against a computational filtering approach (B1). Our second comparison will include our hybrid filtering approach against a crowd-based filtering approach (B2) [30]. For our filtering approach, we will match each idea with approximately five users and then ask for evaluation on a rating scale [13] and then combine the evaluations to derive the average (i.e. mean) [44]. The performance of both filtering approaches will then be evaluated against a baseline. For constructing our baseline, we used the most commonly approach of expert evaluation through consensual assessment technique, which combines the consensus-based classification of a crowdsourcing contribution through several domain experts [9]. Figure 3 displays our planned evaluation procedure.

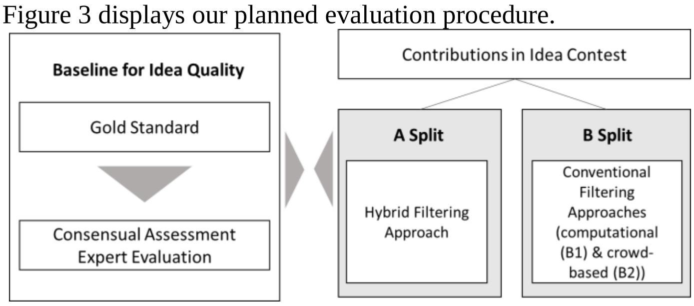


Figure 3. Proposed Evaluation Procedure

The performance of filtering approaches will be assessed through its accuracy. Therefore, we calculate the true positive (correctly identified high quality idea related to baseline) and the false positive (incorrectly identified high quality idea related to baseline) for each filtering approach. The receiver operating characteristic (ROC) plots the true positive rate against the false positive rate. The area under this curve then provides a measure for accuracy. A perfectly accurate filter approach would have an area of 1.0 [8].

8 Conclusion

Our research introduces a novel filtering approach that combines the strengths of both machines and humans in evaluating creative ideas by using machine learning approaches to assign the right user with the required solution knowledge to a corresponding idea. To this end, we propose tentative design principles that we validated in the field with experts on crowdsourcing and system engineering. To the best of our knowledge, this is the first study that takes this topic into account. Our research offers a novel and innovative solution for a real-world problem and contribute to the body of knowledge on idea filtering for open innovation systems by considering the required expertise of crowd evaluations [43-44]. We, therefore, intend to extend previous research on idea filtering in crowdsourcing engagements through combining algorithmic and crowd-based evaluation.

References

1. Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H.: Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition. Journal of Management Information Systems 26, pp. 197–224 (2009)
2. Afuah, A., Tucci, C. L.: Crowdsourcing as a solution to distant search. Academy of Management Review 37, pp. 355–375 (2012)
3. Lakhani, K. R., Jeppesen, L. B.: Getting unusual suspects to solve R&D puzzles. Harvard Business Review 85, pp. 30–32 (2007)
4. Toubia, O., Florès, L.: Adaptive idea screening using consumers, Marketing Science 26, pp. 342–360 (2007)
5. Bjelland, O. M., Wood, R. C.: An inside view of IBM's' Innovation Jam'. MIT Sloan Management Review 50, p. 32 (2008)
6. Buskirk, E. V.: Google Struggles to Give Away $10 Million. Wired Magazine (2010)
7. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: how to benefit from (too) many great ideas. MIS Quarterly Executive 12, pp. 199–211 (2013)
8. Klein, M., Garcia, A.C.B.: High-speed idea filtering with the bag of lemons. Decision Support Systems 78, pp. 39–50 (2015)
9. Blohm, I., Riedl, C., Füller, J., Leimeister, J.M.: Rate or trade? identifying winning ideas in open idea sourcing. Information Systems Research 27, pp. 27–48 (2016)
10. Walter, T.P., and Back, A.: A text mining approach to evaluate submissions to crowdsourcing contests. Proceedings of the Hawaii International Conference on Systems Science (2013)
11. Westerski, A., Dalamagas, T., Iglesias, C.A.: Classifying and comparing community innovation in Idea Management Systems. Decision Support Systems 54, pp. 1316–132 (2013)

12. Rhyn, M., Blohm, I.: A Machine Learning Approach for Classifying Textual Data in Crowdsourcing. Proceedings of the WI 2017 (2017)
13. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: The effect of rating scales on decision quality and user attitudes in online innovation communities. International Journal of Electronic Commerce (17:3), pp. 7–36 (2013)
14. Magnusson, P.R., Wästlund, E., Netz, J.: Exploring users' appropriateness as a proxy for experts when screening new product/service ideas. Journal of Product Innovation Management 33, pp. 4–18 (2016)
15. Nagar, Y., Malone, T.: Making business predictions by combining human and machine intelligence in prediction markets. Proceedings of the International Conference on Information Systems (2011)
16. Kahneman, D., Tversky, A.: Variants of uncertainty. Cognition 11, pp. 143–157 (1982)
17. Kornish, L.J., Ulrich, K.T.: The importance of the raw idea in innovation: Testing the sow's ear hypothesis, Journal of Marketing Research 51, pp. 14–26 (2014)
18. Huang, L., Pearce, J.L. Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. In: Administrative Science Quarterly 60, pp. 634–670 (2015)
19. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. In: MIS Quarterly 28, pp. 75–105 (2004)
20. Peffers, K., Tuunanen, T., Rothenberger, M.A., and Chatterjee, S.: A design science research methodology for information systems research. Journal of management information systems 24, pp. 45–77 (2007)
21. Gregor, S., Jones, D.: The anatomy of a design theory. Journal of the Association for Information Systems 8, p. 312 (2007)
22. Gregor, S., Hevner, A.R. Positioning and presenting design science research for maximum impact. MIS Quarterly 37, pp. 337–355 (2013)
23. Boudreau, K.J., Lakhani, K.R.: Using the crowd as an innovation partner. Harvard Business Review 91, pp. 60–69 (2013)
24. Jeppesen, L.B., Lakhani, K.R.: Marginality and problem-solving effectiveness in broadcast search. Organization Science 21, pp. 1016–1033 (2010)
25. Terwiesch, C., Xu, Y.: Innovation contests, open innovation, and multiagent problem solving. Management science 54, pp. 1529–1543 (2008)
26. Poetz, M.K., Schreier, M.: The value of crowdsourcing: can users really compete with professionals in generating new product ideas? Journal of Product Innovation Management 29, pp. 245–256 (2012)
27. Bayus, B.L.: Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. Management science 59, pp. 226–244 (2013)
28. Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing crowds: Mapping the genome of collective intelligence. MIT Working Paper Series (2009)
29. Surowiecki, J.: The wisdom of crowds: Anchor. (2005)
30. Nagar, Y., Boer, P. de, Bicharra Garcia, A.C.: Accelerating the Review of Complex Intellectual Artefacts in Crowdsourced Innovation Challenges. Proceedings of the International Conference on Information Systems (2016)
31. John, T.: Supporting Business Model Idea Generation Through Machine-generated Ideas: A Design Theory. Proceedings of the International Conference on Information Systems (2016)
32. Dean, D.L., Hender, J.M., Rodgers, T.L., Santanen, E.: Identifying good ideas: constructs and scales for idea evaluation (2006)
33. Bao, J., Sakamoto, Y., Nickerson, J.V.: Evaluating Design Solutions Using Crowds, Proceedings of the Americas Conference on Information Systems (2011)
34. Salganik, M.J., Levy, K.E.C. Wiki surveys: Open and quantifiable social data collection (2012)

35. Soukhoroukova, A., Spann, M., & Skiera, B.: Sourcing, filtering, and evaluating new product ideas: An empirical exploration of the performance of idea markets. Journal of Product Innovation Management 29, pp. 100-112 (2012).
36. Hevner, A. R. A three-cycle view of design science research. Scandinavian Journal of Information Systems 19, p. 4 (2007)
37. Wagenknecht, T., Crommelinck, J., Teubner, T., Weinhardt, C.: When Life Gives You Lemons: How rating scales affect user activity and frustration in collaborative evaluation processes. Proceedings of the WI 2017 (2017)
38. Vaishnavi, Vijay K., and William Kuechler. Design science research methods and patterns: innovating information and communication technology. Crc Press, (2015)
39. Sonnenberg, C., & vom Brocke, J.: Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. Design Science Research in Information Systems. Advances in Theory and Practice, pp. 381-397 (2012).
40. Tversky, A., Kahneman, D.: Availability: A heuristic for judging frequency and probability. Cognitive Psychology 5, pp. 207–232 (1973)
41. Larrick, R.P., Mannes, A.E., Soll, J.B., Krueger, J.I.: The social psychology of the wisdom of crowds. Social Psychology and Decision Making, pp. 227–242 (2011)
42. Deutsch, M., Gerard, H.B.: A study of normative and informational social influences upon individual judgment. The Journal of Abnormal and Social Psychology 51, p. 629. (1955)
43. Mannes, A.E., Soll, J.B., Larrick, R.P.: The wisdom of select crowds. Journal of Personality and Social Psychology 107, p. 276 (2014)
44. Keuschnigg, M., Ganser, C.: Crowd Wisdom Relies on Agents' Ability in Small Groups with a Voting Aggregation Rule. Management Science 63, pp. 818–828 (2017)
45. Amabile, T.: Creativity in context: Westview Press (1996)
46. Hippel, E. von.: "Sticky information" and the focus of problem solving: implications for innovation. Management Science 40, pp. 429–439 (1994)
47. Kulkarni, A., Narula, P., Rolnitzky, D., & Kontny, N.: Wish: Amplifying creative ability with expert crowds. In Second AAAI Conference on Human Computation and Crowdsourcing 2014.
48. Faraj, S., Krogh, G. von, Monteiro, E., Lakhani, K.R.: Special Section Introduction—Online Community as Space for Knowledge Flows. Information Systems Research 27, pp. 668–684 (2016)
49. Yuan, A., Luther, K., Krause, M., Vennix, S. I., Dow, S. P., & Hartmann, B.: Almost an expert: The effects of rubrics and expertise on perceived value of crowdsourced design critiques. In Proceedings of the 19th ACM Conference on CSCW & Social Computing, pp. 1005-1017 (2016)
50. Blei, D.M.: Probabilistic topic models. Communications of the ACM 55, pp. 77–84 (2012)
51. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of Machine Learning Research 3, pp. 993–1022 (2003)
52. Shi, Z.M., Lee, G.M., Whinston, A.B.: Toward a Better Measure of Business Proximity: Topic Modeling for Industry Intelligence. MIS Quarterly 40, pp. 1035–1056 (2016)

Cyber-physische Systeme und digitale Wertschöpfungsnetzwerke

Teilkonferenzleitung

Daniel Beverungen
Christian Janiesch
Martin Matzner

Activity Monitoring Using Wearable Sensors in Manual Production Processes - An Application of CPS for Automated Ergonomic Assessments

Rocco Raso[1], Andreas Emrich[1], Torsten Burghardt[2], Martin Schlenker[3], Thomas Gudehus[4], Oliver Sträter[2], Peter Fettke[1] and Peter Loos[1]

[1] German Research Center for Artificial Intelligence (DFKI), Institute for Information Systems (IWi), Saarbrücken, Germany
{rocco.raso,andreas.emrich,peter.fettke,peter.loos}@dfki.de
[2] University of Kassel, Institute of Industrial Science and Process Management, Kassel, Germany
{torsten.burghardt,oliver.sträter}@uni-kassel.de
[3] Society for Intelligent Textile Products (ITP GmbH), Weimar, Germany
schlenker@itp-gmbh.de
[4] Informations-Technologie für Menschen (ITfM GmbH), Morschen, Germany
tgudehus@itfm.de

Abstract. The automated identification and analysis of human activities in a manufacturing context represent an interesting challenge to support the workers, providing novel solutions for managing and optimizing existing manufacturing processes. If on one hand real-time coupling of events and activities is a relatively easy task for activities which are executed by means of information systems, on the other hand, the coupling of events and human physical activities remains an unsolved problem. In this paper we present a novel paradigm based on the integration of a light-weight, low-cost body sensor network and a software solution based on machine learning for tracking working operations. This enables the fast identification of inconvenient ergonomic behaviors and process aspects, which are objective of workflow analysis, process improvement and optimization activities. To assess the usability and functionality of the system a study under real conditions was conducted in the logistic plant of a big automobile manufacturer.

Keywords: Activity Recognition, Process Identification, CPS, Wearable Sensors

1 Introduction

The paradigms of Industry 4.0 and Industrial Internet aim at enabling the automation of production processes by means of the development of smart factories using the Internet of Things [1]. Smart factories integrate information systems and physical entities to trace components during every production step and to receive information from the machines during their lifecycle. This enables to deduce the workflows of production and business processes for a better process management. Process discovery approaches, by means for instance of process mining techniques, enable to

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

check the conformity of accomplished processes and process models, in order to identify anomalies and inconvenient situations during the process execution [2].

Manual activities still play an important role in the industry. Semi-automated and manual manufacturing systems are largely present in industrial contexts and the complete automation of several manufacturing sectors represents a utopian target [3]. Activities with high human involvement are often necessary and hardly replaceable by machines. Nowadays, even in very automated manufacturing contexts, certain types of assembly tasks, loading and unloading operations cannot be automated.

The goal of this paper is to provide an approach for IoT-based systems for monitoring and assessing work situations in manual work processes. We propose an automated system, which combines wearable technology and machine learning for detecting workflows and operations during logistic processes with high human involvement. We integrate activity recognition approaches with process identification techniques, in order to detect human activities for identifying information about executed processes and possible anomalies. By means of this information, the system is able to learn patterns for correlating inconvenient ergonomic behaviors and process steps, making possible to predict potentially dangerous working activities.

This paper aims at answering following research questions:

- How can we integrate body sensor networks and AI techniques based on machine learning for identifying and analyzing human activities in manufacturing?
- Is it possible to combine ergonomic and process information for predicting potentially inconvenient or dangerous process activities?

The result of our research is an infrastructure based on hardware components and a software framework, which makes possible the automated conformity check between process model and process execution, supporting the workers with relevant feedback.

1.1 Methodology and paper structure

According to Hevner's design science research approach [4], our research follows seven guidelines to be rigorous and relevant. Our newly implemented IT infrastructure for a low-cost body sensor network, presented in section 4 and developed according to the requirements described in section 3, represents the innovative and purposeful artifact (Guideline 1) The introduction section and the related work specify the relevance of the research the problem domain (Guideline 2). The aspects concerning the evaluation of the artifact (Guideline 3) are considered in section 5. A prototypical implementation has been developed to demonstrate our artifact and to show the feasibility of the system and its innovative strength (Guideline 4). The prototypical implementation presented in this paper represents our proof of concept. It enables the real-time recognition of human activities during logistic operations, which permits to deduce information about inconvenient ergonomic behaviors and their correlation with workflow and process steps. Section 1 shows the respect of a rigorous methodological research approach, which reflects the coherence and the consistency of the artifact (Guideline 5). Conceiving the design science as a search process (Guideline 6), in section 2 we show that the presented approach considers the relevance of previous related work. The purpose to permit the communication of the research (Guideline 7) is expressed publishing this paper.

2 Related Work

2.1 Activity Recognition for Process Identification of Manual Operations

The issue of process identification and analysis of manual working activities represents an interesting workbench for testing cutting-edge technologies and redesigning traditional paradigms for manual manufacturing. Tracking working activities enables to gain important information to better control and manage business processes, in different areas of application [5,6]. For instance, process reengineering takes advantage of process identification to shape ideal target processes [7]. In production scheduling and planning, process identification and analysis techniques are employed to model dynamic process planning system [8]. To be able to identify specific processes, it is, of course, necessary to detect activities and relevant process steps during their execution. Different methods and technologies are used for process identification tasks. According to the classification of human motion of Aggarwal and Cai 1999 [9], human motion analysis can be distinguished into body structure analysis and tracking of human motion. From a technical point of view, analysis and tracking of human activities can be done by means of optical tracking or integrating sensor-based cyber-physical systems. Marker-based optical systems can be employed to recognize steps of a process and specific objects, which are useful to infer information about workflows and processes. Systems based on optical tracking with markerless technology can provide classifications of work tasks in industrial environments [10]. Similar approaches have been used in manufacturing to extract the temporal structure of workflows [11, 12]. Sensor-based cyber-physical systems can be wearable or non-wearable. Wearable computers are one approach to track physical activities of humans with a focus on tracking of assembly steps in car manufacturing [13, 14]. Lee and Mase 2002 [15] developed one of the first systems of body sensors to detect activities and location of the user. Kitagawa and Windsor 2008 [16] compare different motion capture approaches from industry. Even if they are very common, traditional systems for activity recognition based on motion capture and optical technology present big limitations: the very large computational power and storage needed, the limitation of the workers’ movements to the field of view of the cameras and the reduction of the workers’ privacy, which limit the acceptance of the system.

2.2 Real-Time Ergonomic Analysis

Within manufacturing and logistic processes, musculoskeletal disorders caused by manual tasks represent one of the main factors of all work-related disorders [17]. Ergonomic analyses are usually done by methods of observation, which evaluate the ergonomic situation after that the exposure to critical postures, or positions, has taken place [18]. Different tools are used to evaluate the ergonomic working conditions. Examples are the “Ovako Working posture Analysis System” (OWAS) [19], the

“Rapid Upper Limb Assessment” (RULA) [20] and the Ergonomic Assessment Worksheet (EAWS) [21]. All ergonomic evaluation methods by observation have own advantages and focus. Since 2009 EAWS is the standard in the automobile industry [22]. These methods of observation serve primarily the structural prevention (external factors for a better working environment; i.e. redesigning the workplace or reducing physical demands within the working conditions). The other part is the behavioral prevention, which focuses on the individual behavior of the worker and tries to improve the working condition through training and education [23]. There are only a few existing systems for providing real-time ergonomic feedback. The RULA score can be used to assess the criticality of the posture, or position, and the modalities for feedback are acoustic and haptic [24]. Battini et al. 2014 combine a set of ergonomic methods (RULA, OWAS etc.) which can be chosen appropriately for the working condition. However, EAWS is not part of it. For the feedback, the system uses the visual modality [25]. Last two cited systems are based on motion capture technology, which, as we already analyzed, presents several limitations.

3 Requirements of the System

3.1 Requirements for Human Activity Identification in Manufacturing

Based on our concept idea and on the shortcomings of the state of the art analysis, we derived specific requirements for our artifact.

Requirement 1. ***The framework must be able to implement the corpus of sensor data.*** In order to automatically recognize patterns to be associated with related human operations, the framework must recognize and implement the available sensor data. Since the framework is based on machine learning techniques, several in-lab tests of the system must be accomplished before the framework reaches a sufficient training set of data, which is necessary for the good operation of the system. Moreover, preparatory tagging operations for the machine learning process of the system must be done before the start of the study under real conditions in the manufacturing plant.

Requirement 2. ***The feedback mechanism does not have to interfere with the working activities of the user.*** Since the manufacturing contexts are often very loud because of the presence of working machine, the use of acoustic feedback will not be considered. Haptic and visual feedback will be used to interface with the user, by means of a mobile device which is attached to the forearm of the user. This enables to perceive haptic feedback and to check the monitor of the device, in case the manual activity which the user is performing allows it in safety. The use of Head Mounted Displays (such as Google Glasses) as a feedback source for the user has been considered and the developed framework can integrate this typology of devices.

Requirement 3. ***The framework must be able to integrate data from different sources.*** In case of ambiguous or unclear identification of a working operation, further external sensor sources must be able to be integrated into the framework for solving ambiguities of interpretation. Data concerning ambiguous or unclear working activities are combined with external data (i.e. position data coming from the Wi-Fi net) and with available process information which has been already detected.

3.2 Requirements for Ergonomics and Feedback Interaction

To prevent critical ergonomic situations the analysis and feedback of postures must follow a designated and reproducible set of feedback criteria, as shown in the introduction and related work section.

Requirement 4. ***The framework must use a standardized procedure for the assessment of physical stress.*** For our system, the Ergonomic Assessment Worksheet (EAWS) [21] has been chosen as a guideline for feedback and input, for the workflow management system (and therefore to fulfill interlinking between structural and behavioral prevention). EAWS differentiates between standing, walking, sitting, and kneeling or squatting. These postures concern the upper body. Our system must enable feedback after exceeding the following EAWS positions: "Diffraction" ~60°; "Rotation" ~30°; "Tilt" ~30° and "Hands over the head".

Requirement 5. ***The framework must provide real-time ergonomic feedback about critical ergonomic postures.*** For the score of the ergonomic assessment through EAWS, points are accumulated over two factors: intensity and duration of the inconvenient posture [26]. For providing feedback about the criticality of the ergonomic posture in real-time, these measures must be combined and adjusted, according to a normalization process. For this reason, our system accumulates points when reaching a specific threshold within the values of the pressure and strain sensors (see EAWS positions above). After exceeding the threshold values over a timespan of two seconds, an alarm is triggered. The purpose of this design is that postures which are ergonomically critical can be avoided in advance. In the EAWS method, points are calculated only when a timespan of four seconds is exceeded. With our system, the user can avoid postures, which would result in a worse EAWS score, because of the possibility to leave the critical ergonomic posture within the two seconds of difference, which is calculated between our feedback system and the point where the EAWS method would calculate points for its score.

4 System Design and Development

4.1 Hardware Architecture

The hardware components are based on a flexible textile shirt with strain sensors, hydrostatic pressure units and an electrical part for analog/digital conversion with serial-send-unit. Sending process will be done with a 5 GHz Wi-Fi connection to a router. The mobile device feedback is also connected to this access point. For the experiment, the shirt was equipped with a high stretchable textile material to fit a variety of different human body families. Furthermore, these sensors have the capability to be physically calibrated in the same kind of variety. With the help of the strain sensors, the system is able to measure the demanded EAWS positions, diffraction, rotation and tilt of the upper body. In this way, the sensors are situated in the middle of the back to the waistband, symmetrically on both sides from the shoulder diagonally to the complementary side at the waistband and at last on each body side left and right. For the investigation of the position of hands in reference to the head, pressure elements attached to the top of the shoulder are used. The electronic components, which are fixed on the user's back, measure the singular values, convert them and send them in a specific order via a serial output to the host.

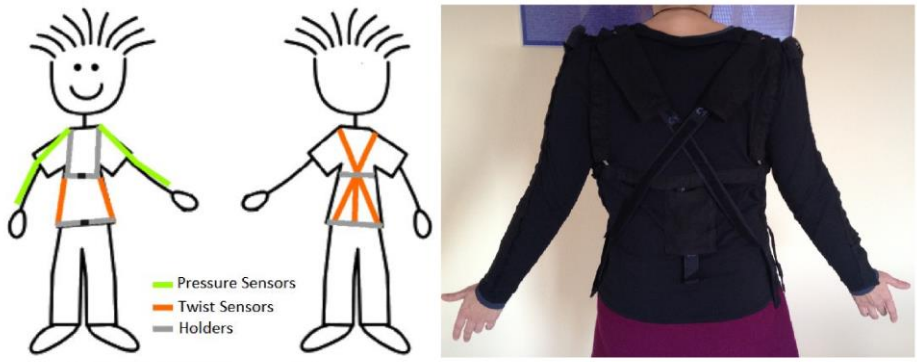


Figure 1. Representation and photo of the body sensor network which we used.

4.2 Software Architecture

Based on the requirements analysis, the software architecture covers the sensor processing, the process monitoring, discovery and analysis, as well as notification mechanisms for the user. The main component of the framework is the Process Engine. Its function is to analyze and process sensor data detected by the sensors and it has direct interface with every software component. Sensor data collected by the Sensor Body Network are sent and received via Wi-Fi and analyzed. The Process Engine deals with the ambiguity of activity recognition by means of applying context knowledge, according to the specificity of every process instance and considering the specificity of the knowledge domain where it operates. Its use of operative BPM tools enables the reasoning over existing model artifacts. Modeling notation techniques (such as eEPC and BPMN) are used from the Process Engine for enabling process type recognition and monitoring the course of detected processes. This connection with operative BPM systems makes possible the evaluation of the process history and the comparison between the planned model of the process and real process. For analyzing and processing complex events, it is necessary to consider the specifically involved entities. Therefore, the Process Engine is able to recognize the different process instances and to infer probabilistic considerations about the involved users. If for instance the system recognized that a specific user is performing an activity, but it is not able to distinguish between two very similar activities, the framework merges together different kinds of data (such as location data or older process instances of the user) to provide an accurate survey of the performed process instances.

A large amount of data from multiple sources is combined to infer specific process patterns and to solve potential equivocal behaviors, these are represented in Fig.2 as "External Sensors". After a process of filtering, normalization, aggregation and segmentation of the received information, the Process Engine initiates the storage of the event data on the Data Layer. The Process Engine is able to correlate newly received data with process histories and models, according to the specificity of every process instance. The system learns from received sensor data, enhancing the knowledge of the system about specific actions. This enables the possibility to provide related feedback to the user. By means of a Mobile Device attached to the forearm of the user, the Process Engine provides haptic and visual feedback about inconvenient or dangerous working operations. This feedback information is not only related to inconvenient ergonomic behaviours, but it is also about probabilistic considerations concerning sensor-based process activity predictions. A further possibility for receiving feedback is represented by the use of Head Mounted Displays, which have been considered for being integrated into the system.

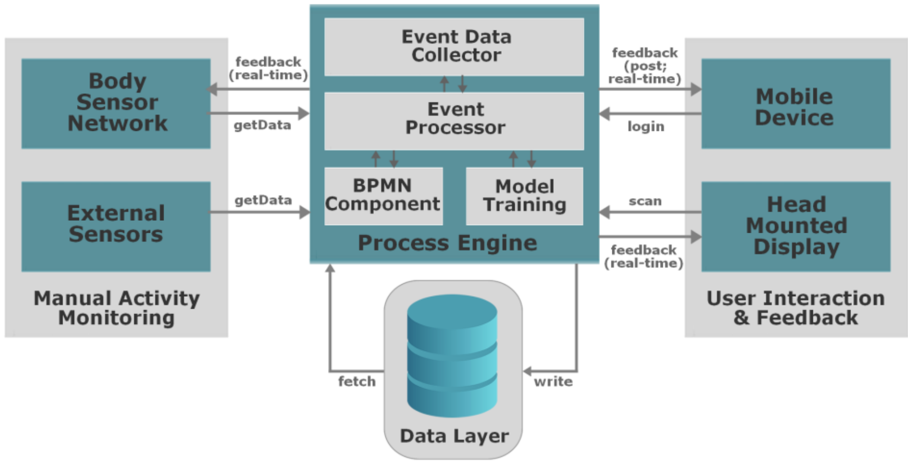


Figure 2. Architecture of the system

4.3 Sensor-Based Process Activity Prediction

The sensors as indicated in Figure 1 can be utilized to monitor the body posture and hence the load for certain body regions over time. As the actual load largely correlates with the task at hand, i.e. the action that the user performs, proactive support to prevent ergonomically bad situations by predicting the next actions, i.e. the next process activities based on the sensor data. In our model, we hence formulate the value ranges of the different sensors with t for the various twist sensors and p for the pressure sensors. We use this sensor data as the input vector to learn situations as a binary or multi-class-classification problem. We either learn, if an alarm occurred as a binary classification problem and the process activity a as a multi-class classification from the set of process activity types A. Due to the high intra-individual differences in the actual sensor values we translated the actual sensor values to an 80-item score scale that interpolates the score according to threshold values determined in an individual calibration being performed with every participant. This causes all t and p values being transformed to t’ and p’ respectively. Hence, the alarm prediction using binary classification can be formulated as follows:

$$\langle t'_1, \dots, t'_5, p'_1, p'_2 \rangle \rightarrow \begin{cases} 1, if\ alarm \\ 0, else \end{cases} \quad (1)$$

In a similar way, a simple multi-class classification that performs the process activity prediction can be formulated as follows:

$$\langle t'_1, \ldots, t'_5, p'_1, p'_2 \rangle \rightarrow a \in A \quad (2)$$

As the sensors capture mainly upper extremity movements, this approach is not entirely appropriate to distinguish under extremity movements. Hence, we also filter the results from (2) with all permitted activity types from the process model.

5 Use Cases

5.1 Execution of the Experiment

We conducted a study under real conditions in the plant of the Volkswagen Konzern After Sales, in Kassel (Germany) and we identified three different processes with high involvement of human operations, which constitute our three use cases. These are represented by three process models which differ for sequence and types of working operations. All use cases present analogous process activities, which represent the pool of detectable manual activities. In the first use case, the worker should move small but heavy car components to collocate them in appropriate places. For doing this, the worker has to load the components on a forklift, drive it to the right place and manually unload the components. In the second use case, the worker must interact with a semi-automated warehouse system, by means of a computer. After receiving requests on the monitor of the computer, the worker should manually load and unload car components from a goods lift to a different goods lift. The components are light and they have small dimensions. In the third use case, the worker should place several large and bulky car components inside of a container and then drive the forklift to bring the container to the warehouse. In every use case, the worker must place a printed barcode on the component before moving it to the appropriate place. In some case, if required, the worker must scan the barcode using a manual barcode-scanner. The experiment has been conducted in a real manufacturing plant by real workers as test persons. The study took place during six different days and every use case has been tested during two different days with 15 users. A total of 30 instances (10 per every use case) have been tested. Every instance lasts about 30 minutes and has been performed by one single user. During the same instance, the user performed the same working activities two times. During the first iteration, in case of inconvenient ergonomic position, the user did not receive any feedback, whereas during the second iteration the user did. Every user wore the described sensor body network and had a mobile device tied to the left forearm for receiving eventual feedback.

5.2 Evaluation

The design of the experiment has an A/B testing scenario. Before starting the experiment, every user rated the workload through a standardized questionnaire based on the NASA Task Load Index (NASA-TLX) [27]. This enables to examine the characteristics of subjective workload within following dimensions: mental, physical and temporal demand, performance, effort and frustration. By means of the questionnaire, loads on the users and design attributes of the system have been rated. During the experiment, all users have been monitored by an eye tracking device, in order to properly evaluate the system [28]. Eye tracking technology makes possible to determine and quantify the usability of the system since the reactions of the users could be recorded and associated with specific process activities. As an indicator of attention, the following parameters from the eye tracking device were analyzed: “Overall fixations”, “Overall gaze duration of the fixations”, “Look path, “Fixations on the display” and “Gaze duration of the fixations on the display”. All parameters were standardized (1/s) to allow comparability over all conditions. At the end of the experiment, all users filled out the NASA-TLX questionnaire again and rated the system with a second questionnaire according to ISO 9241-110 [29]. The second questionnaire enables to examine usability characteristics within following features: suitability for the task, suitability for individualization, conformity with user expectations, self-descriptiveness, controllability and error tolerance. Accuracy, relevance and utility of feedback provided to the users have been evaluated. Sensor data for every instance have been compared to determine the influence of feedback for avoiding inconvenient working operations. The precision of the prediction of process activities in relation to the available sensor data has been also evaluated.

5.3 Results

The items of the ISO 9241-110 have an average score of M=4.97 with an SD=0.49, which represent good results. However, the average score has still potential to be improved. The results show no significant increase in subjective workload in the NASA-TLX. The average values of the items changed from M=9.8 with an SD=3.2 without feedback to M=10.0 with an SD=3.5 at the condition with active feedback from the system to the user. The eye tracking parameters show the same results of questionnaires. No significant changes happen within the two conditions and this probably confirms that the choice of the questionnaires is appropriate and all available results are coherent. The change of the eye tracking parameters is shown in Table 1. This is an indicator that no serious additional demand was needed or generated by our system. The fixation on the display with an average frequency of M=0.63 per minute and an SD=0.37 in combination with the gaze duration of the fixations on the display (on average M=0.96s per fixation and an SD=0.67s) suppose that not much attention had to be drawn away from the manual activity to execute the feedback.

Table 1. Results of eye tracking.

Condition	Overall fixations per second	Overall gaze duration per fixation	Look path per second
Without feedback	*M=0.82 SD=0.28*	*M=1.37s SD=0.46s*	*M=1.21cm D=0.57cm*
With feedback	*M=0.83 SD=0.27*	*M=1.33s SD=0.44s*	*M=1.21cm SD=0.54cm*

Every instance has been analyzed and in particular the quantity of inconvenient manual operations detected by the system by means of sensor data has been considered and analyzed. These values constitute a specific pattern in the sensor data and according to this pattern, feedback for the user are generated by the system. Comparing the iteration without feedback and the iteration with feedback for the user related to the same instance, we observe a reduction of 54% of the quantity of feedback. This would mean that if the users receive feedback from our system, the users will probably decrease by about 54% the quantity of inconvenient behaviors. By means of the evaluation of the artifact, we identified during which process activity more feedback for inconvenient behaviors is sent to the users. We identified 21 typologies of process activities, which are common for all three use cases and then we analyzed the quantity of feedback per every process activity (Fig.3). We can observe that the process activities with the highest rate of feedback for inconvenient behaviors are following: “2-hands lifting” (28,5%), “2-hands arrange” (28,3%) and “2-hands dropping” (18,2%). These results are coherent, if we consider that the highest quantity of inconvenient or dangerous behaviors has been identified during three process activities in which the main task implied the loading of heavy components.

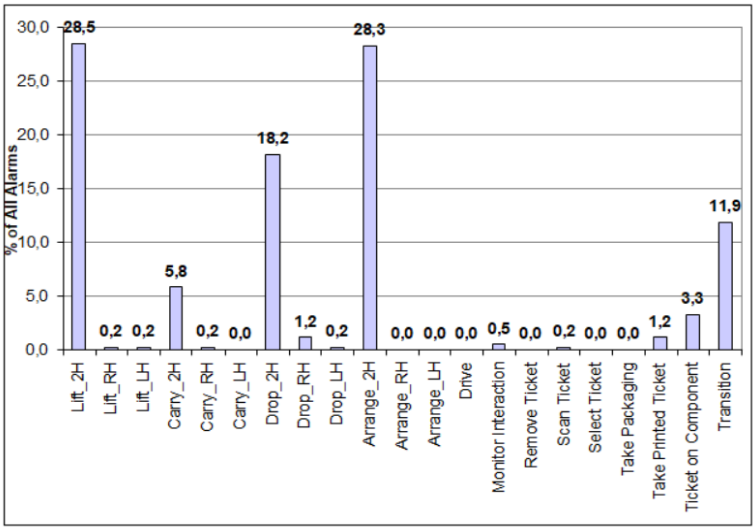


Figure 3. Percentage of alarm feedback per process activity.

In the practical evaluation, the system has been prepared and calibrated for every single user to cater to the individual physical characteristics (e.g. body height and shape) of each participant. The scoring system, as introduced in section 4.3, partially solves the problem of interpolating these values, however, it makes it somewhat more difficult to express the absolute values. For future serial production, the use of confection sizes could help this situation. The customer should be able to wear the shirt simply and comfortable - without doing a complex calibration. As the data showed, the prediction of process activities was limited by the sensor selection in the given suit. While this is appropriate for assessing the ergonomic situation of a person, it is not entirely possible to monitor all potential human activities of an actor in a business process. However, the results indicate, that the process discovery mechanism itself works. The limitation lies in the selection of sensors, which could be extended in future work. Another interesting effect was, that for the second use case, where the process was being followed in a more structured way by the employees, also the discovery accuracy was considerably higher.

6 Conclusion & Outlook

The approach presented in this paper introduced a new approach for discovering business processes from physical human activities captured by sensor data. The sensor data are being analyzed with a focus on ergonomic risks in the work process and associated with the given activities in a work process in order to determine organizational weak spots in the process or workplace design. The approach seems viable and useful from an end user's perspective as well as for a general approach for process discovery based on sensor data. While the low-cost sensor setup ensures a high applicability of the approach in industrial practice, it limits the ability to discover arbitrary kinds of business processes. Hence, for discovery accuracy, future work has to incorporate more exhaustive sensor setups, while for industrial setups cost-benefit analysis has to be performed in order to define a sufficient yet minimal sensor setup for a given set of workplaces.

7 Acknowledgement

This research has been possible thanks to the research project PREFLOW (project ID 16SV7286), which has been funded by the German Federal Ministry of Education and Research (BMBF) under the framework of the KMU-Innovativ Programme for Information and Communication Technologies. We would also like to show our gratitude to the Volkswagen Konzern After Sales Kassel and in particular to Mrs. Christine Schröder and the workers involved during the research and its evaluation.

References

1. Westkämper E, Jendoubi L (2003) Smart Factories – Manufacturing Environments and Systems of the Future. In: Bley, H (Eds.): Proceedings of the 36th CIRP International Seminar on Manufacturing Systems. Pp 13-16.
2. Van der Aalst, W. M., & Weijters, A. J. M. M. (2004). Process mining: a research agenda. Computers in industry, 53(3), 231-244.
3. Frohm, J., Lindström, V., Winroth, M. and Stahre, J. (2008). Levels of Automation in Manufacturing. Ergonomia IJE&HF, vol 30.
4. Hevner, A.R., March, S.T., Park, J. & Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28 (1), pp.75-105.
5. Balzert, S., Kleinert, T., Fettke, P., & Loos, P. (2011). Vorgehensmodelle im Geschäftsprozess-management - Operationalisierbarkeit von Methoden zur Prozesserhebung. IWi, (193), 61.
6. Knoch, S., Ponpathirkoottam, S., Fettke, P., & Loos, P. (2017) Technology-Enhanced Process Elicitation of Worker Activities in Manufacturing, International Workshop on Ubiquitous Business Processes Meeting Internet-of-Things (BP-Meet-IoT-17), 15th International Conference on Business Process Management, September 10-15, Barcelona, Spain, Springer, 2017.

7. Hammer, M., & Champy, J. (2001). Reengineering the Company - A Manifesto for Business Revolution. Harper business, New York USA.
8. Seethaler, R. J., & Yellowley, I. (2000). Process control and dynamic process planning. International Journal of Machine Tools and Manufacture, 40(2), 239–257.
9. Aggarwal, J. K., & Cai, Q. (1999). Human Motion Analysis: A Review. Computer Vision and Image Understanding, 73(3), 428–440.
10. Voulodimos, A., Kosmopoulos, D., Veres, G., Grabner, H., Van Gool, L., & Varvarigou, T. (2011). Online classification of visual tasks for industrial workflow monitoring. Neural Networks.
11. Nater, F., Grabner, H., and Gool, L. V., (2011). Unsupervised workow discovery in industrial environments, ICCV11 WS on Visual Surveillance , 1912–1919.
12. Veres, G., Grabner, H., Middleton, L., Gool, L.V., (2010) Automatic workflow monitoring in industrial environments. In: Proc. of ACCV
13. Kreil, M., Sick, B., and Lukowicz, P. (2014). "Dealing with human variability in motion based, wearable activity recognition," in IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops 2014), March, pp. 36–40.
14. Stiefmeier, T., Lombriser, C., Roggen, D., Junker, H., Ogris, G., and Troester, G. (2006). "Event-Based Activity Tracking in Work Environments," in \i3rd International Forum on Applied Wearable Computing (IFAWC 2006), March, pp. 1–10.
15. Lee, S.-W., and Mase, K., (2002). Activity and location recognition using wearable sensors. IEEE Pervasive Computing, 1(3):24–32,
16. Kitagawa, M., and Windsor, B. (2008). MoCap for Artists: Workflow and Techniques for Motion Capture (1.edition.), Amsterdam, Boston: Focal Press.
17. Meyer, M., Böttcher, M. & Glushanok, I. (2015). Krankheitsbedingte Fehlzeiten in der deutschen Wirtschaft im Jahr 2014 (Fehlzeiten-Report). In P.D.B. Badura, P.D.A. Ducki, H. Schröder, J. Klose & M. Meyer (Hrsg.), Fehlzeiten-Report 2015 (S. 341–548). Springer.
18. Klippert, J., Gudehus, T. & Zick, J. (2012). A software-based method for ergonomic posture assessment in automotive preproduction planning: Concordance and difference in using software and personal observation for assessments. Human Factors and Ergonomics in Manufacturing & Service Industries, 22 (2), 156–175.
19. Karhu, O., Kansi, P. & Kuorinka, I. (1977). Correcting working postures in industry: A practical method for analysis. Applied Ergonomics, 8 (4), 199–201.
20. McAtamney, L. & Nigel Corlett, E. (1993). RULA: a survey method for the investigation of work-related upper limb disorders. Applied Ergonomics, 24 (2), 91–99.
21. Schaub, K., Caragano, G., Britzke, B. & Bruder, R. (2010). The European Assembly Worksheet. In P. Mondelo, W. Karwowski, K.L. Saarela, P. Swuste & E. Occhipinti (Hrsg.), . Gehalten auf der VIII International Conference on Occupational Risk Prevention, Valencia.
22. Toledo, B. (2012). Global ergonomics strategy in Volkswagen: from the product construction, over the planning until the serial process. Work (Reading, Mass.), 41 Suppl 1, 4413–4417.
23. Schlick, C. M., Bruder, R. & Luczak, H. (2010). Arbeitswissenschaft. Springer-Verlag.
24. Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D. & Marin, F. (2013). Innovative system for real-time ergonomic feedback in industrial manufacturing. Applied Ergonomics, 566–574.
25. Battini, D., Persona, A. & Sgarbossa, F. (2014). Innovative Real-time System to Integrate Ergonomic Evaluations into Warehouse Design and Management. Comput. Ind. Eng., 1–10.
26. IAD und AMI (Hrsg.). (2014). Ergonomic Assessment Worksheet V.1.3.4. Institut für Arbeitswissenschaft der TU Darmstadt (IAD).
27. Hart, S. G. & Staveland, L. E. (2005). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Ergonomics: Skill, Displays, Controls, & Mental Workload.
28. Kassner, M. P., Patera, W. R. & Bulling, A. (2014). Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. Whitepaper.
29. Ergonomics of human-system interaction - Part 110: Dialogue principles (ISO 9241-110:2006)

Entwurf einer Referenzarchitektur zur operativen Abwicklung von Performance-based Contracting

Marcel Fisser[1], Thomas Sperber[1], Paul Menning[1], Robert Arnold[1] und Jürgen Anke[1]

[1] Hochschule für Telekommunikation Leipzig, Wirtschaftsinformatik, Leipzig, Deutschland
{marcel.fisser2, thomas.sperber, paul.menning, robert.arnold, anke}@hft-leipzig.de

Abstract. Ziel dieses Beitrags ist die Entwicklung einer technischen Referenzarchitektur (RA) für die operative Abwicklung von Performance-based Contracting (PBC) von vernetzten Produkten. Zur Erstellung der RA werden dabei wesentliche Komponenten aus Architekturen für Big Data und dem Internet of Things (IoT) miteinander verknüpft. Zudem wurden wesentliche benötigte PBC-Komponenten ergänzt und daraus eine instanziierbare und wiederverwendbare RA für PBC-Produkte erstellt. Aus wissenschaftlicher Sicht ist eine Grundlage zur Diskussion über die digitale Standardisierung von PBC-Produkten geschaffen worden. Die Validierung der RA durch Experten aus einem mittelständischen Unternehmen, welches PBC-Produkte auf Basis von IoT anbietet, konnte die Anwendbarkeit der aufgestellten RA zeigen.

Keywords: Performance-based Contracting (PBC), Big Data, Internet of Things (IoT), Industrie 4.0, Referenzarchitektur

1 Einleitung und Grundlagen

Performance-based Contracting (PBC) ist ein Geschäftsmodell, welches an Bedeutung gewinnt [1]. Immer mehr produzierende Unternehmen entscheiden sich für die Einführung von PBC [2], [3]. Die Digitalisierung, in Form von Internet of Things (IoT) und Big Data, bietet hierfür Grundlagen für die Umsetzung [4]. So ist es durch IoT-Techniken möglich, Daten in Echtzeit zu erfassen und über das Internet an den Betreiber zu senden. Die daraus resultierenden großen Datenmengen können mittels Big Data-Technologien gespeichert sowie verarbeitet werden [5], [6].

Bei der Realisierung von PBC für vernetzte Produkte, stellt sich den Unternehmen die Frage, welche IT-Architektur dafür eingesetzt werden soll. Ein möglicher Startpunkt ist die Verwendung einer Referenzarchitektur (RA). Ziel dieses Beitrags ist es, einen ersten Entwurf einer solchen Referenzarchitektur für PBC für Produkte in Anlehnung an IoT und Big Data zu entwickeln und zu evaluieren.

Performance-based Contracting. Je nach Perspektive wird PBC als Vertragsmodell, Geschäftsmodell oder Beschaffungsstrategie bezeichnet [1], [2], [7], [8]. Im Rahmen dieses Papers wird PBC als Geschäftsmodell verstanden. Der Fokus beim PBC liegt auf der zu erbringenden Leistung in Art, Umfang und Qualität und nicht auf Produkten, mit denen diese erzielt wird [7]. Wird ein Produkt im Rahmen des PBC-

Geschäftsmodells angeboten, schließen ein Kunde und der Hersteller des Produktes einen Vertrag. Innerhalb dieses Vertrages ist u.a. festgelegt, welche Leistung der Kunde von dem Produkt und damit auch von dem Hersteller benötigt und erwartet. Diese Leistungsvereinbarung ist innerhalb des Vertrages in den Service Level Agreements (SLAs) definiert. Um diese SLAs zu bewerten und zu überprüfen, werden zusätzlich Metriken festgelegt und eingeführt, die Kennzahlen liefern und damit die SLAs überprüfbar machen [9]. Diese Metriken können die Leistung des Services wertbasiert (nach dem Geldwert der erzeugten Leistung) oder ergebnisbasiert (nach der Menge der direkt erbrachten Leistungseinheiten) bewerten und messen [1]. Können diese Vereinbarungen durch den Betreiber nicht eingehalten werden, sind Regresszahlungen (negative Incentives) zu erwarten [1].

Big Data. Big Data steht für Daten, die zu umfangreich sind, um mit traditionellen Mitteln verarbeitet zu werden. Dies bezieht sich nicht nur auf das Datenvolumen, sondern auch auf die Strukturierung, den Grad der Verschiedenartigkeit und die Geschwindigkeit mit der sie erzeugt werden. Dies ergibt die häufig zur Beschreibung von Big Data verwendeten „3V“: Volume, Variety und Velocity [6], [10], [11]. Oftmals sind mit dem Begriff Big Data auch die Technologien gemeint, welche diese großen Datenmengen erfassen, verarbeiten und analysieren [12]. Die Autoren beziehen sich in dieser Arbeit auf den technischen Big Data Begriff.

Internet of Things (IoT). Laut van Kranenburg beschreibt IoT ein globales, dynamisches Netzwerk bei dem "Dinge" (Maschinen, Sensoren etc.) als selbstständige Einheiten mit eigener Identität und autonomen Fähigkeiten über das Internet miteinander kommunizieren und sich dabei auf Protokolle und Schnittstellen zur bidirektionalen Kommunikation stützen [13]. Für PBC ist dies relevant, da mittels IoT benötigte Daten erfasst und an ein Rechenzentrum gesendet werden können.

Referenzarchitektur. Diese Arbeit stützt sich auf die Definition des Rational Unified Process (RUP), laut der eine Referenzarchitektur ein Muster ist, das für den Einsatz in bestimmten betrieblichen und technischen Kontexten entwickelt wurde und dort nachweislich genutzt wird [14]. Besonderer Vorteil von RAs ist dabei die Wiederverwendbarkeit und Instanziierbarkeit, sowie die flexible Anwendbarkeit auf verschiedene Anwendungsszenarien von PBC [15].

2 Forschungsgegenstand und Problemdefinition

Vorteile der Digitalisierung von PBC. Wie bereits im ersten Abschnitt erwähnt, handelt es sich bei PBC um ein Geschäftsmodell, welches in den meisten Unternehmen bisher wenig bis gar nicht digitalisiert worden ist. Jedoch ist es aus betriebswirtschaftlicher Sicht sinnvoll, die PBC-Prozesse durch IT-Systeme zu unterstützen und damit effizienter zu machen. Gleichzeitig kann die Kundenzufriedenheit erhöht oder neue Geschäftsfelder erschlossen werden [15]. Zudem erlauben die technologischen Entwicklungen der letzten Jahre, dieses komplexe Geschäftsmodell zu digitalisieren. Gerade die Entwicklung in der Sensorik, in der Datenverarbeitung oder im Cloud-Computing, ermöglichen es große Datenmengen von vernetzten Produkten zu erfassen und zu verarbeiten [4].

Forschungsgegenstand und -ziel. Für die Digitalisierung eines Geschäftsmodells wird ein Überblick über die benötigten IT-Komponenten und deren Beziehungen untereinander benötigt. Hierfür wird in den meisten Fällen eine technische Architektur verwendet, die als Grundlage für die Entwicklung des Systems dient. Bei der Recherche nach PBC-Architekturen fällt auf, dass zu diesem Thema derzeit kaum einschlägige Forschungsergebnisse vorliegen. Jedoch gibt es im Kontext von PBC verwandte Architekturansätze, die in dieser Arbeit analysiert werden. Konkret handelt es sich um Big Data- und IoT-Architekturen, deren Zusammenhänge mit PBC im Abschnitt fünf näher erläutert werden. Es ist sinnvoll an dieser Forschungslücke anzusetzen, da die Nachfrage nach PBC in den letzten Jahren gestiegen ist, wie eine Studie des IT-Verbands Bitkom zeigt. Danach sind immer mehr Unternehmen an der Digitalisierung ihres Geschäftsmodells interessiert [16]. Daher ist das Ziel dieser Arbeit die Entwicklung einer RA für Performance-based Contracting (PBC-RA).

Einsatzzweck der Referenzarchitektur. Diese RA richtet sich hauptsächlich an Personen, die sich mit der Digitalisierung ihres PBC-Modells in ihrem Unternehmen beschäftigen oder die die RA als Basis für weiterführende Forschungen verwenden. Die in diesem Beitrag entwickelte RA soll auch nicht nur auf eine Produktgruppe anwendbar sein, sondern bei vielfältigen PBC-Anwendungsgebieten eingesetzt werden können. Dennoch muss die RA im Hinblick auf die prozessuale Gestaltung eingegrenzt werden, da diese sonst den Rahmen dieser Arbeit übersteigt. Im Fokus dieser Arbeit steht deswegen nur die operative Abwicklung von PBC, wie z.B. das Erfassen und Analysieren der Verbrauchsdaten, aber nicht die Vertragsgestaltung.

3 Verwandte Arbeiten

Die Forschungsfrage dieser Arbeit setzt sich aus den vier Kerngebieten "Entwicklung einer RA", "Big Data", "IoT" und "operative Abwicklung von PBC" zusammen, zu denen bereits verschiedene wissenschaftliche Arbeiten veröffentlicht wurden.

Im Bereich Big Data beschreiben Demchenko et al. [6] ein Big Data Architecture Framework (BDAF), welches alle Aspekte des Big Data-Ökosystems abdecken soll und sich aus Big Data Infrastruktur, Big Data Analytics, Datenstrukturen und -modellen, Big Data Life Cycle Management und Big Data Security zusammensetzt. Als Grundlage werden verschiedene Best Practices der Industrie verwendet. Außerdem schlagen Pääkkönen und Pakkala [11] eine technologieunabhängige RA für Big Data-Systeme auf Basis von bestehenden Big Data-Architekturen vor.

Einen Überblick über IoT gibt Ray [5]. Dort werden z.B. funktionale Bestandteile von IoT-Architekturen, Kommunikationsstandards, und verschiedene Sicherheitskonzepte aufgeführt und erläutert. Insbesondere die Informationen und Verweise auf andere Arbeiten zu den letzten beiden Punkten können bei der tatsächlichen Ableitung einer Architektur aus der RA hilfreich sein.

Dem Geschäftsmodell PBC wurde sich bereits aus verschiedenen Betrachtungswinkeln genähert. Selviaridisa und Wynstrab [1] haben diese Betrachtungsweisen zum Beispiel im Rahmen einer Literaturanalyse dargestellt. Eine Gruppe der Betrachter hat das Thema wie Martin [17] geschäftsseitig bearbeitet, indem er in seiner Arbeit beschreibt, wie das Interesse und die Bedeutung von PBC für öffentliche Einrichtungen und Behörden zugenommen hat. Dem lassen sich die Autoren von wissenschaftlichen Arbeiten zu CPS-Systemen gegenüberstellen, welche technische Teilaspekte von PBC untersuchen. Reidt und Krcmar [18] beschreiben eine RA für cyber-physische Systeme zur Unterstützung der Instandhaltung. Johanson und Karlson [19] beschreiben ein auf „Condition Monitoring“ und „simulation-driven availability prediction“ basierendes System, das Cloud Computing-Infrastruktur nutzt, um die Verfügbarkeit von Produkten oder Produktionssystemen sicherzustellen. Diese Arbeiten können als Basis genutzt werden, um konkrete Wartungskomponenten, in die in dieser Arbeit entwickelte RA, zu integrieren. Des Weiteren bildet die Arbeit von Reidt und Krcmar [18] die Basis für die von den Autoren umgesetzte Forschungsmethodik zur Entwicklung einer RA, welche in Kapitel 4 beschrieben wird. Wiederum prozessual eingegrenzt haben sich Sols et al. [3] in ihrem Artikel über die Vertragsfindung bei "Performance Based Logistics", welcher ein Framework zur Vertragsfindung bei PBL skizziert.

Zwischen diesen Forschungen ordnet sich daraus resultierend die Forschungslücke des vorliegenden Beitrags ein, da im Mittelpunkt der Analysen die operative Abwicklung von PBC erforscht und Punkte, wie die Vertragsfindung, nicht betrachtet werden. Außerdem hebt sich diese Arbeit von den anderen technischen Arbeiten ab, da es die umfassende digitale Umsetzung von PBC untersucht und nicht Teilbereiche, wie Verfügbarkeitsüberwachungs- und Wartungssysteme im Fokus stehen.

4 Forschungsmethodik

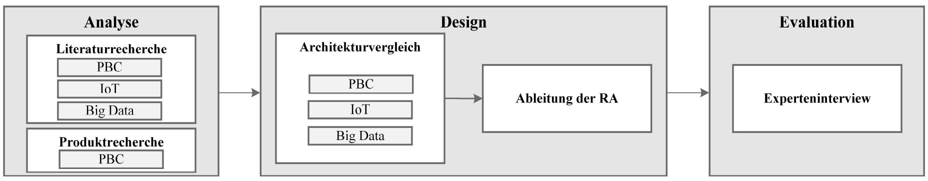


Abbildung 1. Vorgehensmodell

Zur Erstellung einer RA existieren verschiedene Vorgehensweisen, wie z.B. induktive oder deduktive Methoden. Induktive bzw. empirisch-deskriptive Erstellung einer RA bedeutet, dass aus einzelnen unternehmensspezifischen Architekturen eine allgemeingültige RA abstrahiert wird. Deduktive oder analytisch-präskriptive Erstellung einer RA erfolgt unter Einbeziehung theoriebasierter Erkenntnisse [20]. Da aber erst wenige dokumentierte Beispiele für IT-gestütztes PBC existieren und die von Unternehmen eingesetzten Architekturen zumeist nicht öffentlich zugänglich sind, war eine derartige Abstrahierung für diese Arbeit nicht möglich. Daher wurde eine dreistufige Vorgehensweise, in Anlehnung an Reidt und Krcmar [18] angewendet, da diese eine Referenzarchitektur für cyber-physische Systeme zur Unterstützung der Instandhaltung erstellt haben, welche thematisch eng verwandt ist mit der vorliegenden Arbeit und die Autoren ebenfalls das Problem hatten, dass keine offenen Architekturen identifiziert werden konnten. Deshalb haben sie eine Methodik erarbeitet, die auf den Grundkonzepten der gestaltungsorientierten Wirtschaftsinformatik [21] basiert und die Phasen Analyse, Design und Evaluierung beinhaltet. Abbildung 1 zeigt die einzelnen Phasen und deren Teilschritte, wie sie von den Autoren dieser Arbeit durchgeführt wurden.

Analyse. Im ersten Schritt wurde eine Literaturrecherche nach Webster und Watson [22] mit den Suchbegriffen „PBC", „RA", „IoT" und „Big Data" (in verschiedenen Schreibweisen) mithilfe der Suchmaschinen BASE[1] und ScienceDirect[2] durchgeführt, um den aktuellen Stand der Wissenschaft in diesen Bereichen zu ermitteln und relevante Literatur zu finden. Hieraus ergaben sich die RA-Beschreibungen und PBC-Funktionsbausteine, die nachfolgend verwendet werden. Des Weiteren wurden mit einer Produktrecherche zum Thema PBC frei verfügbare Informationen zu verschiedenen PBC-Angeboten, wie z.B. „Power by the Hour" von Rolls-Royce [23], „Pay per Wash"[3] von Winterhalter, „Instant Ink"[4] von HP oder „Cost per Unit"[5] im Bereich Fahrzeuglacke von BASF gesichtet und analysiert, um elementare PBC-Komponenten zu identifizieren und deren Eigenschaften für diese Arbeit zu definieren. Diese Unternehmen wurden per E-Mail kontaktiert und um weitere Informationen zu den einzelnen Produkten gebeten. Außerdem wurden IoT- und Big Data-Architekturen auf Standardelemente untersucht und es wurde geprüft, welche dieser Komponenten zur Erstellung einer PBC-RA verwendet werden können.

Design. Die Designphase setzt sich aus den beiden Schritten Architekturvergleich und Ableitung der RA zusammen. Im Rahmen des Architekturvergleichs wurden sowohl Big Data als auch IoT-RAs untersucht und es wurde geprüft, welche der zuvor identifizierten PBC-Komponenten durch welche Big Data-/IoT-Bestandteile abgebildet werden können. Resultat dieses Vergleichs ist es, dass eine Vielzahl der benötigen Komponenten für die Digitalisierung von PBC-Produkten durch Big Data und IoT abgedeckt werden können. Bei den spezifischen PBC-Funktionsbausteinen zur Leistungsabrechnung und zur Überwachung der Einhaltung der SLAs, auf Basis von ermittelten Metriken, stellten die Autoren jedoch fest, dass diese nicht durch Big Data- und IoT-Bestandteile abgebildet werden können.

Mit den Erkenntnissen aus dem Vergleich wurde im Anschluss die RA abgeleitet. Zuerst wurden die identifizierten Big Data- und IoT-Teile integriert, wodurch das wesentliche Gerüst der RA mit den Komponenten Data Ingest, Data & Analytics, Data Usage und Supporting Components festgelegt wurde. Daraufhin wurde geprüft, ob die fehlenden PBC-Funktionsbausteine in das bestehende Gerüst integriert werden können oder dieses hierfür erweitert werden muss. Hierzu wurden die Anforderungen und Eigenschaften der fehlenden Funktionsbausteine erneut betrachtet und mit dem bestehenden Grundgerüst verglichen. Dabei hat sich ergeben, dass keine zusätzliche Gruppe benötigt wird, sondern die fehlenden Funktionsbausteine durch eigene Komponenten in der Gruppe Data & Analytics integriert werden können.

---

[1] https://www.base-search.net/, zuletzt aufgerufen am 28.11.2017

[2] http://www.sciencedirect.com/, zuletzt aufgerufen am 28.11.2017

[3] http://www.pay-per-wash.biz/de_de/, zuletzt aufgerufen am 28.11.2017

[4] https://instantink.hpconnected.com/de/de, zuletzt aufgerufen am 28.11.2017

[5] http://www.basf-coatings.com/global/ecweb/de_DE/content/products_industries/automotive-oem-coatings/services/index, zuletzt aufgerufen am 28.11.2017

Evaluierung. Die RA wurde, analog zu Reidt und Krcmar [19], mithilfe eines Experteninterviews evaluiert. Dazu wurden in einer zweistündigen Online-Präsentation, verbunden mit einer Telefonkonferenz, drei Experten mit mehrjähriger Berufserfahrung befragt. Dabei handelte es sich um den Vertriebsleiter, den Produktleiter und den Serviceleiter eines führenden mittelständischen Unternehmens, das PBC für industrielle Anlagen bereits seit Jahren erfolgreich einsetzt und schon weitgehend digitalisiert hat. Anfragen an andere Unternehmen, deren Produkte im Rahmen der Produktrecherche analysiert wurden, blieben unbeantwortet, sodass die Evaluierung nur zusammen mit einem Unternehmen durchgeführt werden konnte.

Zur Vorbereitung des Interviews wurde eine Präsentation entwickelt, die die Hochschule, den Forschungsgegenstand und die RA vorstellte. Außerdem wurde ein Fragenkatalog erarbeitet, der sich an den einzelnen Komponenten der RA orientierte und somit als Leitfaden im Experteninterview diente. Das Experteninterview begann mit einer Vorstellung der teilnehmenden Personen und des Unternehmens. Im Laufe des Gespräches wurden die einzelnen Komponenten der PBC-RA vorgestellt und es wurde diskutiert, inwieweit sich diese in der Architektur des befragten Unternehmens wiederfinden lassen. Dabei stellte sich heraus, dass die Funktionen sämtlicher identifizierter Komponenten zugeordnet werden konnten, auch wenn ggf. andere Bezeichnungen innerhalb des Unternehmens hierfür verwendet werden. Es zeigte sich aber auch, dass das Unternehmen die Abrechnung der erbrachten Leistungen bisher noch nicht automatisiert hat, eine entsprechende Umsetzung aber einen Mehrwert für das Unternehmen darstellen würde. Während des Interviews wurde durch die Autoren ein Protokoll geführt, welches im Anschluss zusammen mit der Präsentation an alle Teilnehmer gesendet wurde. Damit konnte eine erste Validierung der PBC-RA anhand eines realen Falls erzielt werden.

5 Eine technische Referenzarchitektur für PBC

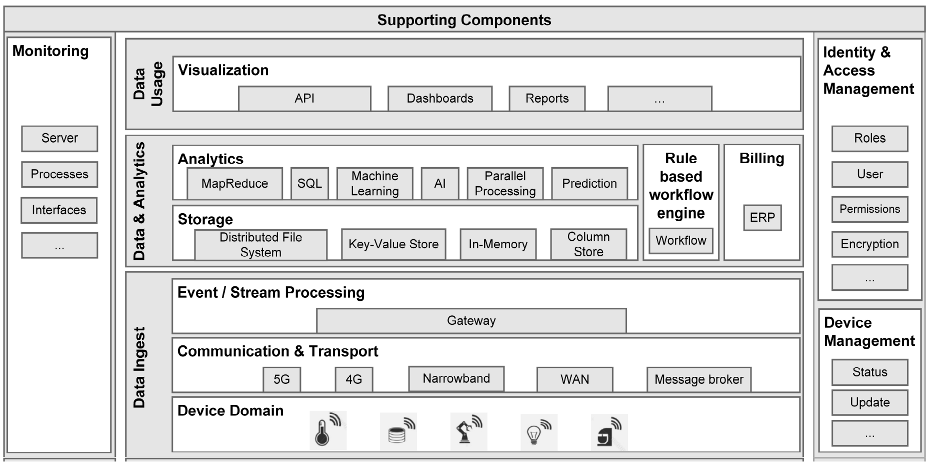


Abbildung 2. Technische Referenzarchitektur PBC

Unterschiedliche Big Data- und IoT-RAs verwenden abweichende Begriffe und Komponentennamen [24], obwohl sich aus deren Beschreibungen ergibt, dass sie häufig dasselbe meinen. Daher wurden für die RA dieser Arbeit Begriffe und Komponentennamen definiert, die die aus Sicht der Autoren die Funktion und Eigenschaft der aufgestellten RA am besten beschreiben. Die RA (Abbildung 2) organisiert die identifizierten Komponenten in den vier Gruppen Data Ingest, Data & Analytics, Data Usage und Supporting Components, die nachfolgend detaillierter beschrieben werden.

5.1 Data Ingest

Die Produkte, die über die IT-gestützte Form des PBC angeboten und betrieben werden, erzeugen bei ihrer Verwendung große Mengen an Daten, da stets die Benutzung in Echtzeit überwacht werden muss. Die aus dem englischen Sprachraum stammende Wortkombination aus „Data" und „Ingest" beschreibt in der erarbeiteten RA die Aufnahme der Daten der Produkte [25]. Grundsätzlich setzt sich diese Gruppe in der vorgeschlagenen RA aus den folgenden Komponenten zusammen:

Device Domain. Die unterste Komponente in dieser Gruppe ist die „Device Domain". Diese Komponente beinhaltet zunächst das Produkt, welches über die PBC-Plattform betrieben wird. Darüber hinaus beinhaltet diese Komponente die Geräte oder Bauteile, wie Sensoren, Mobile Devices oder Embedded Devices usw., die benötigt werden um relevante Messwerte und Daten zu erfassen [26]. Damit werden im späteren Verlauf die spezifischen PBC-Prozesse und Komponenten der RA mit Daten versorgt. Die Device Domain dient in der RA als Datenerzeugungskomponente. Diese Komponente bildet die Basis für die Digitalisierung von Produkten mit dem PBC-Geschäftsmodell, da ohne die erzeugten Zustandsinformationen der Produkte keine IT-gestützte Prozessautomatisierung des Geschäftsmodells möglich ist.

Communication & Transport. Die über die Komponente „Device Domain" erfassten Echtzeitdaten der PBC-Produkte werden über die Komponente „Communication & Transport" in die PBC-Plattform zur weiteren Verarbeitung übermittelt [26–28]. Entscheidendes Kriterium für die Wahl des Kommunikationskanals ist dabei u.a. die benötige und zur Verfügung stehende Bandbreite, um die Daten auf geeignetem Weg transportieren zu können. Beispiele für derartige Kommunikations- und Transportwege sind 4G, zukünftig 5G oder Narrowband-Netze, aber auch Message Broker Systeme für Local Area Network-Szenarien sind denkbar [29].

Event / Stream Processing. Diese Komponente stellt in der RA die wesentliche Integrationskomponente dar, um die in der „Device Domain" produzierten Daten technisch in die PBC-Plattform einzubinden. Stream-basierte Daten werden dabei über ein zentrales Gateway entgegengenommen und an die entsprechenden Funktionsgruppen innerhalb dieser Komponente zur Vorverarbeitung weitergeleitet [26]. Aufgabe dieser Komponente ist es dabei, die Stream-basierten Daten in Echtzeit anzunehmen und aufzubereiten, um anschließend die Speicherung der Daten in der nächsthöheren Komponente „Storage" sicherzustellen [25]. Die Herausforderung ist hierbei die Menge an eintreffenden Daten in naher Echtzeit zu verarbeiten.

5.2 Data & Analytics

Die Komponenten der Gruppe „Data & Analytics“ dienen zum einen der Datenhaltung der erzeugten Daten, zum anderen sind mit den Komponenten „Analytics“ und „Rule based workflow engine“ (RBWFE) die Geschäftslogik und Intelligenz der PBC-RA enthalten. Diese Komponenten sind spezifisch für PBC und weisen damit ein wesentliches Unterscheidungsmerkmal zu Big Data- und IoT-Architekturen auf.

Storage. Unter der Komponente „Storage“ sind alle Funktionsbausteine zu verstehen, die benötigt werden, um die durch die PBC-Produkte erzeugten Daten und Informationen für einen längeren Zeitraum zu persistieren und für Analyseereignisse vorzuhalten [29]. Dabei muss sichergestellt sein, dass genügend Speicherplatz für die großen Datenmengen zur Verfügung steht und dass schnell auf die Daten zugegriffen werden kann. Um dies zu erzielen kann auf wesentliche Kerntechnologien von Big Data-RAs zurückgegriffen werden [11]. Es bieten sich dafür die verteilte Speicherung der Daten auf mehreren Rechnerknoten oder der Einsatz von In-Memory-Speichersystemen an, um einen schnellen und effizienten Zugriff auf die Daten zu ermöglichen. Weitere Technologen, wie spaltenorientierte Speicherformate, Key Value Stores, etc. sind ebenfalls denkbar [28]. Eine Kombination der Technologien ist empfehlenswert, um deren Vorteile zusammenzubringen.

Rule based workflow engine. Ein bedeutender Baustein für die PBC-RA ist die RBWFE. Dieser Baustein bildet einen wesentlichen Teil der benötigten Geschäftslogik ab. Gleichzeitig stellt diese Komponente ein wesentliches Unterscheidungsmerkmal zu Big Data- und IoT-RAs dar.

Die in den Verträgen mit den Kunden festgelegten SLAs in Kombinationen mit den Metriken werden innerhalb dieser RA als sog. „Rules“ bezeichnet [30]. Diese Regeln werden in digitaler Form, für jeden Kunden, aus dem geschlossenen Vertrag in der RBWFE abgelegt. Im laufenden Betrieb der Plattform ist die RBWFE dafür verantwortlich, auf Basis der bereitgestellten und voranalysierten Daten, Metriken zu erstellen bzw. zu errechnen und pro Kunden zu überprüfen, ob die SLAs in Bezug auf die kundenindividuellen Metriken eingehalten werden [30]. Die Überprüfung des Status soll dabei nicht zyklisch oder sporadisch erfolgen, sondern in Echtzeit Auskünfte über den aktuellen Status liefern. Ergebnisse dieser Komponente können dann wieder von der „Storage“-Komponente abgelegt und unter Verwendung der Visualisierungskomponente für Endkunden und PBC-Betreiber angezeigt werden.

Letztendlich setzt sich diese Komponente daher aus zwei Teilen zusammen. Eine Komponente zum Erfassen der Regeln und eine Komponente mithilfe derer pro Kunde und Anwendungsfall Workflows definiert werden können. Definierte Regeln können dann im Workflow überprüft werden und wieder Ereignisse und Aktionen anstoßen, wie zum Beispiel das Versenden von E-Mails.

Analytics. Ein zweiter Baustein für die PBC-RA ist die Komponente „Analytics“ [25], die einen weiteren Teil der Geschäftslogik beinhaltet. Sie hat die Aufgabe, die von den PBC-Produkten und Sensoren generierten Rohdaten aufzubereiten, sodass die RBWFE, die in Form von Regeln beschriebenen SLAs überprüfen kann. Ziel dieser Komponente ist es, die definierten Metriken pro Kunde zu berechnen und bereitzustellen, damit die RBWFE die Auswertung übernehmen kann [30]. Die RBWFE- und „Analytics“-Komponente sind daher voneinander abhängig.

Um die Daten zu analysieren und aufzubereiten, muss die „Analytics“-Komponente u.a. mit großen Datenmengen umgehen können. Hierfür können vornehmlich Big Data-Technologien [31] verwendet werden. Diese können z. B. Batchprozesse in Verbindung mit MapReduce-Algorithmen sein, die eine parallele Verarbeitung der Daten ermöglichen. Aber auch einfache SQL-Abfragen auf den Daten sind denkbar. Die Wahl der einzusetzenden Technologien steht dabei in Zusammenhang zur Ablage der Daten in der „Storage“-Komponente. Darüber hinaus können in dieser Komponente weitere Analysen auf Basis der Daten durchgeführt werden. Mögliche Szenarien sind hier u.a. die Prädiktion der Daten, um das Verhalten und die Bedarfe auf Basis von historischen Daten zu ermitteln.

Billing. Die „Billing“-Komponente dient der Abrechnung der bereitgestellten Leistung. Dabei muss die Abrechnung auf individueller Basis, in Bezug zu den erzielten Metriken und eingehaltenen SLAs pro Kunde erfolgen [30]. Die Datenbasis zur Erstellung von Rechnungen erhält diese Komponenten dabei aus den Komponenten RBWFE und „Analytics“, die ihre Ergebnisse in einem Bereich der „Storage“-Komponente persistieren. Zur Erstellung der tatsächlichen Rechnungen für einen Endkunden können beispielsweise ERP-Systeme eingesetzt werden, die die benötigen Daten aus der „Storage“-Komponente über definierte Schnittstellen beziehen.

5.3 Data Usage

Die Gruppe „Data Usage“ beinhaltet einzig die Komponente „Visualisierung“. Diese stellt den Bereich dar, um die aufgenommenen und verarbeiteten Daten der Plattform für die Benutzer sichtbar zu machen [29]. Dabei können Werkzeuge zum Einsatz kommen, um beispielsweise Auswertungen, aktuelle Echtzeitdaten der Produkte oder ähnliches in Form von Grafiken, Tabellen oder anderen Visualisierungen für den Endkunden und die PBC-Betreiber zur Verfügung zu stellen [25]. Die Daten zur Visualisierung enthält die Komponente aus der Gruppe „Data Storage & Analytics“, in der die benötigen Daten bereits derart bereitgestellt werden, sodass die Visualisierungskomponente [28] diese nur noch anzeigen muss.

5.4 Supporting Components

Die Gruppe der unterstützenden Komponenten ist nicht unmittelbar an der Digitalisierung der PBC-Produkte beteiligt, sondern unterstützt den Betrieb der Plattform und stellt dabei Funktionen bereit, die übergreifend verwendet werden können [26].

Identity- & Accessmanagement. Das „Identity- & Access Management“ repräsentiert [26] einen Platzhalter in der erarbeiteten RA um im weitesten Sinne Sicherheitsmechanismen zu implementieren. Sicherheit ist bei diesen Architekturen ein entscheidender Faktor [28]. Es werden große Mengen, mit zum Teil sensiblen Daten der Produkte, über einen meist öffentlichen Kommunikationskanal transportiert und verarbeitet. Diese Gruppe dient daher zur Absicherung der gesamten Plattform.

Device Management. Alle Geräte in der „Device Domain“ müssen verwaltet, überwacht und gesteuert werden [26]. Dabei muss die Möglichkeit geschaffen werden, über die Komponente „Communication & Transport“ mit den Gerätschaften und Sensoren zu kommunizieren, um bspw. stets den Überblick über die Anzahl der verbauten Geräte und Sensoren bei Kunden zu behalten, um die Devices ggf. mit Softwareupdates zu versorgen oder auch Devices falls nötig zu deaktivieren [32]. Eine weitere wesentliche Aufgabe dieser Komponente ist die Wartung der Geräte und Maschinen, die über die PBC-Plattform betrieben werden. Es geht hierbei darum, dass regelmäßig einzuhaltende Wartungsintervalle, Verschleiß an Geräten, Störungen, drohende Ausfälle sowie Defekte automatisch erkannt und an eine zentrale Stelle des Betreibers gemeldet werden, sodass Servicetechniker gezielt und mit den passenden Ersatzteilen zum richtigen Zeitpunkt bei dem Kunden vor Ort sind, um das Gerät zu reparieren und den reibungslosen Betrieb weiterhin zu gewährleisten.

Monitoring. Die Komponente „Monitoring" bezieht sich auf die PBC-Plattform selbst. Die Plattform setzt sich aus diversen Komponenten und unterschiedlichen Technologien zusammen. Damit spielt diese eine wichtige Rolle, da ohne die korrekte Funktionsweise der Plattform das Produkt für den Endkunden nicht funktioniert und SLAs nicht eingehalten werden können. Aus diesem Grund muss sichergestellt werden, dass innerhalb dieser Komponenten die Plattform selbst überwacht wird, um einen reibungslosen Betrieb derselben sicherzustellen [33]. Das Monitoring sollte dabei eine große Bandbreite von Überwachungsmaßnahmen unterstützen, die sich vom einfachen Schnittstellentest bis hin zur Überprüfung von einzelnen Servern eines Serverclusters erstrecken.

6 Zusammenfassung, Diskussion und Ausblick

Ein erstes Ergebnis dieser Arbeit ist, dass eine einheitliche Definition des Begriffs PBC derzeit nicht existiert. Weiterführende Forschungsarbeiten zu diesem Thema können dazu beitragen, herauszufinden welche Geschäftsmodelle tatsächlich zu PBC gehören und welche eigenständige oder untergeordnete Geschäftsmodelle darstellen.

Darüber hinaus konnte die Arbeit zeigen, dass die Digitalisierung des Geschäftsmodells noch am Anfang der Entwicklung steht und daher auch RAs, welche mit der in dieser Arbeit aufgestellten vergleichbar sind, zur Digitalisierung dieses Geschäftsmodells nicht oder nur teilweise existieren. Die Intention dieser Arbeit ist es, einen ersten Schritt in die Richtung einer allgemeingültigen RA für die operative Abwicklung von PBC zu unternehmen und eine Grundlage für zukünftige Diskussionen zu diesem Thema zu bilden. Da eine vielseitig einsetzbare RA erstellt werden sollte, wurden Informationen zu Produkten aus verschiedenen Branchen miteinander verglichen. Dabei wurde erkannt, dass PBC ähnliche IT-Komponenten zur Digitalisierung des Geschäftsmodells benötigt. Die anschließende Validierung der RA mittels Experteninterview hat ergeben, dass die aufgestellte RA alle Komponenten beinhaltet, die derzeit bei der Digitalisierung ihres Produktes benötigt werden.

Die durchgeführte Validierung der aufgestellten RA konnte jedoch nur für ein Produkt und eine Branche durchgeführt werden. Daher konnte mit dieser Validierung keine Allgemeingültigkeit für die aufgestellte RA nachgewiesen werden. Gleichwohl untermauert die Bestätigung des Unternehmens, die Richtigkeit und Anwendbarkeit der RA für industrielle Anlagen. Um einen besseren Nachweis der Allgemeingültigkeit zu erbringen, sollte die RA zusammen mit weiteren Unternehmen, die PBC-Produkte in IT-gestützter Form betreiben, validiert und untersucht werden.

Eine weitere Validierungsmöglichkeit besteht darin, ein neues Produkt auf Basis der RA zu erstellen, um Vor- und Nachteile der Architektur anhand eines praktischen Beispiels zu evaluieren und Schwachstellen in der Architektur aufdecken zu können.

Weiterhin besteht eine offene Frage darin, welche konkreten Technologien (Tools, Soft- und Hardwarekomponenten, Frameworks, Übertragungswege o.ä.) sich am besten dazu eignen die RA in die Praxis umzusetzen. Eine entsprechende Beurteilung ist somit ebenfalls ein mögliches Thema für weitere Forschungsarbeiten.

Auch ist es von den Autoren beabsichtigt sich bei der aufgestellten RA nur auf die operativen Aspekte und Abläufe von PBC zu beziehen. Aspekte wie die Bereitstellung von Produkten, Vertragsverhandlungen oder konkrete Wartungsprozesse sind nicht Teil der Arbeit. Die fehlenden Aspekte sind dabei aus Sicht der Autoren aber relevant und sollten in weiterführenden Arbeiten betrachtet und bewertet werden.

Wie viele andere RAs fokussiert sich die aufgestellte RA ausschließlich auf die technischen Aspekte einer Architektur. Laut Cloutier et al. [34] existieren jedoch neben der technischen Dimension noch die Dimensionen „Business Architecture“ und „Customer Context“, welche in diesem Beitrag nicht betrachtet wurden. Eine Betrachtung dieser Dimensionen in Kombination mit den Ergebnissen dieser Arbeit ist aus Sicht der Autoren sinnvoll, um das Konzept einer RA zu vervollständigen.

References

1. Selviaridis, K., Wynstra, F.: Performance-based contracting. A literature review and future research directions. International Journal of Production Research 53, 3505–3540 (2014)
2. Shick, R.A.: Government contracting. A public solutions handbook. Routledge, New York (2016)
3. Sols, A., Nowick, D., Verma, D.: Defining the Fundamental Framework of an Effective Performance-Based Logistics (PBL) Contract. Engineering Management Journal 19, 40–50 (2015)
4. Bauernhansl, T. (ed.): Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien und Migration. Springer Vieweg, Wiesbaden (op. 2014)
5. Ray, P.P.: A survey on Internet of Things architectures. Journal of King Saud University - Computer and Information Sciences (2016)
6. Demchenko, Y., Laat, C., Membrey, P.: Defining architecture components of the Big Data Ecosystem (2014)
7. NGIP: Performance Based Contracting, http://www.nigp.org/docs/default-source/New-Site/global-best-practices/performancebased
8. Glas, A.H.: Public Performance-based Contracting. Ergebnisorientierte Beschaffung und leistungsabhängige Preise im öffentlichen Sektor. Springer Fachmedien Wiesbaden, Wiesbaden (2012)
9. Jiahuan Lu: The Performance of Performance-Based Contracting in Human Services. Maryland (2014)
10. Sagiroglu, S., Sinanc, D.: Big data. A review, 42–47 (2013)
11. Pääkkönen, P., Pakkala, D.: Reference Architecture and Classification of Technologies, Products and Services for Big Data Systems. Big Data Research 2, 166–186 (2015)

12. Reichert, R. (ed.): Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie. transcript-Verl., Bielefeld (2014)
13. van Kranenburg, R., Dodson, S.: The internet of things. A critique of ambient technology and the all-seeing network of RFID. Institute of Network Cultures, Amsterdam (2008)
14. Kruchten, P.: The rational unified process. An introduction. Addison-Wesley; Safari Books Online, Boston, Sebastopol, CA (2003)
15. Henderson, J., Venkatraman, N.: Strategic alignment. Leveraging information technology for transforming organizations. IBM SYSTEMS JOURNAL, 472–482 (1999)
16. Bundesministerium für Wirtschaft und Energie: Digitale Geschäftsmodelle. Themenheft Mittelstand-Digital (2017)
17. Martin, L.: Performance-Based Contracting for Human Services. Administration in Social Work 29, 63–77 (2005)
18. Reidt, A., Krcmar, H.: Referenzarchitektur für Cyber-physische Systeme zur Unterstützung der Instandhaltung (2016)
19. Johanson, M., Karlsson, L.: Service architectures for product and production availability: A system of systems approach (2016)
20. Rosemann, M.: Komplexitätsmanagement in Prozeßmodellen. Methodenspezifische Gestaltungsempfehlungen für die Informationsmodellierung. Gabler, Wiesbaden (1996)
21. Hevner, A.: A Three Cycle View of Design Science Research (2007)
22. Webster, J., Watson, R.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quaterly, xii–xxiii (2002)
23. Smith, D.: Power-by-the-hour. The role of technology in reshaping business strategy at Rolls-Royce. Technology Analysis & Strategic Management 25, 987–1007 (2013)
24. Mauro, A. de, Greco, M., Grimaldi, M.: What is big data? A consensual definition and a review of key research topics. In:, pp. 97–104. AIP Publishing LLC (2015)
25. Dong, W., Kerrigan, M., Tsantanis, N., Grepps, P. and Nachman-Ghnassia, S.: Big Data Analytics Guidebook. Version 4.0.2, https://www.tmforum.org/resources/best-practice/gb979-big-data-analytics-guidebook-r16-5-1/#
26. Krco, S., Pokric, B., Carrez, F.: Designing IoT architecture(s). A European perspective. IEEE, Piscataway, NJ (2014)
27. Jayavardhana, G., Rajkumar, B., Slaven, M., Marimuthu, P.: Internet of Things (IoT). A vision, architectural elements, and future directions. Future Generation Computer Systems 29, 1645–1660 (2013)
28. Chen, M., Mao, S., Liu, Y.: Big Data. A Survey. Mobile Networks and Applications 19, 171–209 (2014)
29. Singh, D., Tripathi, G., Jara, A.J.: A survey of Internet-of-Things. Future vision, architecture, challenges and services. In: 2014 IEEE World Forum on Internet of Things (WF-IoT), pp. 287–292. IEEE (2014)
30. Gajurel, A.: Performance-Based Contracts for Road Projects. Springer India, New Delhi (2014)
31. NTT DATA Corporation: Big Data Reference Architecture. Ver. 1.0, http://www.nttdata.com/global/en/shared/pdf/bigdata_reference_architecture.pdf
32. Chen, G., Huang, J., Cheng, B., Chen, J.: A Social Network Based Approach for IoT Device Management and Service Composition. In: A Social Network Based Approach for IoT Device Management and Service Composition, pp. 1–8. IEEE (2015)
33. Georgiou, A.: Storing Data Flow Monitoring in Hadoop, http://cds.cern.ch/record/1596239/files/Storing_Data_Flow_Monitoring_in_Hadoop.pdf
34. Cloutier, R., Muller, G., Verma, D., Nilchiani, R.: The Concept of Reference Architectures. Wiley InterScience, 14–25 (2008)

Data Analytical Processes in IS

Teilkonferenzleitung

Christoph Flath
Hansjoerg Fromm
Ralph Grothmann
Thomas A. Runkler
Thomas Setzer

Anchor and Adjustment Detection on Empirical Forecast Series

Florian Knöll[1], David Roßbach[1]

[1] Karlsruhe Institute of Technology (KIT), Institute of Information Systems and Marketing, Karlsruhe, Germany
knoell@kit.edu, usehm@student.kit.edu

Abstract. Human involvement in forecasting processes can lead to predictions being influenced by cognitive biases. Anchoring and Adjustment is a cognitive bias that typically reduces the accuracy of forecasts. Considering the influence of forecasts in corporate processes, improving forecast accuracy can be of crucial importance. This study aims to use the newly developed Bandwidth Model to improve forecast accuracy and to detect Anchoring and Adjustment effects in forecasting processes. We demonstrate the potential of the Bandwidth Model by the regression of Anchoring and Adjustment effects on forecast errors, which exhibit higher explanatory power than standard volatility metrics. The results suggest that a forecast support system can use the Bandwidth Model to improve forecast accuracy in enterprises.

Keywords: Anchoring and Adjustment Model, Corporate Cash Flow, Forecasting Process, Forecast Revision, Cognitive Bias

1 Introduction

Cognitive biases can influence human behavior. According to [20] these biases can be divided into three groups: Representativeness, Availability, and Anchoring and Adjustment. Representativeness is a judgmental heuristic in which probabilities of events are valued based on the similarity to other well known events. Second, availability describes the heuristic that a person uses easily available examples to predict probabilities. Third, Anchoring and Adjustment takes place when a person consciously selects a certain numerical value but is influenced by environmental variables, like other predicted values. Especially Anchoring and Adjustment has a strong influence on forecast processes [8, 17, 19]. One needs constantly to account for previous forecasts within forecasts processes, as these forecasts can have influence on the new forecasts. Forecast accuracy can be improved by detecting the influence on the forecast process through Anchoring and Adjustment.

There are several authors who focus on the benefits of revisions in the forecasting process and use statistical methods to evaluate human prognoses and thus enable improvements to be made. For example, the authors of [15] state that revised forecasts are significantly more accurate than unrevised. The authors of [1] and [2] show that

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

underreacting to new information is widely spread, while [7] and [9] state that overreaction is common as well.

Other authors focus on how to determine the influence of Anchoring and Adjustment and which forecasts are affected by Anchoring and Adjustment the most [4, 10, 14, 2]. Most of them focus on single forecasts [10] or the mean of the previous forecasts [14]. The authors of [11] analyze timing, magnitude, and trend in revision processes. They state, depending on the volume of the revisions, early and later revisions can reduce forecast errors. Early revisions are beneficial for evenly distributed revisions, while in case of concentrated revisions the effect can be detrimental and later revisions can reduce error levels. The authors of [16] tested the effect of anchoring on Australian earning forecast. They state that analysts tend to overreact for positive revisions and underreact for negative ones. Results in [13] suggest that Anchoring and Adjustment patterns can occur at different levels, for example for organizational influences. Finally, the impact of a decision support system on forecast accuracy was tested by [3] and a clear positive effect was found. The restrictions of this approach are treat by authors like [6]. They analyze macroeconomic forecasts and found that even additional information is not sufficient to improve the GDP forecast.

The usefulness of Anchoring and Adjustment in the forecasting process depends on the correct detection on forecasting series. The authors of [12] developed the *Bandwidth Model (BWM)* and the *Logarithmic Bandwidth Model (LBWM)* to determine the influence of Anchoring and Adjustment on the forecast process. They tested the performance of both models on synthetic forecast processes, which generate several possible patterns that occur in real data. Both Bandwidth Models show a better performance than previously used models.

This paper aims to apply the approach on empirical forecast series. The objective is to investigate whether the model of [12] shows an improved identification of anchoring not only on simulated, but also on real forecast series. Therefore, we formulate the guiding research question:

Research Question 1: Can the Bandwidth Models be used to determine a probability for anchoring on forecast series?

The empirical forecast series originate from an international corporation that is based in Germany. The corporation's activities in various business divisions allow a better transferability of the measured effects to other areas. Further, the use of empirical data allows us to develop an *Empirical Bandwidth Model (EBWM)*. The ability of these three models to improve the precision of revisions will be compared to standard metrics of volatility.

The paper is structured as follows: Forecast process and notation is introduced in Chapter 2. Chapter 3 introduces the Bandwidth Model, the Logistic Bandwidth Model, and the Empirical Bandwidth Model. Afterwards, in Chapter 4 the outcomes of the three models are used to predict the forecast errors. Standard volatility metrics are used as a benchmark to show the impact of Anchoring and Adjustment, as well as the application of the Bandwidth Models for forecast correction.

2 Literature Review

In the literature, the methodology for determining whether a particular forecast or forecast series is influenced by Anchoring and Adjustment (A&A) has been supplemented by five important publications [4, 10, 14, 2, 12]. The author of [4] assumes that the value with the highest amplitude exerts an anchor effect. Therefore, if a proposed forecast lies between the old forecast and the forecast with the highest amplitude, an anchoring is assumed. In addition, the distance between anchor and old forecast must be significantly smaller than that between the anchor and the new forecast. The methodology of paper [2] suggests that the last forecast causes an anchor effect. A new forecast is therefore only influenced by its predecessor. These two approaches concentrate strongly on individual important forecasts with potential anchor effects. Therefore, the authors of [14] assume that not individual values trigger anchoring, but that the mean value of the previous forecasts exerts such an influence. Anchoring therefore occurs when a new forecast lies between the old forecast and the mean value of the previous forecasts. In [10] the authors assume that the direction of the adjustments of the forecasts is important, for example, if successive forecasts are adjusted upwards, anchoring can be responsible. In order to improve these identification methods, [12] has developed a new approach. This new approach, which is examined in this paper on real-world forecast data, does not consider a single forecast or parameter like the mean value, but rather each forecast to determine probabilities for the presence of anchor effects.

3 Notation and Data

The sequence of initial forecast and adjusted forecasts is referred to as *forecasting process*. For each forecast process one *actual item (A)* with the terminal realization value is provided. For each forecasting process a human expert generates the t-th forecast F_t in every quarter of the year until the actual value is realized. An example five-step forecasting process of the sample corporation with a maximum lead time of 5 quarters contains five forecasts each. Figure 1 visualizes the structure the forecasting process.

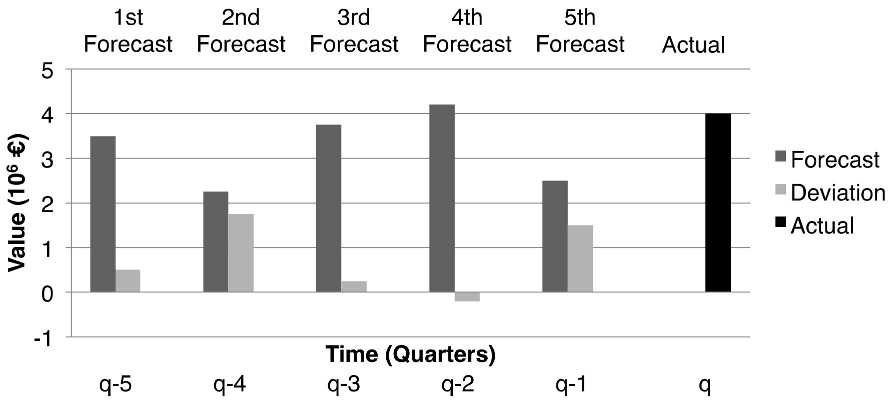


Figure 1. The Empirical Forecast Process.

Several loss functions can be implemented to measure the forecast error. Typically, the difference between the forecasts and the actual is used. However, if the values of the forecast items over different forecast processes vary widely, differences in error may be difficult to interpret and error relative to the actual volume may be preferable. The percentage error loss function takes the actual volume into account, and we calculate the percentage error between the last forecast F_{q-1} and the actual.

$$Err = \frac{F_{q-1} - A}{|A|} \tag{1}$$

The different volumes of the forecasts should also be taken into account when analyzing forecast revisions. The *Revision (Rev)*, as presented in Equation 2, is defined as the difference between two successive forecasts and the first forecast and thus avoid this problem.

$$Rev_t = \frac{F_{t+1} - F_t}{F_t} \tag{2}$$

The used data in this study is a set of forecast processes that contains real-world cash flow forecasts and realizations. It is provided by a large diversified multinational corporation that generates annual revenues in the medium double-digit billion Euro range. The data contains information of judgmental forecasts and actuals of invoice issued and received with additional information such as the currency, region, and business units. The financial cash flow data was selected in accordance with the corporation's recommendations, leading to 34.057 observations. Further, the cleaning process contains the following steps: Forecast processes with missing values or without revision volumes are removed. In addition, the upper and lower 5% percentile is withdrawn to account for outliers. Overall, the data used in the analyses bases on 26.401 observations of forecast processes, with six columns (one actual and five forecasts).

4 The Bandwidth Model

One of the previously mentioned models to detect Anchoring and Adjustment was introduced by [14]. They focus on the mean as anchor value. This model neglects strict positive or negative anchor effects. The model in [12] considers these effects with the BWM. It uses underlying patterns that are based on A&A to predict the final error. These underlying patterns are predicted through the strength of the revisions. The BWM works as follows: Each revision is assigned to one of three groups. These groups are the *Up*-Group, the *Down*-Group, and the *Const*-Group. Positive revisions above a threshold α are assigned to the *Up*-Group (value 1). Negative revisions below -α are allocated to the *Down*-Group (value -1). The remaining revisions are assigned to the *Const*-Group (value 0). The assignment to the groups is shown in Equation 3.

$$\begin{aligned} &Rev_t \in Up \;\Leftrightarrow\; Rev_t > \alpha \\ \textbf{or }&Rev_t \in Down \;\Leftrightarrow\; Rev_t < -\alpha \\ \textbf{or }&Rev_t \in Const \;\Leftrightarrow\; |Rev_t| \leq \alpha \end{aligned} \tag{3}$$

Allocation to the three groups is based on a threshold value that ensures low susceptibility to minor changes in the forecast. As a result the selected groups focus on major changes. The categorization into three groups, as an important part of the model, restricts the model simultaneously. Assigning to three groups without further differentiation, neglects the differences within these groups. Further, the selected α has a high influence on the results. Choosing an optimal α value is crucial to improve the detection of A&A.

Two further models will be introduced to overcome this limitation. One of these models was introduced in [12]. These two models base on the allocation to different groups, which complement the assignment of functions for further differentiation of revisions. Therefore, the requirements for the assignment function are examined in more detail before the models are introduced. The assignment functions needs to fulfill the Equations 4-6.

$$f^+(0) = 0 \tag{4}$$

$$f^+(\max\ Rev_t) = 1 \tag{5}$$

$$f^{+\prime}(Rev_t) \geq 0, \forall\ Rev_t \in\ Rev^+ \tag{6}$$

The functions value should be 0 for a revision where no adjustment has taken place. The smallest revision should further be assigned to 100% to the *Down*-Group, the largest to 100% to the *Up*-Group. The function should monotonically increase as a higher revision should result in higher assigned values. As assigning functions that can fulfill these conditions, the logarithmic growth function and the empirical distribution function will be examined.

The first model to be considered with an assignment function is the LBWM. This model uses the logarithmic growth function, which induces a sigmoid process. This sigmoid process assigns small anchor probabilities for small revisions (*Rev* close to zero), and assigns large probabilities otherwise. Near the threshold α the transition is steady. The BWM categorizes revisions in three different groups. However, the logarithmic growth function distinguishes between positive and negative revisions.

According to the Conditions 4-6 the revisions of 0 will be assigned to the positive group. The logarithmic growth function will be modified in a way that a weight on every revision can be assigned.

Equation 7 shows the standard logarithmic growth function f_x. The function depends on the saturation limit G, a parameter k influencing the strength of the growth, the functions values for revisions of size 0, and the exponential function *exp*.

$$f_x = \frac{G}{1+exp^{(-kGx)}(\frac{G}{f(0)} - 1)} \tag{7}$$

First, the saturation limit G is set to 1. In addition, the function should assume its turning point at the threshold α equal 0.5. For this purpose, the parameter k is determined as shown in Equation 8.

$$k = \frac{ln(\frac{1}{\mu} - 1)}{\alpha} \tag{8}$$

The parameter μ is introduced to ensure that the function does not convert to zero and the parameter α shifts the Equation 7. The overall effect of both parameters is stated in Equation 9.

$$f(0) = 0.5 \textbf{ and } f(\text{-}\alpha) = \mu \tag{9}$$

Equations 10 and 11 present the final function for the LBWM. A distinction is made between positive and negative revisions, which reflect the sign before the revision by means of two different functions. These two equations are obtained through a re-shift of the function with –α.

$$f_{log}^{+}(Rev_t) = \frac{1}{1+exp^{\left(\frac{\frac{\ln(\frac{1}{\mu}-1)}{\alpha}}{(-Rev_t+\alpha)}\right)}} \tag{10}$$

$$f_{log}^{-}(Rev_t) = \frac{1}{1+exp^{\left(\frac{\frac{\ln(\frac{1}{\mu}-1)}{\alpha}}{(Rev_t+\alpha)}\right)}} \tag{11}$$

Figure 2 shows an example of how the LBWM can assign revisions (x-axis) to different probabilities for anchoring (y-axis). The figure provides the combination of the formulas for positive and negative revisions. For example, the revision $Rev_t = 4$ would be considered as effected by anchoring with a probability of about 43%.

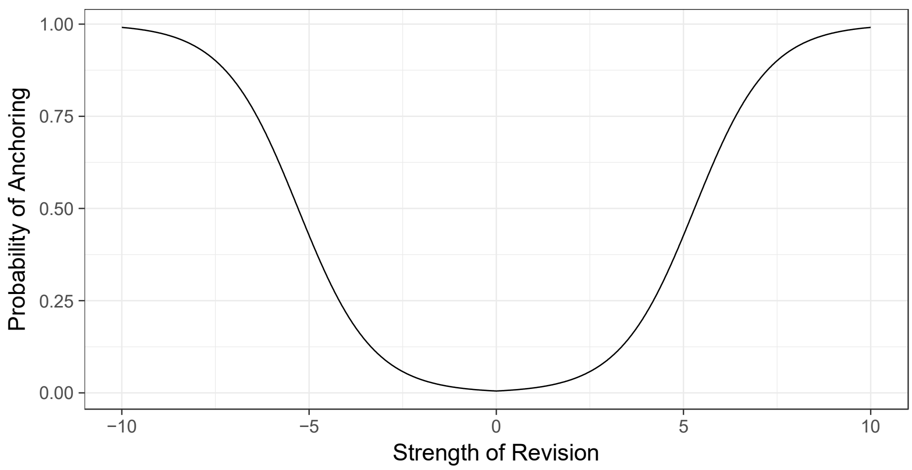


Figure 2. Logistic Function for Positive and Negative Revisions.

The second model with an assigning function is the Empirical Bandwidth Model. The empirical distribution function in this model complies with the Equations 4-6, which is why no further adjustments are necessary. The empirical distribution function assigns revisions in two different groups. The function assigns different revisions to each pair of unequal function values. The function grows in dependence to the sample size. The results of the empirical distribution function depend on the quality and amount of the data basis. The more values the data basis contains, the closer the function is to a continuous one. The dependency on the quality of the data is eminent. The combination of a small data base and poor data can lead to wrong predictions. The data set is thus used entirely and is not split into training and test data.

In total, three different models for the detection of A&A effects were explained. The Bandwidth Model classifies all forecasts in three different groups. The Logistic Bandwidth Model and the Empirical Bandwidth Model consider intergroup differences. The Logistic Bandwidth Model complemented the forecasts by a weight for A&A effects and classifies them in two different groups. The Empirical Bandwidth Model assigns weights based on the range of the forecast and optional training data. These three models will be evaluated in a next step. This evaluation is performed by predicting the forecast error through the models. As a benchmark, regressions with several standard measures for volatility will be used.

5 Empirical Analysis

The evaluation uses the forecast processes of the international corporation to compare regression results of the three models with the results of the benchmark models, intending to show potentials for forecast correction. The benchmark models extract several standard metrics for volatility for revisions, such as the range of the revisions, the sum of the amount of volatility, or the highest absolute revision. The different benchmark models use linear regressions on the forecast series. The values of these series are used to predict the *Err*. Table 1 shows the results of several standard metrics for volatility of the cleaned dataset with 26401 items. For example, the regressions consider revisions with the highest amplitude, or the range between the largest and smallest revision, or the squared revisions.

Table 1. Regression Results of Volatility Metrics.

Model	Intercept	Coefficients	Adj. R^2
$\lvert Err\rvert \sim \beta_0 + \Sigma_{i=1}^{4} \beta_i \lvert Rev_i\rvert$	0.38***	-1.61×10^{-5} -1.80×10^{-5} 4.66×10^{-5} -4.67×10^{-6}	-5.93×10^{-5}
$\lvert Err\rvert \sim \beta_0 + \beta_1 \max \lvert Rev_i\rvert$	0.38***	1.12×10^{-5}	-2.96×10^{-5}
$Err \sim \beta_0 + \beta_1 \{Rev_i : \max\lvert Rev_i\rvert\}$	0.02	-7.76×10^{-6}	-3.55×10^{-5}
$\lvert Err\rvert \sim \beta_0 + \beta_1 \sigma\,(\cup\, Rev_i)$	0.38***	9.91×10^{-6}	-3.55×10^{-5}
$\lvert Err\rvert \sim \beta_0 + \beta_1 (\max\lvert Rev_i\rvert - \min\lvert Rev_i\rvert)$	0.38***	5.27×10^{-6}	-3.51×10^{-5}
$\lvert Err\rvert \sim \beta_0 + \Sigma_{i=1}^{4} \beta_i\, Rev_i^2$	0.38***	-1.61×10^{-5} -1.80×10^{-5} 4.66×10^{-5} -4.67×10^{-6}	-5.92×10^{-5}

* indicates a significance level of 0.1, ** of 0.05, and *** of 0.01

The regression results on volatility metrics exhibit small revision coefficients, which are all not significant to the 0.1 p-value. The best metric achieved an Adj. R^2 (share of explained variation, devalued by the number of components) of -2.96×10^{-5}. This metric uses the strongest revision of a forecast series as predictor for *Err*. Overall, none of the standard metrics provide an Adj. R^2 sufficiently high to improve forecasts. However, the results of the metrics can be used as a benchmark for the Bandwidth Models.

The first model applied to the forecast sequence is the BWM. In the initial paper [12] uses a fixed value of 0.5 for parameter α, which is adopted here. Equation 12 regresses the *Err* to examine the explanation performance (Adj. R^2) of the linear regression.

$$Err \sim \beta_0 + \sum_{t=1}^{4} \beta_t Up_t + \sum_{t=1}^{4} \gamma_t Down_t + \varepsilon \qquad (12)$$

Apart from the intercept, both coefficients are significant. The Adj. R^2 of the BWM on the empirical forecast series is 0.03. The result is shown in the second column of the Table 2.

The Logarithmic Bandwidth Model is the first evaluated model with the logistic assigning function. The LBWM replaces the *Up* and *Down* terms in Equation 12, which is shown in Equation 13. The influence is measured with a linear regression.

$$Err \sim \beta_0 + \sum_{t=1}^{4} \beta_t f_{log}^{+}(Rev^{+}) + \sum_{t=1}^{4} \gamma_t f_{log}^{-}(Rev^{-}) + \varepsilon \quad (13)$$

In a practical implementation, the optimal α value can be estimated with historical training data. The adjustment of the α parameter can improve the results of the Bandwidth Models. Two possible ways to predict the α parameter were considered. Either α is the same for all revisions or a different α is used for earlier forecasts in a series than for later ones. For the case of equal α values for revision we optimize them through the algorithm *BFGS-B* [5]. BFGS-B is an iterative method for solving unconstrained nonlinear optimization problems. Figure 3 shows the effects on the Adj. R^2. The optimal value is reached for α = 0.36.

Table 2. Regression Results of the Bandwidth Models.

Coefficients	BWM	LBWM: Equal α	LBWM: Different α	EBWM
β_0	0.02	0.02***	0.01***	0.00***
β_1	0.12***	0.03***	0.04***	0.03***
β_2		0.07***	0.07***	0.07***
β_3		0.16***	0.15***	0.13***
β_4		0.31***	0.30***	0.25***
γ_1	-0.13***	-0.04***	-0.04***	-0.03***
γ_2		-0.07***	-0.06***	-0.05***
γ_3		-0.15***	-0.14***	-0.11***
γ_4		-0.28***	-0.36***	-0.19***
$Adj. R^2$	0.03	0.07	0.08	0.06
* indicates a significance level of 0.1, ** of 0.05, and *** of 0.01				

The optimal α values for different α for all forecasts are shown in Table 3. Table 2 shows the result of both variants of the LBWM. Column three shows the LBWM with one optimized α value and column four shows the LBWM with different α values for each revision. The LBWM with one specific α reaches an Adj. R^2 of 0.07. The part of the explained variance appears promising when compared to the results of [18]. The authors asked amateurs to estimate the value of houses with a given anchor value. At least 17% of the variance could be explained by the anchor value. However, this is the determination of a single anchor value. In contrast to their work, our forecasts are not single anchor values, but series of anchored forecasts. For this reason, an Adj. R^2 of 0.07 appears promising. As expected, the coefficients for positive revisions are positive and the negative revisions are negative. The influence of the coefficients increases over time, stating that newer forecasts provide more information. Moreover, the intercept is significantly different from zero.

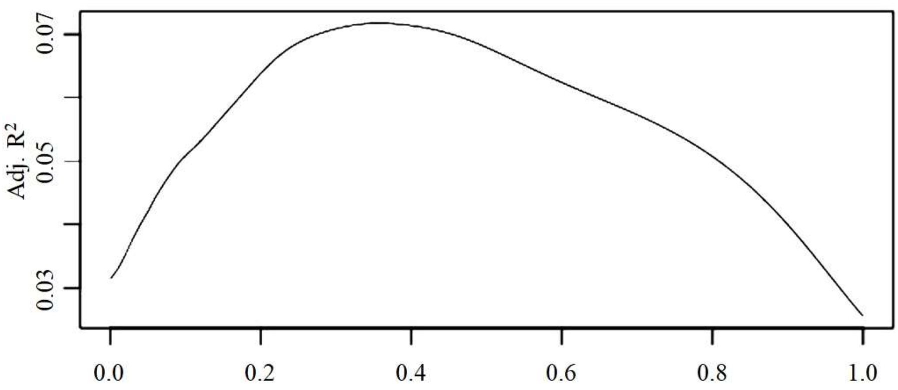


Figure 3. Comparison of α Values.

The LBWM with different α values further improve the results to an Adj. R^2 of 0.08, while the influence of the coefficients increase over time. Further, the α show a positive trend. The comparison with the standard metrics for volatility shows the potential (Adj. R^2) of the LBWM is higher.

Table 3. Optimization of α Values for the Logarithmic Bandwidth Model

Coefficients	β_1	β_2	β_3	β_4	γ_1	γ_2	γ_3	γ_4
α Values	0.36	0.23	0.25	0.31	0.39	0.26	0.33	0.51

The Empirical Bandwidth Model is the second version of a model with assignment function after the LBWM that uses the empirical function to assign the *Up* and *Down* groups. Equation 14 shows the regression model of the BWM extended upon the EBWM. The structure of the results is equal to the results for the Logistic Bandwidth Model.

$$Err \sim \beta_0 + \sum_{t=1}^{4} \beta_t f_{emp,t}^{+}(Rev^{+}) + \sum_{t=1}^{4} \gamma_t f_{emp,t}^{-}(Rev^{-}) + \varepsilon \qquad (14)$$

The results in Table 2 show an increase of the coefficients for later revisions. Positive revisions result in positive coefficients and negative revisions in negative coefficients. The Adj. R^2 of the EBWM is 0.06, which is below the values of the LBWM. However, an increase of the sample size is expected to further improve the Adj. R^2 of the EBWM.

Overall the BWM shows better results than the benchmarks. The results of [12] were thus also confirmed on empirical forecast series. However, BWM was surpassed by the LBWM and the EBWM. Especially the LBWM with different α values shows a far higher Adj. R^2 than the standard metrics and a slightly higher value than both other Bandwidth Models. The detected Adj. R^2 of 0.08 shows the importance of A&A effects on forecast series to predict the forecast error. However, the difference of the enhanced models is not very high, as is the influence of α optimization. This suggests that the used assigning function is not as important as the fulfillment of the assigning functions Conditions 4-6. In the work, the regressions use all available information (revisions) to estimate the anchor results, i.e. it is an in-sample test. Thus, the paper shows the improvement of recognizability by optimizing alpha values, while both LBWMs (without and with optimization) achieve comparable Adj. R^2 results. Further evaluations and out-of-sample tests of the forecast series could be carried out in order to ensure the stability of alpha values on the forecast processes. The specific optimization must then take place in the respective context, which is derived from the generalizability of the methodology.

The LBWM and the EBWM tend to show a similar process for the revisions independent from the used assigning function. Positive revisions have positive coefficients and negative revisions always negatives. This suggests that negative revisions of experts tend towards an underestimation whereas positive revisions tend towards an overestimation.

6 Conclusion and Outlook

In this paper we analyzed three models to determine A&A effects on forecast series - the Bandwidth Models: BWM, LBWM and EBWM. These models can relate forecast revisions to the influence of A&A effects. Using this relation, the forecast models provide input for linear regressions to estimate the forecast error. In comparison, the forecast errors were regressed with standard metrics for volatility of the forecast processes. The evaluation based on short empirical forecast series of a large corporation, where forecast accuracy is essential for corporate success.

The results show that all Bandwidth Models exhibit a higher explanatory power on the forecast error than the standard metrics. The increased Adj. R^2 implies that the forecast accuracy is influenced by cognitive biases. Utilizing the information of detected A&A effects for human forecast processes in the case of real-world forecast improved the regression results for error estimation. As a result, the research question can be answered positively. The Bandwidth Model can be used to determine probabilities for anchoring.

This study has some limitations. The introduced LBWM and EBWM were tested on one particular set of empirical corporate data. However, we recommend the analysis of the method on further data samples - and, for example, in other domains.

Overall, we recommend evaluating the Bandwidth Models in a Forecast Support System. The usage of the error regressions that utilize the Bandwidth Models offers a high potential of forecast improvement by the correction of the A&A bias. The correction can be employed within an information system such as a forecasting support system. The forecast support system can identify forecasts that exhibit a high probability for *A&A*. The managerial analysis of these identified forecasts allows experts to optimize forecast revisions and to improve managerial processes.

References

1. Abarbanell, J.S., Bernard, V.L.: Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior. The Journal of Finance 47(3), 1181–1207 (1992)
2. Amir, E., Ganzach, Y.: Overreaction and Underreaction in Analysts' Forecasts. Journal of Economic Behavior & Organization 37(3), 333–347 (1998)
3. Arnott, D., O'Donnell, P.: A Note on an Experimental Study of DSS and Forecasting Exponential Growth. Decision Support Systems 45(1), 180–186 (2008)
4. Bromiley, P.: Do Forecasts Produced by Organizations Reflect Anchoring and Adjustment? Journal of Forecasting 6(3), 201–210 (1987)
5. Byrd, R.H., Lu, P., Nocedal, J., Zhu, C.: A Limited Memory Algorithm for Bound Constrained Optimization. SIAM Journal on Scientific Computing 16(5), 1190–1208 (1995)
6. Cho, D.W.: Do Revisions Improve Forecasts? International Journal of Forecasting 18(1), 107–115 (2002)
7. Clements, M.P.: Rationality and the Role of Judgement in Macroeconomic Forecasting. The Economic Journal, pp. 410–420 (1995)
8. Dalrymple, D.J.: Sales Forecasting Practices: Results from a United States Survey. International Journal of Forecasting 3(3-4), 379–391 (1987)
9. Ehrbeck, T., Waldmann, R.: Why are Professional Forecasters Biased? Agency versus Behavioral Explanations. The Quarterly Journal of Economics 111(1), 21–40 (1996)
10. Harvey, N., Bolger, F., McClelland, A.: On the Nature of Expectations. British Journal of Psychology 85(2), 203–229 (1994)
11. Knöll, F., Dorner, V., Setzer, T.: Relating Cash Flow Forecast Errors to Revision Patterns. In: Relating Cash Flow Forecast Errors to Revision Patterns, pp. 1217–1228. MKWI - Prescriptive Analytics in IS, Universitätsverlag Ilmenau (2016)
12. Knöll, F., Roßbach, D.: The Anchor and Adjustment Bandwidthmodel: Exploring Cognitive Biases in Simulated Forecast Series. In: 51th Hawaii International Conference on System Sciences (HICSS), 2018. IEEE (2018)
13. Knöll, F., Setzer, T., Laubis, K.: On Predictability of Revisioning in Corporate Cash Flow Forecasting. In: 51th Hawaii International Conference on System Sciences (HICSS), 2018. IEEE (2018)
14. Lawrence, M., O'Connor, M.: Exploring Judgemental Forecasting. International Journal of Forecasting 8(1), 15–26 (1992)
15. Lim, J.S., O'Connor, M.: Judgmental Forecasting with Time Series and Causal Information. International Journal of Forecasting 12(1), 139–153 (1996)
16. Marsden, A., Veeraraghavan, M., Ye, M.: Heuristics of Representativeness, Anchoring and Adjustment, and Leniency: Impact on Earnings' Forecasts by Australian Analysts. Quarterly Journal of Finance and Accounting, pp. 83–102 (2008)
17. Mentzer, J.T., Cox, J.E.: Familiarity, Application, and Performance of Sales Forecasting Techniques. Journal of Forecasting 3(1), 27–36 (1984)
18. Northcraft, G.B., Neale, M.A.: Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions. Organizational Behavior and Human Decision Processes 39(1), 84–97 (1987)
19. Phillips, L.D.: A Theory of Requisite Decision Models. Acta Psychologica 56(1), 29–48 (1984)
20. Tversky, A., Kahneman, D.: Judgment under Uncertainty: Heuristics and Biases. In: Utility, Probability, and Human Decision Making, pp. 141–162. Reidel (1975)

Value Based Pricing meets Data Science: A Concept for Automated Spare Part Valuation

Clemens Wickboldt[1], Natalia Kliewer[2]

[1] Freie Universität Berlin, Professur für Wirtschaftsinformatik, Berlin, Germany, clemens.wickboldt@fu-berlin.de
[2] Freie Universität Berlin, Professur für Wirtschaftsinformatik, Berlin, Germany, natalia.kliewer@fu-berlin.de

Abstract. Turning data into value is an exciting challenge for Data Science in times of an exponential growing amount of data. Maintenance, Repair and Overhaul companies are facing pricing related decision problems on a daily basis. The industry sits on vast amount of data. Due to lacks of transparency in the surplus part market and missing concepts to efficiently use internal data, existing information is not used exhaustively to improve data-based part utilization decisions. An early-stage concept for automated spare part valuation which classifies pricing data before applying appropriate valuation methods is presented and hereby combines methods from multiple disciplines. Information from heterogeneous sources is aggregated, transformed and then supports machine learning methods to automatically determine a Fair Market Value for surplus spare parts. Handling incomplete historical data sets as well as validating the calculated Fair Market Value are some of the challenges which become visible.

Keywords: Automated Value Determination, ETL, Data Classification, Spare Parts, Fair Market Value

1 Introduction

'We are drowning in information but starved for knowledge.' stated author John Naisbitt in his book Megatrends in the year 1984 [1]. Until today, decision makers still complain in a very similar way about 'drowning in data while thirsting for information' [2]. As this is a general problem, it's especially true for market participants with a vast number of different products, for example the aircraft spare parts industry. The UK's Royal Airforce alone is managing around 685.000 line items [3], the largest commercial airplane Airbus A380 consists of around 2.5 million individual parts[1]. The aircraft parts industry is an especially interesting example since aircraft maintenance has to face high safety regulations but also operate in a very short timeframe to avoid very expensive aircraft downtime. Maintenance Repair and Overhaul (MRO) companies, especially in the aircraft industry, are constantly confronted with managing the logistics of spare parts. An inexpensive repair and maintenance in a timely manner requires an efficiently

[1] Airbus Press Office 2017: Facts & Figures, June 2017

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

organized material supply. It cannot always be resorted to new parts, either for cost reasons or because certain parts are not available on the market. Additionally, constant pulling and replacing of parts leads to a high number of used parts in stock which not always have a direct use, in other words they are surplus. The market for aircraft surplus spare parts is characterized by a great number of different products but just a few market participants. Transactions are done via e-mail or telephone. Although trading platforms are listing offers for some parts, prices are rarely published. The compartmentalized structure of traders and MRO companies in the aircraft industry may be one reason why there is still no end-to-end auction based market place for surplus parts. Valuating spare parts on the market or in stock in this environment is done today in a time consuming and error-prone manual manner by industry experts.

There are several ways to determine a price as foundation for the part utilization decision, including auction based market places and prediction models based on historical data. As soon as context based value information is needed for internal decision making, external determined prices fail to include company-specific information and don't provide a sufficient basis for decision-making. Predicting a value based on historical data however may work well for market players with a huge amount of structured and high quality historical data. In this paper, we propose a concept for combining methods from multiple disciplines with the goal of automated part valuation. The aircraft industry is used as an example. This work addresses problems of the MRO industry in general. Existing pricing approaches which had been already discussed in [4] are shortly described in Section 2.1. In cooperation with a large aircraft MRO provider, we gained insights in existing spare part utilization processes which are described in Section 2.2. In Section 3 Pareto Principle and Product Lifecycle are used to classify pricing data which is a basis for an extension of the developed system architecture for automated spare part valuation (ASPV) of [4]. Section 3 closes with approaches for handling the validation of the resulting values. Section 4 summarizes this work and gives an outlook for what will be discussed during following research.

2 State of the Art of Pricing and Surplus Part Supply Chain

This section gives an overview about existing pricing approaches and the surplus spare part supply chain to understand the need for a part's value as basis for the utilization process.

2.1 Pricing

It has been shown that price has the greatest impact on a company's earnings before interest and taxes [5]. Existing literature provides approaches for pricing. Before proceeding to automatically finding a surplus part's value, a short overview about the underlying terminologies is given. According to [6], '*Price* in business market is what a customer firm pays a supplier for its product offering'. Furthermore, [7] state, '*Value* in business markets is the worth in monetary terms of the technical, economical service and social benefits a customer company receives in exchange for the price it pays for a market offering'. Pricing approaches can be categorized into three different groups [8]:

Cost Based Pricing uses data from cost accounting and for example considers original purchase price and internal costs (e.g. repairing) to determine a selling price. It doesn't take the competition or market into account and is therefore considered as the weakest approach.

Competition Based Pricing observes price levels of the competition and uses market prices for orientation in price setting but dismisses company-internal information such as inventory or repair costs of defective spare parts.

Value Based Pricing uses a predefined value as the basis for determining the price. As only value based pricing takes company internal information into account, this approach is now described in more detail to discuss its potential to serve as a foundation for decision making and price setting.

This work focuses on a business to business secondary market for surplus parts which mainly differentiates between the following more granular value terms. *Base Value* is considered as the economic value and assumes balanced supply and demand [9] as well as completely informed market participants and thus is considered as a hypothetic value [10]. *(Current) Market Value* is defined as determined value after manual analysis [10] or the 'most likely trading price' and is used synonymously with Fair Market Value [9]. [11] state that the *Fair Market Value (FMV)* is '(..) the price at which the property would change hands between a willing buyer and a willing seller when the former is not under any compulsion to buy and the latter is not under any compulsion to sell, both parties having reasonable knowledge of the relevant facts'.

Still, many companies are not able to benefit from Value Based Pricing. One obstacle is that the value has to be determined before it can be used as an argument for pricing [8]. Value Based Pricing and value estimation methods are often related to product introductions in primary markets. State-of-the-art literature about Value Based Pricing takes the customer point of view for value determination. Common methods are surveys or conjoint analysis [12]. This is applicable to this scenario in a limited way only. Conducting surveys for this large quantity of different products seems to be a disproportionate effort. Also, as soon as context based values are needed, external determined prices fail to include company-specific information and don't provide a sufficient foundation for a company's decision-making. Therefore, an approach is needed which includes company internal information and is feasible to a large number of parts.

2.2 Understanding the Spare Part Supply Chain

As pointed out in [13], products at the end of the traditional supply chain may be surplus but still have value. As shown in Figure 1, surplus parts are stored, depending on their value.

Expert interviews with a large MRO provider showed that parts with value but no internal demand will be put on a trade stock and sold to the surplus market via direct sale or auction. From the same surplus market, parts could be bought (B). If there is internal demand they are stored in a pool stock for (future) replacement (R). The pool stock could also be filled by parts pulled out of the operations, such as an aircraft or production facility (P). If they have no value at all, they will be scrapped. Because of the fact that parts circulate in a loop but not return to the manufacturer, this process is called Alternative Closed Loop Supply Chain Process.

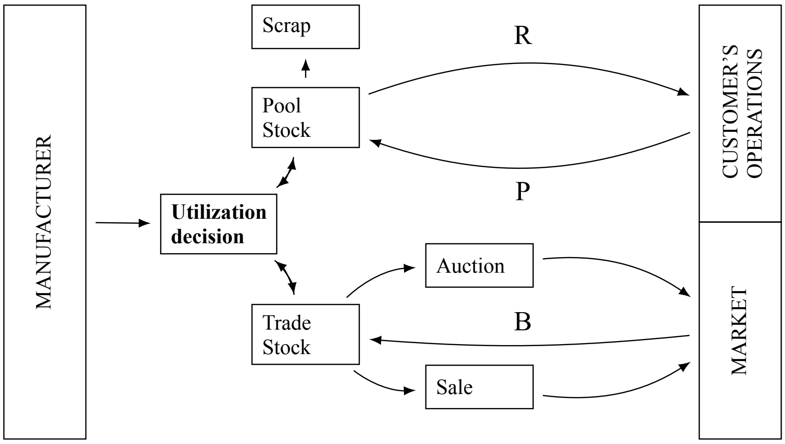


Figure 1. Alternative Closed Loop Supply Chain Process from a MRO point of view

Studies have shown that simple disposal of parts which have failed or appeared to have failed is a major cost contributor within the whole supply chain [13]. In this situation, reverse logistics are important from an economic[2] and ecological[3] point of view. [13] also points out, that extra revenues are realized if there is an ability to liquidate parts with value in secondary markets. MRO companies therefore need to estimate a value as characteristic parameter and foundation for price derivation in order to support internal decision-making processes.

3 Framework for Automated Spare Part Valuation

From characteristics of the surplus market, an overview of relevant pricing literature and value definitions, the following research goals result:

- Design a framework in order to automatically determine a FMV for surplus spare parts and present it to the end user in a meaningful way (*ASPV framework* in following).
- Find a way to prepare data as foundation for automated valuation of surplus spare parts.
- Validate the calculated FMV.

The resulting ASPV framework should be able to determine a part's FMV under consideration of the company's specific situation. The calculated value contains

[2] Valuable parts will not be disposed but could be used or sold.
[3] Probably fewer parts will be disposed if the value is recognizable.

information as displayed in Table 1. The variables are categorized by time (historic and current) and by point of view (company-internal and -external).

Table 1. Variables for FMV calculation

Time	*Internal*	*External*
Historic	Purchase price, offer price, selling price	Original launch price
Current	--	Market price

We suggest a combination of a set of data analytics methods to transparently determine and present a FMV. Based on the results from [14], proposing prediction methods for determining prices for the used car market, an application of prediction methods seems to be promising in order to evaluate secondary market items. There is a wide variety of machine learning methods, from simple Linear Regression Models to Neural Networks which may perform differently in this specific scenario[4].

The ASPV framework differs from state-of-the-art Value Based Pricing methods as shown in Table 2. State-of-the-art Value Based Pricing and value estimation methods are related to product introductions in primary consumer markets. The proposed ASPV framework focuses on business to business secondary markets in the MRO industry. State-of-the-art Value-Based Pricing focuses on an individual consideration of a small number of products which is an obvious scope for pricing related customer surveys and conjoint analysis. On the contrary, this work aims to automated valuation of thousands of different products from multiple variables (see Table 1), present the valuation in a transparent way and in a second phase increase objectivity in contrast to the status quo results from manual value determination.

Table 2. Differentiation from state-of-the-art Value Based Pricing methods

Dimension	*State-of-the-art Value Based Pricing*	*ASPV framework*
Target	Primary markets	Secondary market
Audience	B2C	B2B
Scope	Small number of products	Thousands of products
Automation	Individual consideration	Automated valuation

3.1 Classification of Pricing Information

As Value Based Pricing is effort intensive, it should not be used on all products. To address the effort issue, we propose a rule-based segmentation of the products based on their need for an approximation on the one side or for exact value determination on the other side. To reduce complexity in the classification stage we prefer this rule-based approach over machine learning.

[4] An extensive overview can be found in 15. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics, 2009.

[16] take a closer look at the Pareto principle which led to a classification of products into A, B and C groups. The 'A' group (the 'vital few'), consisting of approximately 20% of the attributes (surplus spare parts in this case), accounts for 80% of the phenomenon (value in this case); the 'B' group, i.e. the next 30% of the items, accounts for 10% of the phenomenon, and the 'C' group (the 'trivial many'), which contains 50% of the items, accounts for also 10% of the phenomenon. Given that, [17][5] recommend the following procedure:

- **Classification** – Find attributes for sorting products into A, B and C groups,
- **Differentiation** – set differentiation policy for each class,
- **Allocation** – allocate (pricing) effort according to classification and differentiation.

Based on the Pareto principle and illustrated in Table 3, category A parts which make about 20% of the parts but are responsible for 80% of the inventory's value would be valuated via the following ASPV framework. Category B and C parts which only contribute a combined 20% of the value but 80% of quantity could be just valued via market data. If competition based pricing is not feasible because of a lack of data, there is no alternative option than using any internal price information as basis for the FMV.

Table 3. Classification prior to choosing valuation method

Category	*Amount*	*Value*	*Method*
A	20%	80%	Valuation via ASPV framework
B	30%	10%	Competition Based Pricing
C	50%	10%	Competition Based Pricing

Another perspective on product importance is by classification into stages of a product lifecycle. [18] gives an overview of a product's lifecycle progress in four phases, beginning with an introduction phase where the amount of sold products grows slowly. The following growth phase is characterized by even more rising sales. Due to market saturation or introduction of improved products, sales reach a plateau and later on decline until obsolescence. [19] extend this concept and state that the number of new products and the demand for remanufactured products or spare parts is positively correlated but also delayed in time. As spare parts are removed some time later from operations, the number of parts on the surplus market peaks after the maturity phase of new parts (see Figure 2).

To refine the allocation of parts, the consideration of the part's product lifecycle is helpful. Products wander through a product lifecycle which is characterized by a dependence from product value to demand and distribution on the market. The value of the parts is very dependent on the product lifecycle and therefore the demand of the underlying operations. Nevertheless, the product lifecycle of the new part can be an estimate of the demand for the spare part and therefore an indicator for the part's importance for valuation.

[5] Quoted from 16. Grosfeld-Nir, A., B. Ronen, and N. Kozlovsky, *The Pareto managerial principle: when does it apply?* International Journal of Production Research, 2007. **45**(10): p. 2317-2325..

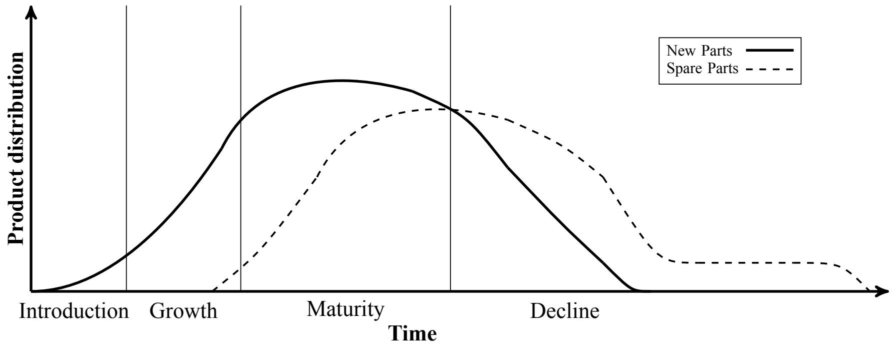


Figure 2. Product lifecycle based on [19]

As extension to the procedure of [17] and in the context of this scenario, the following procedure is therefore recommended:

- **Classification –** Find attributes for sorting products into A, B and C groups,
- **Differentiation –** set differentiation policy for each class,
- **Refine –** differentiate further by putting in perspective of product lifecycle,
- **Apply –** use ASPV framework for A parts and competition based pricing for B and C products,
- **Revise –** repeat process regularly based on improved data.

3.2 ASPV Framework

Even with a smaller number of parts for valuation, the process from raw data to information is a walk through the jungle of data analytics methods. An extension to the existing Extract Transformation and Load (ETL) process of [20] is proposed to address the particularities of pricing data and provide a system for method evaluation. The system architecture of the ASPV framework is visualized in Figure 3.

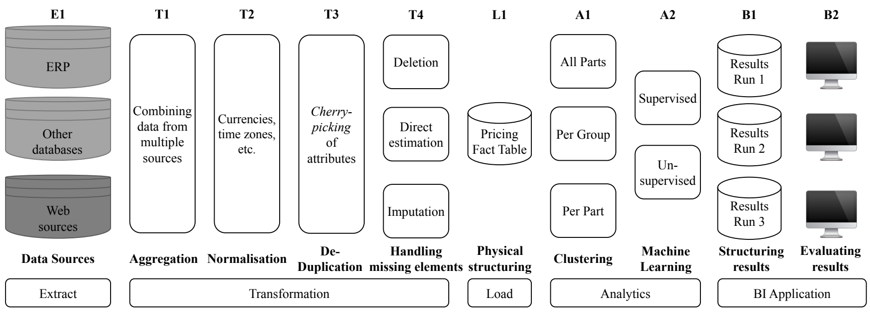


Figure 3. ASPV framework

The first step is the extraction of raw data from available data sources which include internal ERP systems, other internal databases and web sources from trading platforms

(E1). Every data source contains complementary pricing information, that is why data pre-processing is of high importance prior to applying machine learning methods. Extracted data is now aggregated (T1) and subsequently normalized concerning currencies, time zones etc. (T2). The resulting datasets containing pricing information are now uniquely distinguishable by the combination of part number and date.

	partNo	partDate	partGroup	purchasePrice	offerPrice	sellingPrice	originalLaunchPrice	marketPrice
1	(string)	(datetime)	(string)					(double)
2	(string)	(datetime)	(string)				(double)	
3	(string)	(datetime)	(string)	(double)				
4	(string)	(datetime)	(string)		(double)			
5	(string)	(datetime)	(string)			(double)		

1	(string)	(datetime)	(string)	(double)	(double)	(double)	(double)	(double)

Figure 4. Deduplication of aggregated data

As there are usually multiple pricing points per part, duplicative datasets are combined (T3) as shown in the example in Figure 4. Multiple incomplete datasets are combined to one complete dataset. In case of multiple available data, aggregation or selection methods need to be selected according to the business process environment. Still missing elements need to be handled (T4). Two simple and widely used concepts are listwise and pairwise deletion. In case of listwise deletion a complete record is deleted when at least one attribute is missing. This might be the quickest and most effortless approach but also loses much data. Pairwise deletion on the other hand doesn't always ignore incomplete datasets but considers records with connected attributes. Consider the following example: the correlation between attribute A and attribute B is calculated. Pairwise deletion takes all records (including incomplete) into account where attribute A and B still exist. A detailed overview about handling missing data methods is given in [21]. The result of the preceding transformation is a clean pricing fact table (L1) which is the basis for applying analytics methods. Now parts are combined in clusters of different size (A1). Clustering depends on the underlying characteristics of the parts, e.g. number and size of part groups. Machine Learning methods can then be applied on different aggregation levels (A2). Resulting is a FMV per part number and part date (B1). Finding a suitable machine learning method depends on the underlying data structure. The framework's performance can be evaluated by repeatedly running the process in multiple combinations from E1 to A2 and comparing the results (B1) in B2.

3.3 The Challenge of Validation

After collecting the results, the question of validation still stands out. In which combination of methods does the ASPV framework determine a good or correct FMV as target variable? Which FMV is the best or what is the target number of the FMV? We propose a two-step process for using the ASPV to find a FMV: First, automate the valuation process, second increase objectivity of the automated valuation to improve the status quo.

Supervised machine learning methods could be used to automate the valuation process. Predicted FMVs are validated with a subjective opinion of a domain expert. By that the prediction model will be trained to deliver the same quality as the domain expert.

To realize a substantial contribution to existing approaches the ASPV has to be more objective than the domain expert. The FMV should be in context to the input data which are the basis for FMV calculation. In reality, that might lead to choosing a value between reference values such as purchase and selling price. As this seems straightforward for validation, the question may arise, why taking all the effort and complicated models to calculate a FMV to, in the end, simply taking a mean value as reference. Achieving higher objectivity could be realized by deriving a model which leads to a FMV-based decision which contributes most to the overall business profitability. A FMV-based decision which achieves a more profitable business leads to the conclusion that this FMV was correctly determined regarding the definition given by [11]. This determination could be done by comparing time periods before using the FMV and periods while using the FMV for decision making. The challenges coming with this approach are:

1. Separating the FMV from its exogenous influence factors such as
   (a) macroeconomic factors like overall economic growth,
   (b) general profitability within the industry sector,
   (c) general market conditions,
2. and then again, handling emerging missing values which might prevent this separation.

The third approach is not only the most complicated to realize but also the most objective.

4 Conclusion and Outlook

To find methods to determine a surplus spare part's value, existing pricing literature was analyzed. As Cost Based and Competitive Based Pricing are not applicable, Value Based Pricing is the right approach to find a Fair Market Value. The concept of the ASPV framework is an expansion of Value Based Pricing for manual value determination in primary markets to automated valuation in secondary markets.

The vast amount of different parts makes an automated part valuation inevitable. This is impeded by missing data and a non-transparent market structure. A generic ASPV framework for reliable value determination of spare parts enables companies to recognize a part's value right away and leads to a more sustainable use of parts. Due to the information gain, it would lead to a more efficient market overall. In the end, an automated value determination prevents bad part utilization decisions and could lead to a more efficient and sustainable use of surplus material. The concept for automated spare part valuation is a promising alternative for value determination and pricing in secondary markets and thus may serve as a foundation for building a generic surplus part trading platform to overcome market transparency issues if the obstacles of validation are overcome.

An early-stage blueprint of a framework for automated spare part valuation which serves as a guide for segmenting pricing data and aligning software layers from data aggregation to transformation, analytics and evaluation has been provided (research question 1). Future research has to deal with refining the alignment of the methods for segmentation of pricing data (research question 2) by benchmarking the approach against competitive concepts such as mean estimation, a k-nearest neighbors clustering, Bayesian Personalized Ranking or approaches without segmentation. Also the methods for predicting a FMV (research question 3) must be evaluated with real world data. Only experiments with real world data are able to give us an insight on whether a solution based on machine learning is even applicable to find a FMV.

References

1. Naisbitt, J. and J. Cracknell, *Megatrends: Ten new directions transforming our lives*. 1984, Warner Books New York.
2. Neely, A. and Y. Jarrar, *Extracting value from data–the performance planning value chain.* Business Process Management Journal, 2004. **10**(5): p. 506-509.
3. Eaves, A.H. and B.G. Kingsman, *Forecasting for the ordering and stock-holding of spare parts.* Journal of the Operational Research Society, 2004. **55**(4): p. 431-437.
4. Wickboldt, C. and N. Kliewer. *A framework for value-based pricing in aircraft surplus markets*. in *MKWI Conference 2016*. 2016.
5. Hinterhuber, A., *Towards value-based pricing—An integrative framework for decision making.* Industrial Marketing Management, 2004. **33**(8): p. 765-778.
6. Anderson, J.C., J.B. Thomson, and F. Wynstra, *Combining value and price to make purchase decisions in business markets.* International Journal of Research in Marketing, 2000. **17**(4): p. 307-329.
7. Anderson, J.C. and J.A. Narus, *Business marketing: understand what customers value.* Harvard business review, 1998. **76**: p. 53-67.
8. Hinterhuber, A., *Customer value-based pricing strategies: why companies resist.* Journal of Business Strategy, 2008. **29**(4): p. 41-50.
9. Kelly, D.B., *Forecasting aircraft values: An appraiser's perspective.* Air-finance annual, 2008.
10. Ackert, S., *Basics of Aircraft Market Analysis.* Aircraft Monitor, 2012.
11. Mercer, Z.C. and T.S. Brown, *Fair Market Value vs. The Real World.* Business Valuation Review, 1999. **18**(1): p. 16-25.
12. Anderson, J.C., D.C. Jain, and P.K. Chintagunta, *Customer value assessment in business markets: A state-of-practice study.* Journal of Business-to-Business Marketing, 1992. **1**: p. 3-29.
13. Blumberg, D.F., *Introduction to management of reverse logistics and closed loop supply chain processes.* 2004.

14. Voß, S. and S. Lessmann, *Resale Price Prediction in the Used Car Market.* Tristan Symposium VIII, 2013.
15. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics, 2009.
16. Grosfeld-Nir, A., B. Ronen, and N. Kozlovsky, *The Pareto managerial principle: when does it apply?* International Journal of Production Research, 2007. **45**(10): p. 2317-2325.
17. Ronen, B., I. Spiegler, and N. Kozlovsky, *The use, misuse and abuse of the Pareto principle.* Working paper, 2005.
18. Blythe, J., *Essentials of marketing*. 2008: Pearson Education.
19. Östlin, J., E. Sundin, and M. Björkman, *Product life-cycle implications for remanufacturing strategies.* Journal of cleaner production, 2009. **17**(11): p. 999-1009.
20. Kimball, R. and M. Ross, *The data warehouse toolkit: the complete guide to dimensional modeling*. 2011: John Wiley & Sons.
21. Allison, P.D., *Missing data. Series: Quantitative applications in the social sciences.* Thousand Oaks. CA, 2002.

Understanding Anomalies: Visualizing Sensor Data for Condition Monitoring of Manufacturing Machines

Daniel Olivotti[1], and Dennis Eilers[1]

[1] Leibniz Universität Hannover, Institute for Information Systems, Hannover, Germany
{olivotti,eilers}@iwi.uni-hannover.de

Abstract. In a more and more service oriented manufacturing industry new data driven challenges like predictive maintenance arise. For example, machine learning models can use sensor data to predict anomalies during machine operation. Such models usually base on experiences from past data to train these algorithms. However, since components are often used in different machines with different and partly unique or new domains, experiences about mutual interferences are missing. In this study we try to address this issue by introducing a visualization technique for an intuitive anomaly detection which allows domain experts and engineers to monitor the condition of a machine over time. The heat map based visualization highlights unusual operation measurements dependent on different sensor data combinations. With domain and engineering knowledge, the insights can be used to identify case based reasons for a changing behavior. The application is tested with a demonstration machine.

Keywords: anomaly detection, condition monitoring, sensor data, manufacturing machines, visualization

1 Introduction and Motivation

In manufacturing industries services play an important role. Machine builders as well as component suppliers are offering physical products along with services, so called product-service systems, more and more [1], [2], [3]. Services lead to further revenue channels and long-term customer satisfaction [1], [4].

Examples for these services are condition monitoring and predictive maintenance. Potential downtimes of machines in manufacturing industries usually result in high breakdown costs. To counteract this, maintenance helps to keep machine availability at a high level. With the help of predictive maintenance, anomalies can be detected early, and the necessary measures can be taken. This results in high cost savings as well as high machine output. The combination of actual sensor values with prediction models in order to forecast a degradation of components is a challenging task [5]. Machine builders integrate components from different component suppliers in their machines. Individual requirements and preferences of machine owners are considered and increases the variety of machines. The interplay between components in machines is essential when anomaly detection is performed, and breakdown possibilities of machines are calculated. Component suppliers have detailed knowledge of their own

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

products but combining this company specific knowledge and different sensor data for a whole machine can lead to a better understanding of observed phenomena.

In this study, we introduce a heat map based visualization technique to better identify the reasons for anomalies, depending on different sensor data combinations. The idea is to use machine learning models trained with time series operation data (e.g. motor speed, or torque) in order to learn how the normal operation of a unique and individual machine looks like. In a subsequent step the models are applied to constantly predict future operations states such that actual and predicted values can be compared. The resulting residuals are then visualized on a heat map color scale depending on different sensor values of different machine components. The visualization allows to intuitively identify clusters of unusual operation states in the context of sensor data like the temperature or the vibration of certain components. The technique also allows continuous updates to assess mutual interferences between components over time. By using this visual analytics approach in combination with machine learning models, application domain experts and engineers can discuss possible reasons or improvements regarding unusual observations more easily. We illustrate the functionality by using a unique demonstration machine which is built to test the interplay of different components during operation.

The remainder of this paper is structured as follows. The research background with focus on predictive maintenance and condition-based monitoring as well as visualizations for data analytics applications is stated out in Section 2. The constructed demonstration machine along with a description of relevant sensor data are presented in Section 3. The anomaly detection approach based on Artificial Neural Networks (ANN) is presented in Section 4. In Section 5 the heat map based visualization applied to the sensor data of the demonstration machine is focused. The results and limitations are discussed in Section 6. Conclusions and an outlook are presented in Section 7.

2 Research Background

Maintenance of machines is a widely discussed topic in the literature. Predictive maintenance is not based on fixed time schedules but aims to predict failures before they occur which reduces machine downtimes and costs of maintenance. Condition monitoring and predictive maintenance received increasing attention recently due to the availability of massive amounts of machine data. The goal is to identify possible anomalies in the operation of machines in an early stage of the impending problem to prevent major failures and corresponding downtime cost or more expensive repair measures [6]. [7] define predictive maintenance as a condition-based preventive maintenance approach. The actual condition derives from sensors or models or combination of both. [8] divides prognostic models in physical models, knowledgebase models and data-driven models. There exists for example applications of assessing the condition of wind turbines by analyzing sensor data [6], [9]. In the literature, several methodologies like statistical analysis or artificial intelligence are discussed. Various models and algorithms are used for condition-based maintenance [8], [10], [11], [12]. Also for machinery diagnostics and prognostics several research projects are already conducted [10], [13]. Predictive maintenance is seen as IT-based service in the context of the Internet of Things [14]. A service platform is required to offer such services in value networks [15]. Visualization techniques which goes further than simple line or bar charts or the display of key performance indicators are rarely discussed in literature.

Today, more and more individualized machines offer a great opportunity for the production of smaller lot sizes and thus also more individual products but also increases the complexity of tasks like predictive maintenance since no reference values and experiences with certain combinations of machine components are available. Therefore, individual application domain and engineering knowledge is required to be still able to identify possible anomalies or unusual behavior. Visualization is a suitable vehicle to trigger human inference from massive amounts of data [16]. Using visualization to support data analyst can be summarized under the term visual analytics which is first defined by [17]. For comprehensive surveys we refer to [18], [19]. Studies like [20], [21] even show that with help of visualizations human can outperform pure machine learning models in certain analytics task. This shows that human reasoning and domain knowledge is especially important in a sparse data space and with noise and uncertainty in the data. In addition, visualization can improve the communications which is essential for the overall team performance [22]. This is necessary to discuss possible reasons and solutions for identified problems with machines. Promoting a visual literacy can establish a common language on the same level of complexity [23] between different expert groups with different proficiencies which also helps to facilitate efficient problem identification. Therefore, visualization in general can help to tackle the challenges which arise in today’s service based manufacturing.

Since nowadays the condition of a machine can be measured by a huge amount of different sensor data, for example to capture the temperature or the vibration, a visualization technique must be able to handle large and diverse datasets. Heat maps are often applied to large amounts of data to provide an intuitive and easy understandable overview in a two-dimensional data space. They are mainly used to show particularly distinctive patterns in certain regions and are therefore suitable for detecting unusual patterns during the operation of a machine. Researchers have shown the applicability of heat maps for several different problem domains. For example [24] use heat maps to illustrate results of a risk management model, [25] visualizes the density of forecasting ensembles over certain time steps and [26] use 2D and 3D heat maps to visualizes performance measures of regression models.

Standard approaches for processing sensor data in condition monitoring and predictive maintenance applications use simple threshold values as problem indicators or more advanced machine learning algorithms to classify machine-related anomalies [27]. Both require experience and, in particular, previous data about a certain machine type. This is often not the case with custom designs, so there is no experience about the interaction of the components. The visualization approach in this study should addresses the question of how to better understand and assess normal or unusual behavior of machines where no or only little experience and past data are available. Therefore, we use the afore mentioned ideas to propose a comprehensible and intuitive heat map visualization which supports application domain experts and engineers in assessing machine conditions to identify the reasons for a problem based on large amount of sensor data.

3 Demonstration Machine and Sensor Data

A demonstration machine within a project with an automation and engineering company is used to provide data in a flexible way. The demonstration machine represents a practical example of a manufacturing machine. The machine is equipped with hoist and conveyor applications to move goods. These exemplary applications are typically for the involved company. Each of these applications has different technical requirements and specifications based on positioning, movement and dynamic. The interplay of sensor data from different applications and motion axes is not known in advance and therefore the anomaly detection, presented in the following sections, is used. Especially when goods with different sizes, weights and position need to be processed, the physical strain cannot be estimated simple by the mean or intervals. In total seven axes provide a circular material flow of goods in the demonstration machine. Two hoist axes are equipped each with a conveyor and are on the left and right side of the machine. In the center of the machine three conveyors are arranged one above each other. All axes are powered by electrically-geared motors. These geared motors are powered by inverters.

A programmable logic controller (PLC) send commands to the inverters and from there they are passed on to the geared motors. Actual status information and sensor data is transferred in reverse order to the PLC. In the PLC a logic controls the process based on defined states and actions. The inverters provide data within cycle times of less than 10 milliseconds (ms). These is identified as a stable cycle time for recording data as the necessary fieldbus communication is not interfered. From the PLC the data is provided to an industrial PC where the data is stored. There exist several protocols to provide data over the network e.g. MQTT and OPC Unified Architecture. Voltage, current, torque and rotational speed of the engines are recorded and used for anomaly detection. These are also the main process variables for the movement control of the seven axes. Other data from inverters and geared-motors are not considered at the moment. The built-in components are able to retrieve this process parameters without the usage of additional sensors. A visualization terminal displays relevant parameters to the machine operator.

4 Time Series Anomaly Detection with Artificial Neural Networks

In this section we describe our analysis approach to detect anomalies of the demonstration machine. In order to evaluate the current operation status, we investigate a time series of torque measured in newton metre (Nm) which reflects the different positions of a weight on a conveyor belt. Figure 1 shows a plot of the torque values in

Nm for six cycles. This time series captures the occurring movement patterns during the ongoing operation and thus serves as a measure for the condition of the machine.

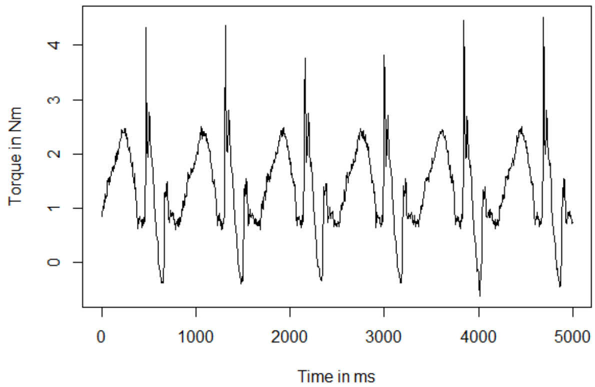


Figure 1. Time series of torque in Nm

To identify an unusual behavior during the operation we use the following anomaly detection approach: Firstly, we sample 60 cycles of sensor data which defines the expected behavior of the demonstration machine in normal conditions. Secondly, ANNs are used to learn and afterwards predict the sensor values for the next time step. Thirdly, the measured value at the next time step is compared to the predicted value and the resulting residuals are evaluated based on a confidence level of how likely an occurrence of this magnitude actually is under the assumption of normal operation conditions. In our example we use simple feed-forward ANNs trained with lags of the 20 past steps of the time series. 30 ANNs with different random weight initializations are trained to build an ensemble based on averaged predictions. To prevent the problem of overfitting we use an early stopping approach. We therefore split the available training data into 80% actual training data and 20% validation data. The 80% portion is used for the iterative learning process, while the 20% portion simulates the out-of-sample results. The actual out-of-sample performance is then observed on a completely independent holdout test dataset. The training process is terminated as soon as the error on the validation data no longer decreases. The ANNs consist of two hidden layers with 50 and 25 hidden neurons using the hyperbolic tangent as activation function. ANNs in general serve as a machine learning method for classification and regression problems and are applied to a variety of forecasting problems [28]. They are able to learn complex, non-linear patterns from data and are reasonably robust against noisy data [29]. To implement the described approach, we use the H2O framework for deep learning (version 3.10.5.3) and perform all necessary data pre-processing and result analysis in R (version 3.4.1) and the H2O R interface. Since ANNs and their learning process is not the focus of this work we refer to [30]. But there is much leeway to improving the models further for example by an extensive hyper parameter optimization (number of hidden layers and neurons), different ensemble methods and more sophisticated regularization techniques like dropout. In this study however, we explicitly focus on a simple approach which performs reasonably well on our data set to demonstrate the applicability of the approach and the subsequent anomaly visualization. Figure 2 compares the actual measurement with the predicted values on an out-of-sample test set, while Figure 3 shows the residual distribution (root-mean-square error: 0.05156, mean absolute error: 0.03864).

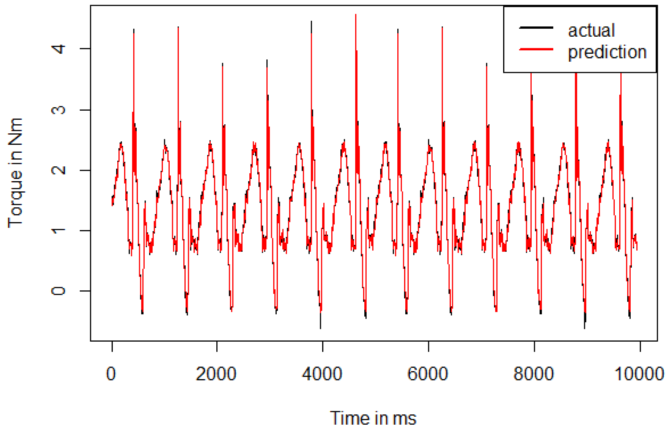


Figure 2. Actual time series vs. prediction

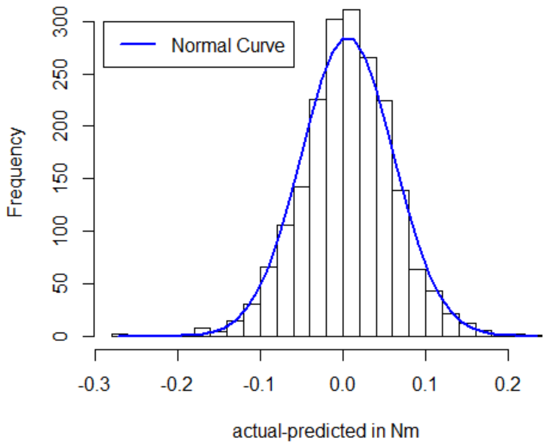


Figure 3. Residual distribution and normal distribution fit

The presented procedure is a standard approach in time series anomaly detection [31] and serves as an early warning system to trigger a subsequent investigation of the reasons for the deviations by means of the analysis method presented in the following section.

5 Heat Map based Visualization

In this section we use the residuals of the previously described time series forecast approach as an anomaly indicator during the operation of the demonstration machine to find the reasons for possible problems. In real world industry scenarios, the interplay of different components in different areas of application of machines can lead to unexpected wear or malfunctions, which should be detected as early as possible, corrected before major failures occur and avoided in the future. This is a challenging task especially with individually-engineered machines due to the lack of experience and missing data. Therefore, the goal is to better understand possible reasons for an unusual behavior of machines based on engineering and application domain knowledge. For this purpose, we propose to use a heat map based residual visualization technique, first introduced by [32] which was previously applied as an evaluation method for machine learning regression models. The idea is to visualize model errors in the context of two different possible features (independent variables) which may have an influence on the distribution of the errors and therefore contain valuable information for explaining or mitigating the problem. To illustrate the functionality, we first artificially simulate two sensor measure time series from two components C-A and C-B of the demonstration machine in addition to the already available torque time series. The first sensor measures the vibration of C-A and the second sensor the vibration of C-B over time. Figure 4 shows the time series during an out of sample period which was not part of the ANN training process.

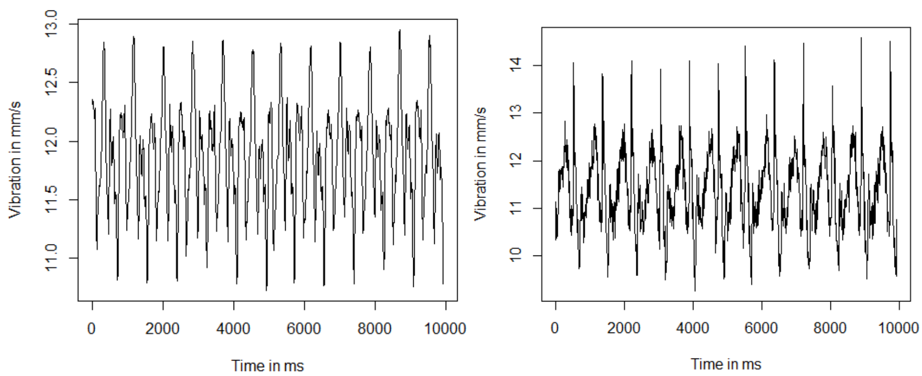


Figure 4. Time series of simulated vibration data for two different components

The idea is to intuitively link the residuals of the time series forecast with measured data of any sensor from any component of the machine to highlight possible dependencies and reasons for observed unusual behavior. In this example we use the two simulated sensors to construct a two-dimensional data space which results in a simple scatter plot as shown in the left plot of Figure 5. We incorporate the density information of the distribution as suggested by [33] to ensure a meaningful representation of the data even for large sample sizes. Darker regions represent a higher density. To incorporate the information about possible anomalies, each data point is now assigned its corresponding residual from the forecasting application. The values of the residuals should be the basis for the color scale of a heat map which incorporates the forecasting error as a further dimension into the graphic. It is assumed that each residual represents the local model results depending on the values of the sensor data. To show the dependencies between two different sensor data sets and the residuals by means of a smoothed color scale in a heat map approach, a regular grid with the respective color codes needs to be calculated. Since the data points are usually not on a regular grid, a method for weighting the influence of each residual on each point on the regular grind in the data space is necessary. Therefore, we use a two-dimensional Gaussian kernel which ensures a strong influence of a residual near the position of the corresponding data point and a decreasing influence with increasing distance. A detailed mathematical description of the procedure can be found in [32]. The results of this simulation can be seen in the right plot of Figure 5.

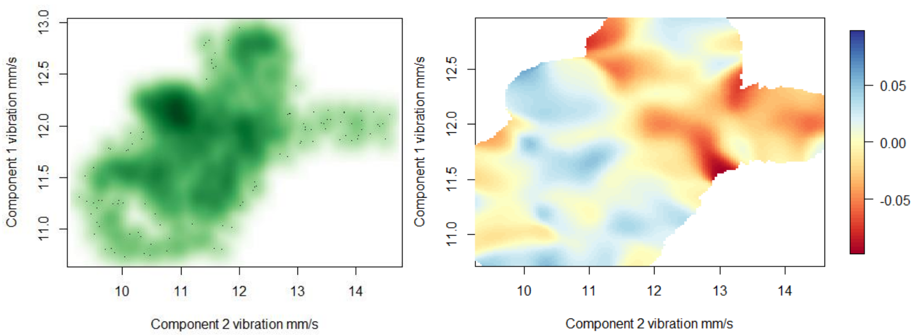


Figure 5. Scatter plot (left) and heat map (right) of residuals depending on sensor data

The color represents the local model errors when predicting the measured Nm of the machine for the next time step which therefore highlights possible unusual behavior of the machine depending on the vibration of component C-A and C-B. In this example, we chose a diverging color palette with 64 divisions. Blue regions represent high positive deviations from the prediction model (actual value in this state is higher than expected by the model) and red regions represent high negative deviations from the prediction model (actual value in this state is lower than expected by the model). In this example, one can observe an overestimation of torque in regions of the data space with higher vibration of both sensors. Underestimations can be observed in conjunction with less vibrations. If the sensor data and therefore the respective component of the machine would have no influence on the torque the expected heat map would have no visible patterns which represents a normal distribution of the residuals. Here, however, various reasons can be responsible for the observed patterns. For example, wear on the motor, incorrectly placed loads or frictions of various components should be taken into account. In applications with many more and different kinds of available sensor data, the visualization can be extended by using a scatter plot matrix to get a quick and comprehensible overview of sensor values/pairwise sensor combinations and the corresponding possible anomalies during the machine operation. The presented technique also allows the user to monitor unusual behavior over time by comparing the heat maps/scatter plot matrix from different time periods which can reveal changes in the condition of special components in the context of the overall machine behavior. Therefore, the approach links unusual machine behavior with sensor data in an intuitive graphical view which can be better interpreted by human experts.

6 Discussion and Limitations

In this first application example, we have shown how heat map based visualizations can be used to monitor the condition of individual machines in the context of measured sensor data from different machine components. The idea follows a two-step approach: Firstly, a machine learning model is used to identify deviations between an expected operation state and the actual observed state of a machine. The resulting residuals of the model are used to construct the heat map which visualizes these deviations on a defined color scale, depending on two different sensor datasets which construct the two-dimensional data space. This enables the user to identify possible clusters of unusual machine behavior in specific regions of sensor measurements. The connection between the anomalies (color scale) and sensor data (axes) can help to identify possible reasons and solutions for identified problems. Such a combination of a machine learning and visual analytics approach can be beneficial especially in situation where only little experience and past data about the machines are available. In addition, the visualization is easy to interpret which can facilitate discussion about the results between people with different background like engineers, data scientist or application domain experts. The technique also enables a monitoring over time by comparing the visualizations from different periods. This allows an assessment about structural changes in the condition of certain components or the combinations of certain components which can be used to perform predictive maintenance measures.

However, both steps of the proposed approach face some difficulties and drawbacks. First the machine learning model currently uses one simple time series of torque for the training process which represents the operation status of the machine. It is assumed that the data for the training represent a normal operation of the machine. The anomaly detection is therefore only able to identify deviations from an assumed ideal situation but there is no possibility to optimize the initial operation behavior. A further limitation is the method to learn the normal operation behavior. Currently a simple feed-forward ANN is used, with some basic optimization and regularization procedures. In further research this will be benchmarked or replaced by more sophisticated algorithms like recurrent neural networks, especially Long Short-Term Memory (LSTM), which are more suitable for time series prediction. The modeling process in general needs therefore further optimization and benchmarks to find the best solution for learning normal operation behaviors of machines.

Regarding the visualization approach, there also exist several limitations and drawbacks. Visualizations in general always require human judgements which includes errors and wrong conclusions. Heat maps in particular are restricted to two dimensions and an additional color scale. The increasing number of sensors, however, make the visualization of all possible sensor combinations more difficult. A scatter plot matrix can help to mitigate the problem. Another problem of heat maps is that they are hardly comparable to each other since the color scale is relative to the measured values and residuals. Quantitative comparisons between different points in time for example can lead to wrong conclusions. Currently also no information about the density is available in the representation, so the heat maps must be shown in conjunction with scatter plots which contain density information.

Another important part of this study is the demonstration machine which provides sensor data from key process parameters. This is advantageous because this data is already used for the control of the motion axes and no additional sensors needed. Cycle times of less than 10 ms are currently used to process sensor data from the PLC to the industrial PC. A tradeoff between low cycle times and high field bus occupancy rate need to be considered. The cycle time must be defined for each parameter to provide the required data quality. In general, additional sensors provides more insights in the condition of machines. The costs of additional sensors must be considered. Common practice is to equip key components for the functions of machines with appropriate sensors. Frequency and volume of sensor data can lead to high amount of data. Policies how long which data is store need to be defined. Interdependencies of components in machines are often complex to model. In some cases, they can solely be found out experimentally. The presented heat map based visualization techniques enables first insights on dependencies of different components. Domain experts or engineers can analysis the affected components in detail. This knowledge can be stored in a knowledge management system to get insights for future applications in advance.

Overall it can be argued that the presented approach can help to trigger human creativity and judgement despite the mentioned difficulties and drawbacks. Especially a visualization approach can only be a support for human analysts who must draw their conclusions based on the provided systems. But to measure the actual benefit of the proposed technique is a challenging task, which is why the evaluation is currently a major limitation of our study.

7 Conclusions and Outlook

Condition monitoring and predictive maintenance aim to reduce maintenance efforts while keeping machine availability high. This is achieved by the use of sensor data from various components and machines. To detect anomalies, the interplay of components in machines is essential. In this article, a heat map based visualization technique is applied to sensor data from manufacturing machines. By combining information of different sensors in a visualization, domain experts or engineers are facilitated to find possible reasons for anomalies during machine operation.

In further research, more components from different component suppliers should be included in the demonstration machine. Long-time test runs enable an evaluation of the presented application of heat map based visualizations in detail. In this article we limited our analysis on different components in machines. An analysis of machines in interchained production lines is also planned to predict machine park availability.

We will also test different machine learning models to better understand the current state of a machine. Especially time series models like LSTM networks might be suitable for this purpose. For live updates of the heat map visualization, the performance of the calculations need to be improved. The current approach is implemented in the programming language R. To increase the performance, other lower level languages like C or languages for high performance computing like Julia might be better choices. Different cloud solutions are also taken in consideration since service-based solutions should be provided.

References

1. Oliva, R., Kallenberg, R.: Managing the transition from products to services. Int. J. Serv. Ind. Manag. 14 (2), 160–172 (2003)
2. Schrödl, H., Bensch, S.: E-Procurement of Cloud-based Information Systems – a Product-Service System Approach. In: Proceedings of the Thirty Fourth International Conference on Information Systems (ICIS), Milan, Italy, December 15–18. (2013)
3. Neff, A.A.: Developing a Maturity Model for Product-Service Systems in Manufacturing Enterprises. In: Proceedings of the Pacific Asia Conference on Information Systems (PACIS), Jeju Island, Korea, June 18–22. (2013)
4. Barrett, M., Davidson, E., Prabhu, J., Vargo, S.L.: Service Innovation in the Digital Age: Key Contributions and Future Directions. MISQ 39 (1), 135–154 (2015)
5. Kaiser, K.A., Gebraeel, N.Z.: Predictive Maintenance Management Using Sensor-Based Degradation Models. IEEE Trans. Syst., Man, and Cybern. - Part A: Syst. and Hum. 39 (4), 840–849 (2009)
6. Garcia, M.C., Sanz-Bobi, M.A., Del Pico, J.: SIMAP: Intelligent System for Predictive Maintenance Application to the health condition monitoring of a windturbine gearbox. Comput. Ind. 57, 552–568 (2006)
7. Mobley, R. D.: An Introduction to Predictive Maintenance, Second Edition, Butterworth-Heinemann, Oxford, UK (2002)
8. Peng, Y., Dong, M., Zuo, M.J.: Current status of machine prognostics in condition-based maintenance: a review. Int. J. Adv. Manuf. Technol. 50, 297–313 (2010)
9. Tchakoua, P., Wamkeue, R., Ouhrouche, M., Slaoui-Hasnaoui, F., Tameghe, T.A., Ekemb, G.: Wind Turbine Condition Monitoring: State-of-the-Art Review, New Trends, and Future Challenges. Energies 7, 2595–2630 (2014)
10. Jardine, A.K.S., Lin, D., Banjevic, D.: A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech. Syst. Signal Process. 20, 1483-1510 (2006)
11. Ahmad, R., Kamaruddin, S.: An overview of time-based and condition-based maintenance in industrial application. Comput. Ind. Eng. 63, 135–149 (2012)
12. Tsang, A.H.C.: Condition-based maintenance: tools and decision making. J. Qual. Maint. Eng. 1 (3), 3–17 (1995)
13. Kothamasu, R., Huang, S.H., VerDuin, W.H.: System health monitoring and prognostics - a review of current paradigms and practices. Int. J. Adv. Manuf. Technol. 28, 1012–1024 (2006)
14. Wortmann, F., Flüchter, K.: Internet of Things - Technology and Value Added. Bus. Inf. Syst. Eng. 57 (3), 221–224 (2015)

15. Gerloff, C., Cleophas, C.: Excavating the Treasure of IoT Data: An Architecture to Empower Rapid Data Analytics for Predictive Maintenance of Connected Vehicles. In: Proceedings of the Thirty Eighth International Conference on Information Systems (ICIS), Seoul, South Korea, December 10-13. (2017)
16. Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G.: Visual Analytics: Definition, Process, and Challenges. In: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (eds.) Information Visualization, Lecture Notes in Computer Science, pp. 154–175. Springer, Berlin, Heidelberg (2008)
17. Thomas, J.J., and Cook, K.A.: A Visual Analytics Agenda. IEEE Comput. Graph. Appl. 26 (1), 10–13 (2006)
18. Sun, G.-D., Wu, Y.-C., Liang, R.-H., Liu, S.-X.: A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges. J. Comput. Sci. Technol. 28 (5), 852–867 (2013)
19. Endert, A., Hossain, M.S., Ramakrishnan, N., North, C., Fiaux, P., Andrews, C.: The Human Is the Loop: New Directions for Visual Analytics. J. Intell. Inf. Syst. 43 (3), 411–435 (2014)
20. Tam, G.K.L., Fang, H., Aubrey, A.J., Grant, P.W., Rosin, P.L., Marshall, D., Chen, M.: Visualization of Time-Series Data in Parameter Space for Understanding Facial Dynamics. Comput. Graph. Forum 30 (3), 901–910 (2011)
21. Tam, G.K.L., Kothari, V., Chen, M.: An Analysis of Machine- and Human-Analytics in Classification. IEEE Trans. Vis. Comput. Graph. 23 (1), 71–80 (2017)
22. Bresciani, S., Eppler, M.J.: The Benefits of Synchronous Collaborative Information Visualization: Evidence from an Experimental Evaluation. IEEE Trans. Vis. and Comput. Graph. 15 (6), 1073–1080 (2009)
23. Fransecky, R.B., Debes, J.L. Visual Literacy: A Way to Learn–A Way to Teach: Association for Educational Communications and Technology, Washington D.C. (1972)
24. Buehler, K.S., Pritsch, G.: Running with Risk. McKinsey Q. 4, 40–49 (2003)
25. Köpp, C., von Mettenheim, H.-J., Breitner, M.H.: Decision Analytics with Heatmap Visualization for Multi-Step Ensemble Data. Bus. Inf. Syst. Eng. 6 (3), 131–140 (2014)
26. Klemm, P., Lawonn, K., Glaßer, S., Niemann, U., Hegenscheid, K., Völzke, H., Preim, B.: 3D Regression Heat Map Analysis of Population Study Data. IEEE Trans. Vis. Comput. Graph. 22 (1), 81–90 (2016)
27. Lee, J., Kao, H.-A., Yang, S.: Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. Procedia CIRP 16 (Supplement C), 3–8 (2014)
28. Adya, M., Collopy, F.: How Effective Are Neural Networks at Forecasting and Prediction? A Review and Evaluation. J. Forecast. 17 (5–6), 481–495 (1998)
29. Schocken, S., Ariav, G.: Neural Networks for Decision Support: Problems and Opportunities. Decis. Support Syst.11 (5), 393–414 (1994)
30. Bishop, C.M.: Neural Networks for Pattern Recognition, Oxford University Press. Oxford (1995)
31. Malhotra, P., Vig, L., Shroff, G., Agarwal, P.: Long Short Term Memory Networks for Anomaly Detection in Time Series. In: 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, April 22-24 April. (2015)
32. Eilers, D., Köpp, C., Gleue, C., Breitner, M.: It's Not a Bug, It's a Feature: How Visual Model Evaluation Can Help to Incorporate Human Domain Knowledge in Data Science. In: Proceedings of the Thirty Eighth International Conference on Information Systems (ICIS), Seoul, South Korea, December 10–13. (2017)
33. Carr, D.B., Littlefield, R.J., Nicholson, W.L., Littlefield, J.S.: Scatterplot Matrix Techniques for Large N. J. Am. Stat. Assoc. 82 (398), 424–436 (1987)

Datenmanagement im Zeitalter der Digitalisierung

Teilkonferenzleitung

Christine Legner
Boris Otto
Dirk Stelzer

Barriers to Adopting Data Pools for Product Information Sharing – A Literature Review

Thomas Schäffer[1] and Dirk Stelzer[2]

[1] Hochschule Heilbronn, Studiengang Wirtschaftsinformatik, Heilbronn, Germany
Thomas.Schaeffer@hs-heilbronn.de

[2] Technische Universität Ilmenau, Fachgebiet Informations- und Wissensmanagement, Ilmenau, Germany
Dirk.Stelzer@tu-ilmenau.de

Abstract. Data pools, i.e., master data repositories that support companies in exchanging information on trade items with business partners in a standard format, have been rated as ideal tools for product information sharing. However, it is widely accepted that adoption of data pools has fallen short of expectations. There is an obvious mismatch between the benefits and advantages associated with the use of data pools and their actual diffusion. This paper presents a theory-based literature review into barriers to adopting data pools. The findings may support providers of data pools to improve their services and adopters of data pools to lay the foundations for an efficient use of these electronic intermediaries for product information sharing.

Keywords: Data Pools, Data Quality, Global Data Synchronization, Master Data, Product Information Sharing, Technology Adoption and Diffusion

1 Introduction

Product information sharing denotes the inter-organizational synchronization of master data relating to products [1]. The term master data refers to critical business objects of an organization that rarely undergo changes [2]. Master data synchronization means achieving consistent information values for items or products within and between organizations [3]. Product information sharing is a problematic, error-prone, labor-intensive, and costly process in most industries [4]. Companies may adopt a variety of standards, technologies, tools, and intermediaries to support product information sharing [1, 5, 6]. One of the most powerful tools for enhancing inter-organizational synchronization of master data is a data pool. A data pool is a repository that supports trading partners in obtaining, maintaining, and exchanging information on trade items in a standard format through electronic means [4]. Data pools combine standards and technologies to exchange product information from manufacturers' to retailers' internal databases through the use of electronic catalogs. These catalogs contain data describing articles and they also coordinate data

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

exchanges [5]. Instead of multiple bilateral exchanges a data pool acts like a clearing center. Manufacturers transfer data to the pool only once and retailers may collect the data from the pool [7]. Data pools assume the role of B2B intermediaries in data synchronization, but they do not focus on matching buyers and sellers [8]. The Global Data Synchronization Network (GDSN) [9] is a global network of specialized data pools. However, companies may also use other data pools that are not part of GDSN.

Gopal and McMillian characterize GDSN as one building block of successful collaboration in the supply chain [3]. De Cobière and Rowe [5] have analyzed various forms of interconnections to achieve data synchronization between companies. They come to the conclusion that data synchronization as supported by GDSN is an "ideal" form of product information sharing. However, it is widely accepted that the potential of data pools has not been fully realized, and that the adoption of data pools has fallen short of expectations [1, 4, 6, 7].

There is an obvious mismatch between the benefits and advantages associated with the use of data pools and their actual diffusion. Accordingly, several scholars ask for more intensive research into the issues of data pool adoption [1, 4, 7].

Some scholars have explored adoption and usage of data pools [1, 4, 5, 7, 10, 11]. They have identified issues and limitations of data pool adoption and usage in manufacturing and retailing companies. However, to the best of our knowledge, no author has yet attempted to provide a comprehensive overview of issues that companies face when adopting data pools for product information sharing.

The objective of our research is to identify reasons for the low adoption rate from previous research and to draw conclusions for providers and users of data pools and for future research. We aim to answer two research questions:

- What are the reasons for the low adoption rate of data pools?
- What conclusions can be drawn from identifying barriers to adopting data pools?

This paper is organized as follows. First, we give a short overview of our methodology. Next, we describe findings of our literature review. We then discuss limitations of our research. We conclude with a summary of our results and suggestions for providers and users of data pools, and for future research.

2 Method

2.1 Literature Review

We conducted a literature review to identify prior research into adoption barriers and issues when using data pools for product information sharing. We followed the guidelines provided by Webster and Watson [12] and vom Brocke et al. [13] to identify relevant publications. As a first step, we examined IS journals and IS conference proceedings using the AIS Electronic Library, ScienceDirect, Google Scholar, and SpringerLink. We conducted electronic searches in titles on the following search term: ["data pool" OR "Datenpool" OR ("electronic" AND ("catalog" OR "catalogue"))]. Furthermore we did a full text search on the following search term: ["master data pool" OR "Stammdatenpool"]. These searches identified a total of 793 publications. After analyzing each article's abstract, keywords, or the full article when necessary, we excluded 757 articles that were duplicates or did not appear to be concerned with or relevant to our research focus. As a third step, we performed forward and backward searches in relevant articles to identify further sources. A total of 45 publications were read in full and coded. We excluded all papers that did not seem to be peer reviewed publications or that only stated the keywords mentioned in the search term without elaborating on these concepts. We also excluded all papers that did not mention issues, limitations, and weaknesses of data pools, or adoption barriers, respectively. Out of the 45 coded articles, 23 include passages of interest. A complete list of all papers and a documentation of the coding procedure is available from the authors upon request. Table 1 in section 3 gives an overview of the 23 publications that elaborate on barriers to adopting data pools. In the remainder of this paper we focus on these papers.

2.2 Selection of a Theoretical Basis

In the course of our literature review we have identified several candidates that may serve as theoretical or conceptual basis for coding and analyzing the papers. In the following section we briefly discuss the most widely used theories and models.

Coordination theory is the most commonly used theory in research on product information sharing [1, 5, 6]. However, coordination theory is not a comprehensive theoretical basis for assessing barriers to the adoption of technologies or tools.

The transaction cost approach may help to assess economic aspects of adopting data pools for product information sharing. However, apart from a few rare exceptions, e.g., [3, 8, 14, 15], prior research has not attempted to quantify transaction costs or benefits related to data pool adoption and usage. Thus, literature does not yet provide enough empirical data for such research.

Madlberger [7] has explored adoption barriers of data pools from the perspective of **diffusion and adoption of standards**. She identified network effects, standardization costs, excess inertia, and path dependencies as potential barriers or facilitators of adopting data pools. However, as data pools are a combination of standards, technology, and data, diffusion and adoption of standards alone falls short of a comprehensive analysis of adoption barriers of data pools [1, 5].

The technology acceptance model is another candidate for a theoretical basis. However, the level of analysis of this model is the individual. Decisions to adopt data pools or to refrain from doing so, however, are made on a group, organizational, or even industry level [7]. Consequently, this model is inadequate for our research.

The diffusion of innovations theory [16] explains how innovations are adopted by social systems, i.e., individuals, groups, organizations, industries, or societies. An innovation is an idea, practice, or object that is perceived as new. Thus, adopting a data pool in a company can be regarded as an innovation [7, 10]. The most important characteristics of innovations that explain the rate of adoption are: relative advantage, compatibility, complexity, trialability, and observability. The diffusion of innovations theory seems to be well suited for our research objective.

Information systems (IS) success models – the most famous models were published by DeLone and McLean [17, 18] – conceptualize system quality, information quality, and service quality as important antecedents to a system's usage, organizational impact, or its net benefits. Data pools can be regarded as inter-organizational IS [4, 11]. Given the fact that quality is a key characteristic in many publications on the adoption of data pools for product information sharing [6, 7, 11, 19], IS success models may also be well suited for our research.

Usually, diffusion of innovations theory and IS success models are used for explaining the adoption of innovations or success of IS. In the present paper, however, we strive for identifying barriers to adoption, i.e., we do not search for enablers but for inhibitors of adoption. Consequently, the characteristics of innovations and the success factors must be interpreted contrariwise.

Diffusion of innovations theory focuses on the following characteristics of innovations that influence their adoption [16]: **Relative advantage**, the degree to which an innovation is perceived as better than the idea it supersedes; **compatibility**, the degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of potential adopters; **complexity**, the degree to which an innovation is perceived as difficult to understand and use, **trialability**, the degree to which an innovation may be experimented with on a limited basis; and **observability**, the degree to which the results of an innovation are visible to others.

From the IS success models we deduced the following antecedents to system usage: **System quality** denotes desirable characteristics of an IS, e.g., functionality, ease of use, or system flexibility and reliability [17]. **Information quality** describes desirable characteristics of the system outputs, e.g., relevance, understandability, accuracy, completeness**,** understandability, timeliness, usability [17]. **Service quality** denotes quality of the service or support that users receive from the IT support personnel. Key service quality characteristics are responsiveness, accuracy, reliability, technical competence, and empathy of the personnel staff [17].

2.3 Coding Scheme

We used these criteria or, more precisely, their opposite, for identifying barriers to adopting data pools in a first iteration of our literature analysis. After having analyzed all 23 papers, we found that none of the two approaches is suitable to categorize each of the barriers mentioned in the literature. Furthermore, some of the criteria do not seem to be relevant for the adoption of data pools as they are not mentioned in the literature. We did not identify a single reference to issues relating to trialability, observability, or issues with service quality of data pools. We therefore deleted these factors from the second iteration of our analysis. Furthermore, we discovered that issues relating to system and information quality are inherent characteristics of data pools. Another characteristic that directly relates to data pools and that is mentioned frequently in the literature is cost, i.e., fees for access, licenses, and usage that users of data pools have to pay. We therefore decided to complement system and information quality by cost of data pools and label these three criteria as characteristics of data pools. They are perceived as barriers to adoption.

Apart from factors directly relating to characteristics of data pools, some authors [3, 10] suggest to distinguish two more perspectives on adopting data pools.

- Inter-organizational factors denote environment characteristics that influence the adoption of data pools. When analyzing the literature for inter-organizational factors, we identified the following categories:
  - As data pools are subject to network effects they require a critical mass of data providers and data recipients to become attractive [1, 7].
  - The decision to adopt a data pool is depending on the perceived commitment of relevant business partners to use the same data pool and on the trust that business partner make available high quality data [4].
  - External IT providers have supported many companies in inter-organizational synchronization of master data before they decided to adopt a data pool. Therefore, external IT providers may act as gatekeepers that may facilitate or inhibit data pool adoption [20].
- Intra-organizational factors denote factors relating to users of data pools, i.e., data suppliers and data recipients. We identified three categories of such factors:
  - Internal capabilities (e.g., knowledge of personnel, processes or process areas, technology and infrastructure, and organization structures) of data providers and data recipients may impede or prevent the adoption of data pools [1, 6].
  - As the adoption of a data pool requires considerable internal adaptations and investment, a high management commitment is indispensable [4].
  - The adoption of a data pool requires considerable efforts for internal adaptation, e.g., setup costs for electronic integration. Perceived high internal costs may become an inhibitor of data pool adoption [1, 3, 6, 21].

Figure 1 gives an overview of the methodology used in our research.

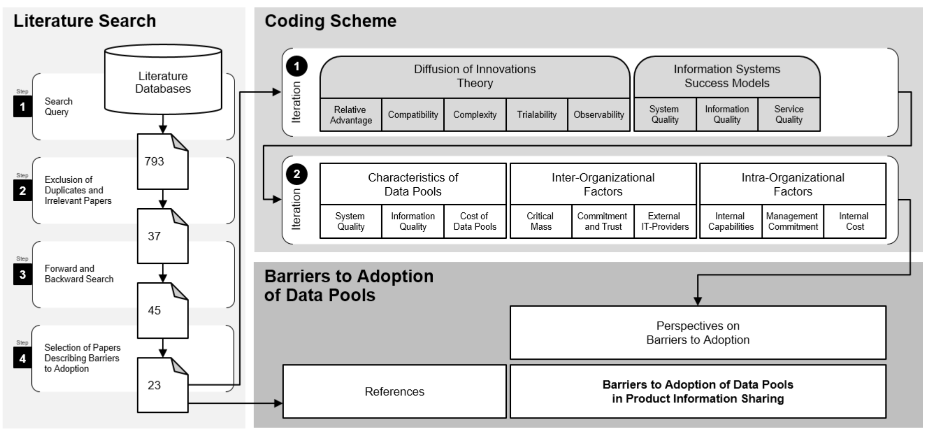


Figure 1. Overview of the Methodology

3 Findings

In the following section, we briefly present barriers to adopting data pools for product information sharing that we identified when analyzing the literature. We use the coding scheme presented in the previous section as a structure for our findings.

3.1 Characteristics of Data Pools

Data pools are combinations of standards and technologies to exchange product information [5]. Consequently, some barriers to adoption relate to standards used by data pools and others concern data pool technologies.

System Quality

Data pools do not always seem to apply adequate mechanisms for ensuring the quality of product master data included in the pool [22]. They trust that data providers supply data of the desired quality. However, this is not always the case. This leads to insufficient data quality in some cases [8, 21].

Both manufacturers and retailers miss feedback loops to improve data quality. They articulate the need for providing and receiving feedback on the quality of data available in the data pool. Currently, data pools do not offer such mechanisms [11].

Several companies use more than one data pool. As each data pool has a unique user interface, users complain about complicated usage of data pools [6].

Process standards supporting usage of data pools do not seem to be sufficient yet. For example, the standards do not include definitions of responsibilities for plausibility checks or liability agreements for incorrect data [23].

Data pools do not necessarily provide options for definition and handling of product subsets, product families, or product bundles [24, 25]. Moreover, GDSN does not provide options for identifying specific physical products. This complicates supply chain management activities, e.g., product recalls or drug tracking [26]. These shortcomings make it difficult to use data pools in day-to-day operations.

Information Quality

Most data pools restrict their offerings to a specific industry or product category. GDSN, for example, specializes on fast-moving consumer goods [8]. Retailers that also procure apparel or automotive parts usually have to obtain master data from more than one data pool. In other words, as data pools do not cover all product data required by many companies, completeness of information is an issue [27, 28].

Several users complain about a limited number of attributes [6]. Data pools do not provide all attributes needed for executing business processes. Prices and discounts, for example, are often negotiated on a bilateral basis. Consequently, these data have to be exchanged directly between business partners [5, 7, 20, 23]. This requires maintaining additional communication structures. Given that most companies must establish and support dyadic communication channels in addition to exchanging data with data pools, many companies question the cost-effectiveness of using data pools [4, 8, 23].

Some papers report on complaints about missing or wrong identification numbers for logistic units, e.g., packaging units or pallets [29]. This is a result of inadequate validation mechanisms of data providers as well as of data pools [22].

Another weakness is closely related to characteristics of standards. GS1, for example, prescribes several mandatory fields, e.g., height, width, and depth of products [27, 30]. Although this information is relevant for most products, this does not apply, for example, to nets of onions. However, suppliers of onions are forced to enter corresponding data [11]. This demotivates data providers and lowers their readiness to provide high quality data. This corresponds to a finding reported by Legner and Schemm [1] who pointed out that data synchronization needs to be enhanced by providing context information. Otherwise data may often be misinterpreted.

Inadequate timeliness is another issue in many data pools [11]. In case of new product launches delays in the provision of data may impede the efficient execution of business processes.

Moreover, data availability is only guaranteed for a limited time in most data pools. Relevant data are not always available throughout the product life cycle [24].

Cost of Data Pools

Prices charged by providers of data pools, i.e., fees for access, licenses, and usage, are perceived as too high by many suppliers and recipients of data [1, 3, 4, 6, 7]. This problem is aggravated by the lack of transparency of data pools' license models [14].

3.2 Inter-Organizational Factors

Critical Mass of Data Pool Users

Like most intermediaries in electronic business, data pools are subject to network effects, excess inertia, and path dependencies [4, 7, 30]. Data pools will not be able to attract enough data recipients unless they have already attracted enough data providers, and vice versa. However, several of the issues and drawbacks of data pools presented in the last section do not make data pools appear particularly attractive. This, in turn, prevents more companies from participating in data pools.

Commitment and Trust

Business relationships built upon a data pool demand a long-term commitment from all relevant business partners. Additionally, product information sharing via data pools cannot be achieved without a strong commitment to provide high quality data [27]. However, many companies are not sure whether all trading partners are sufficiently dedicated to provide data on an adequate quality level and to maintain long term business relations with a data pool [4, 30].

External IT providers

Boukef Charki et al. point to the power of external IT providers that may act as gate keepers for adopting data pools [20]. If these IT providers perceive data pools as threats to their business models they might try to prevent their clients from participating in data pools.

3.3 Intra-Organizational Factors

Internal Capabilities

Insufficient internal capabilities may prevent data suppliers and data recipients from efficiently participating in product information sharing via data pools. Major problem areas are personnel, business processes, technical infrastructure, and organizational structures.

Small organizations regularly lack IT skills, knowledge, and capabilities for inter-organizational data synchronization. In larger organizations IT personnel is often little aware of the product information sharing process in their own organizations [11]. Master data experts are often not directly involved in this process. Thus, there is no central unit to make a well-informed decision about adopting data pools [1, 6, 8, 30].

In many companies, responsibilities for master data management, data quality, and data governance are scattered over several lines of business. Other companies have not defined adequate strategies for master data management and inter-organizational data synchronization [1, 6, 15, 30, 31]. This again, prevents many companies from adequately assessing potential benefits of using data pools.

Using data pools requires adjustments of internal IS. This is not only a lengthy and complex procedure but it also requires technical skills and capabilities. However, staff members responsible for product information sharing often neither have adequate IT skills, nor sufficient support from their IT departments to manage necessary adjustments [4, 6, 15].

Inadequate organizational structures often impede efficient product information sharing. In many companies, sales and procurement staff are responsible for exchanging product data with trading partners [23]. Often, they are not sufficiently motivated to improve product information sharing and they usually do not have necessary technical capabilities to adequately assess the potential benefits of using data pools [1, 6, 23].

Management Commitment

Many senior managers do not seem to be ready to invest in data pools because they do not fully understand the necessity of automated product information sharing and the economic benefits of using data pools [4, 15, 30].

Internal Cost

Organizations that consider adopting a data pool are confronted with significant setup costs for electronic integration and additional cost for internal adaption of personnel capabilities, business processes, and organizational structures [15, 30]. In some cases they are afraid of the additional effort required, the lengthy implementation procedure, and the perceived high costs for installing, updating, and maintaining new systems to compile and exchange data with a data pool [15, 32].

Consequently, several scholars have identified lack of internal readiness as a major inhibitor of adopting data pools for synchronizing master data [1, 3, 4, 15]. Table 1 gives an overview of barriers to adopting data pools as identified in the literature.

Table 1. Overview of barriers to adopting data pools

	Characteristics of Data Pools			*Inter-Organizational Factors*			*Intra-Organizational Factors*		
Reference	System Quality	Information Quality	Cost of Data Pools	Critical Mass	Commitment and Trust	External IT-Providers	Internal Capabilities	Management Commitment	Internal Cost
Becker et al. 2007 [25]	x								
Becker et al. 2008 [31]							x		
Boukef Charki et al. 2011 [20]	x	x				x	x		
Chen & Prater 2013 [15]							x	x	x
Dalmolen et al. 2015 [11]	x	x					x		
de Corbière & Rowe 2011 [32]		x							x
de Corbière & Rowe 2013 [5]		x	x	x					
Gopal & McMillian 2005 [3]	x								x
Hüner et al. 2011 [29]		x							
Karpischek et al. 2012 [28]	x	x							
Karpischek et al. 2014 [27]	x	x					x		
Legner & Schemm 2008 [1]	x						x		x
Madlberger 2011 [7]		x		x			x		x
Nakatani & Chuang 2012 [30]	x	x	x	x	x		x	x	x
Nakatani et al. 2006 [4]	x	x		x	x		x	x	x
Schäffer & Stelzer 2017 [6]	x	x	x	x					
Schemm & Legner 2008 [8]	x	x		x			x		
Schemm 2009 [14]			x				x		x
Schemm et al. 2006 [21]	x						x		x
Schemm et al. 2008 [23]	x		x				x		x
Winkelmann et al. 2008 [24]	x	x							
Zhou et al. 2006 [26]	x								
Zhou et al. 2011 [22]	x	x							

4 Discussion

To the best of our knowledge this is the first review of research into issues that companies face when adopting data pools for product information sharing. Our research has identified a significant number of limitations and drawbacks that can be understood as barriers to adoption. All in all, these barriers may explain why data pools are used so little for product information sharing.

Our review has shown that many companies evaluate system and information quality and the price/performance ratio of data pools as unfavorable. Moreover, many users perceive relative disadvantages of a data pool when they compare it to more traditional forms of data synchronization, e.g., bilateral exchanges of product master data via EDI connections or spreadsheet files [1, 5–7]. Moreover, companies that have not already implemented master data management and internal data synchronization systems perceive data pools as not being consistent with their internal structures, systems, and processes.

Our review has revealed that none of the theories and conceptual approaches presented in section 2.2 fully covers all barriers that were identified in the literature. IS success models support the identification of unfavorable properties of data pools. However, these models are not appropriate when it comes to identifying and analyzing barriers that are no inherent characteristics of data pools. Diffusion of innovations theory puts the focus on users of data pools, i.e., providers and recipients of master data. However, characteristics put forward by this theory are too abstract for a detailed identification of barriers to adopting data pools. Based on the literature we have developed a more specific scheme that helped us to identify and to classify barriers to adopting data pools for product information sharing.

As several of the papers included in our review are more than ten years old, some of the issues with data pools may have been superseded in the meantime. Some of the barriers shown in table 1 are mentioned in a few papers only. It might be helpful to explore whether these properties are relevant for individual companies in specific situations only or whether they keep many companies from adopting data pools. Most of the papers included in our review focus on GDSN data pools. Our review does not distinguish between these data pools and others that are not part of GDSN. More empirical research is needed that explores current issues relating to data pools.

5 Conclusions

From the findings of our review we draw conclusions for providers and users of data pools and for future research.

Implications for Providers of Data Pools

Providers of data pools should strive for improving the price/performance ratio of their services, either by improving system and information quality or by lowering fees for access, licenses, and usage of data pools.

Several users criticize that they have to maintain bilateral communication channels for exchanging master data with business partners in addition to using data pools. If data pools supported private communication channels that companies could use to exchange confidential information on a bilateral basis, e.g., prices or discounts, this would probably be perceived as a helpful functional enhancement.

Implications for Users of Data Pools

Before adopting data pools, firms must implement company-wide master data management and internal data synchronization mechanisms. They should also adapt knowledge and capabilities of staff members, business processes, IT infrastructure, and organization and governance structures. Otherwise, they take the risk of incompatibilities between their internal structures and data pools.

When calculating a business case for data pool adoption, companies should not only criticize the perceived high costs of data pools. They should rather compare it to all benefits and costs of the error-prone, labor-intensive, and costly processes of traditional master data synchronization.

A detailed business case could also help to ensure senior management commitment for adopting a data pool.

Implications for Scholars

It is noticeable that only few of the barriers to adopting data pools can directly be related to the theories and models discussed in section 2.2. It should be considered to put forward a new theoretical basis for conducting research on data pool adoption.

The coding scheme presented in section 2.3 was derived from theories and conceptual approaches and from literature on data pools for product information sharing. However, we assume that more research could lead to more detailed schemes for structuring and categorizing barriers to the adoption of data pools.

We recommend more empirical research into current issues relating to data pools, e.g., in the form of case studies or surveys.

A more detailed analysis into different categories of data pools, e.g., distinguishing between GDSN data pools and others, could help to obtain a more accurate picture of barriers to adoption of specific data pools.

References

1. Legner, C., Schemm, J.: Toward the Inter-organizational Product Information Supply Chain - Evidence from the Retail and Consumer Goods Industries. JAIS 9, 119–150 (2008)
2. Loshin, D.: Master Data Management. Morgan Kaufmann, Amsterdam (2008)
3. Gopal, G., McMillian, E.: Synchronization: A Cure for Bad Data. SCMR 9, 58–62 (2005)
4. Nakatani, K., Chuang, T.-T., Zhou, D.: Data Synchronization Technology. CAIS 17, 962-996 (2006)
5. de Corbière, F., Rowe, F.: From Ideal Data Synchronization to Hybrid Forms of Interconnections: Architectures, Processes, and Data. JAIS 14, 550–584 (2013)
6. Schäffer, T., Stelzer, D.: Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing. In: WI 2017 Proceedings (2017)
7. Madlberger, M.: Can Data Quality Help Overcome the Penguin Effect? The Case of Item Master Data Pools. In: ECIS 2011 Proceedings (2011)
8. Schemm, J., Legner, C.: The Role and Emerging Landscape of Data Pools in the Retail and Consumer Goods Industries. In: HICSS 2008 Proceedings (2008)
9. GS1: Global Data Synchronisation Network (GDSN) Operating Roadmap for GS1 Data Excellence. Version 7.4. Brüssel (2015)
10. de Corbière, F.: Interorganizational Information Systems and Data Quality Improvement: The Case of Product Information in the French Large Retail Industry. In: ICIQ 2007 Proceedings (2007)
11. Dalmolen, S., Moonen, H., van Hillegersberg, J.: Industry-wide Inter-organizational Systems and Data Quality: Exploratory Findings of the Use of GS1 Standards in the Dutch Retail Market. In: AMCIS 2015 Proceedings (2015)
12. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, xiii–xxiii (2002)
13. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: ECIS 2009 Proceedings (2009)

14. Schemm, J.W.: Zwischenbetriebliches Stammdatenmanagement. Lösungen für die Datensynchronisation zwischen Handel und Konsumgüterindustrie. Springer-Verlag Berlin Heidelberg (2009)
15. Chen, H.-C., Prater, E.: Information System Costs of Utilizing Electronic Product Codes in Achieving Global Data Synchronization within the Pharmaceutical Supply Chain Network. IJISSCM 6, 62–76 (2013)
16. Rogers, E.M.: Diffusion of Innovations. Free Press, New York (1995)
17. Petter, S., Delone, W.H., McLean, E.R.: Information Systems Success: The Quest for the Independent Variables. JMIS, 7–62 (2013)
18. Delone, W.H., McLean, E.R.: Information System Success: The Quest for the Dependent Variable. ISR 1, 60–95 (1992)
19. de Corbière, F. de: Data Quality and Interorganizational Information Systems. The Role of Electronic Catalogues. In: AMCIS 2009 Proceedings (2009)
20. Boukef Charki, N., Charki, M.H., Limayem, M.: Making & Giving Sense to the Global Data Synchronization Network Standard Adotpion. In: ECIS 2011 Proceedings (2011) [Spelling error adopted from original paper.]
21. Schemm, J.W., Legner, C., Österle, H.: Global Data Synchronization – Hype oder Lösung für das überbetriebliche Produktstammdatenmanagement zwischen Konsumgüterindustrie und Handel? In: Tagungsband Handelsinformationssysteme 2006 (2006)
22. Zhou, D., Nakatani, K., Chuang, T.T.: Data quality in Collaborative Commerce. IJIQ 2, 264–278 (2011)
23. Schemm, J.W., Legner, C., Österle, H.: Global Data Synchronization - Lösungsansatz für das überbetriebliche Produktstammdatenmanagement zwischen Konsumgüterindustrie und Handel? In: Becker, J., Knackstedt, R., Pfeiffer, D. (eds.) Wertschöpfungsnetzwerke. Konzepte für das Netzwerkmanagement und Potenziale aktueller Informationstechnologien, pp. 173–192. Physica-Verlag HD, Heidelberg (2008)
24. Winkelmann, A., Beverungen, D., Janiesch, C., Becker, J.: Improving the Quality of Article Master Data: Specification of an Integrated Master Data Platform for Promotions in Retail. In: ECIS 2008 Proceedings (2008)
25. Becker, J., Winkelmann, A., Beverungen, D., Janiesch, C.: Bereitstellung von Artikelstammdaten. HMD 44, 45–56 (2007)
26. Zhou, D., Chuang, T.-T., Nakatani, K.: A Conceptual Data Quality Framework for Collaborative Commerce. AMCIS 2006 Proceedings (2006)
27. Karpischek, S., Michahelles, F., Fleisch, E.: Detecting Incorrect Product Names in Online Sources for Product Master Data. EM 24, 151–160 (2014)
28. Karpischek, S., Michahelles, F., Fleisch, E.: The Not so Unique Global Trade Identification Number: Product Master Data Quality in Publicly Available Sources. In: ICEC 2012 Proceedings (2012)
29. Hüner, K.M., Schierning, A., Otto, B., Österle, H.: Product Data Quality in Supply Chains: The Case of Beiersdorf. EM 21, 141–154 (2011)
30. Nakatani, K., Chuang, T.T.: Global Data Synchronisation: Deployment and Implementation Issues of a Global Standard-based Information System. IJSS 8, 157 (2012)
31. Becker, J., Winkelmann, A., Beverungen, D., Janiesch, C.: Stammdatenkonzept für elektronische Verkaufsförderungsaktionen im Wertschöpfungsdreieck Hersteller, Händler und Kunde. In: Becker, J., Knackstedt, R., Pfeiffer, D. (eds.) Wertschöpfungsnetzwerke. Konzepte für das Netzwerkmanagement und Potenziale aktueller Informationstechnologien, pp. 193–207. Physica-Verlag HD, Heidelberg (2008)
32. de Corbière, F., Rowe, F.: Adoption Factors of Electronic Data Exchange and Technology: Can We Distinguish Two Phases? In: ECIS 2011 Proceedings (2011)

Towards Sociotechnical Management of Intra-Organisational Knowledge Transfer

Jonas Oppenlaender, Jesse Josua Benjamin, Claudia Müller-Birn

Institute of Computer Science, Freie Universität Berlin
{jonas.oppenlaender,jesse.benjamin,clmb}@fu-berlin.de

Abstract. The internal transfer of knowledge is a critical success factor for research organisations. However, little emphasis is placed on the intra-organisational transfer of knowledge due to the focus on knowledge transfer to external stakeholders. We introduce our research in progress project that aims to uncover opportunities for the intra-organisational transfer of knowledge at a research organisation. We propose a sociotechnical system that supports the management of the internal transfer of knowledge. Furthermore, we introduce the organisational role of the Knowledge Transfer Manager, which is a central part of the proposed sociotechnical system.

Keywords: knowledge transfer, knowledge management, knowledge sharing, research information network, linked data

1 Introduction

Making efficient use of knowledge is a critical success factor in research organisations [1–2]. Taking a communication theory perspective [3], the intra-organisational transfer of knowledge in research organisations is realized via the exchange of information between members of the organisation. This internal transfer of knowledge (IKT) requires coordination mechanisms [4]. This coordination of the IKT in today's research organisations is left either unmanaged or in the hands of supporting functions, such as knowledge and technology transfer offices [5]. However, little emphasis is placed on the intra-organisational transfer of knowledge due to the focus of the knowledge transfer activities on external stakeholders [6].

In this paper, we take a computer-supported cooperative work perspective of knowledge management [7]. Thus, we propose a sociotechnical system that supports both the self-directed social sharing of knowledge between researchers and the directed management of the inner-organisational advice-seeking relationships in research organisations. We propose that a new organisational role, the Knowledge Transfer Manager (KTM), is in charge of managing the interpersonal relationships at the organisation. We use semantic web technologies for integrating heterogeneous data distributed in disparate information silos in the research organisation with the aim of supporting the KTM in the identification of actionable opportunities for ITK based on this data. Our overarching aim is to generate useful and actionable insights into the

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

research activities and competencies of the research organisation, to support the active management of the IKT and to overcome the "stickiness" [4] impeding the transfer of knowledge in research organisations.

The contributions of this paper are as follows. We explain the need for managing the transfer of knowledge within research organisations. We outline a sociotechnical system addressing these challenges and enabling research organisations to actively manage their internal knowledge transfer. We discuss the role of the Knowledge Transfer Manager in this system and explain the benefits of our approach. Lastly, we give an illustrative example of a potential for knowledge transfer.

The paper is structured as follows. We locate our project in the research area and highlight some of the challenges of internal knowledge transfer in research organisations in section 2. We explain our preliminary research questions and design in section 3. The proposed solution is presented in section 4 and we conclude in section 5.

2 Knowledge Management in Research Organisations

From a strategic perspective [8, p. 32ff.], the intra-organisational knowledge transfer may be oriented along two strategies. The traditional view of knowledge management follows a repository model [7, p. 532]. According to this strategy, the intra-organisational transfer of knowledge is implicitly assumed as a consequence of codifying knowledge in systems, such as knowledge bases (e.g. Wikis) and research information systems. However, this transfer of knowledge is not guaranteed [9]. The hope of knowledge transfer through codification is frequently not realised [10, p. 352].

The second generation of knowledge management [7, p. 546] follows a personalization strategy [11]. This strategic view recognises that knowledge is a tacit or explicit property inherent in the individual [12]. The strategy encourages individuals to share their knowledge directly with other individuals with the support of recommendation systems (e.g. expertise locators) and, more recently, in social networks and research networking systems (RNS).

The RNS are systems which support individual researchers in identifying and maintaining relationships to other researchers [13]. The RNS may, for example, help junior researchers who lack extensive personal networks in finding information [14]. The RNS are gaining in popularity in universities [15], with the VIVO system as a prominent example [16]. VIVO codifies the competencies and research activities of researchers in a repository accessible by other researchers at the research organisation.

When following the personalization strategy, the intra-organisational transfer of knowledge may be inhibited by many factors. As research funding has moved to a result-based model [17], the transfer of knowledge to external stakeholders has been the main centre of attention of the research on knowledge transfer [18]. Research on factors influencing intra-organisational knowledge transfer in research organisations, such as universities, is, therefore, sparse [19, p. 124]. However, we can draw conclusions from research in other contexts. The three key aspects that influence intra-organisational knowledge transfer are technology, organisational structure and culture [20]. We discuss each of these challenges briefly in the context of research networking.

Technology. Information technology can foster the transfer of knowledge by extending the individuals' reach beyond their personal networks [8, p. 95]. The lack of an appropriate system and absence of mechanisms of coordination have been identified as significant barriers to knowledge sharing [21]. While RNS have experienced an uptake in recent years [15], these systems follow the codification strategy. Finding implicit opportunities for networking with other researchers in these systems is a time-intensive process. The researchers do not have the time to use the RNS to its full potential due to their high workload [22].

Organisation. Leadership plays a pivotal role in intra-organisational knowledge transfer "through providing opportunities for and managing the processes" [19, p. 125]. Knowledge and technology transfer offices have been created in most universities and research centres to manage the dissemination of knowledge. However, little emphasis is placed on the intra-organisational transfer of knowledge due to the commercial focus on transfer of knowledge to external stakeholders [6]. This focus weakens the position of knowledge managers in the organisation. The knowledge and technology transfer managers have a "dual identity" that requires them to consider the needs of both external and internal stakeholders [23]. The influence that knowledge transfer managers can exert on members of the research organisation regarding improving the IKT is, therefore, limited.

Culture. Culture has been identified as one of the most important factors that impedes intra-organisational knowledge transfer [24, 25, 26]. While researchers were found to have a positive attitude towards the sharing of knowledge [19], the knowledge sharing practices vary between the disciplines [27]. Knowledge hiding [28] or hoarding [29] behaviour may impede the sharing of knowledge. Having a working social network and intact communication channels may help the individual to overcome these knowledge transfer barriers. However, even if the individual researchers maintain their social relations, the researchers are unlikely to encounter new knowledge in their immediate personal networks [8, p. 95]. Expanding the researchers' knowledge requires relationship maintenance and the expansion of the individuals' social networks.

The information above underlines the need for an active approach to managing the intra-organisational transfer of knowledge in research organisations. Since RNS follow the codification strategy, taking full advantage of RNS, such as VIVO, would require the augmentation of the systems with additional reasoning capabilities [16, p. 25]. Our proposed system fills this gap by combining codification of research project information with personalization facilitated by the provision of reasoning capabilities.

3 Research Questions and Research Design

Motivated by the need for an integrative and adaptive approach to the management of the IKT, we formulate our design problem as follows[1]: *Accelerate knowledge exchange in a research organisation by designing a sociotechnical system that satisfies the requirements of the knowledge transfer manager and researchers to support the*

[1] Following the template for technical research questions given by [30].

knowledge transfer manager in matching researchers who have mutual interests and complementary expertise.

We are in the very early stages of this project. Our preliminary knowledge goals are:

1. What are the data sources available in research organisations that can be integrated? What relationships can be inferred between the members of the organisation, based on these data sources?
2. How can we present the inferred relationships to produce insights that are useful and actionable for both knowledge managers and researchers?

As we are building and evaluating a complex computational artefact, we follow the Design Science approach [31] and the Design Science Methodology by Peffers et al. [32]. Given the project context and problem statement above, we will elicit and define the requirements of the solution in a user-centred approach with researchers at a research organisation. The system envisioned consists of a prototypically implemented sociotechnical architecture. Following the design and implementation of this prototype, we will demonstrate and evaluate the solution in a research organisation in Germany. The prototype will be evaluated on two levels (data and system) and with two stakeholder groups (KTM and researchers). Our work will also include the application of measurement instruments for knowledge transfer [24].

4 Sociotechnical Management of Internal Knowledge Transfer

Researchers are flexibly matched for the transfer of knowledge based on *in-situ* project requirements in our approach to knowledge management. Our proposed architecture (Figure 1) aims to support the facilitation and management of the relations and communications between the members of the organisation.

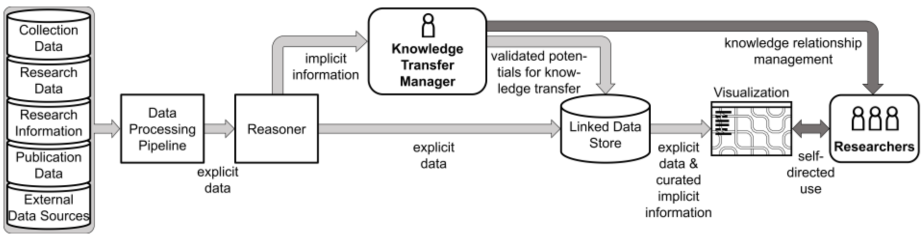


Figure 1. High-level architecture of the sociotechnical system proposed

The main data sources of the system are structured descriptions of research projects at the research organisation. We will acquire several other heterogeneous data sources (such as research collection data and publication metadata). A data processing pipeline integrates the heterogeneous data, describes it with the Resource Description Framework (RDF) vocabulary [33] and stores it in a Linked Data store. The ontological basis for the conversion to RDF is the "Core Dataset Research" [34], a normative recommendation for the description of research projects in the German scientific system. Together with representatives of the research organisation, this ontology [35] will be extended to describe the competencies, methods and attributes of current research activities in research organisations. We will employ reasoning mechanisms (e.g. rule reasoning or feature-based machine learning) on this research project data to identify the implicit potentials for intra-organisational knowledge transfer.

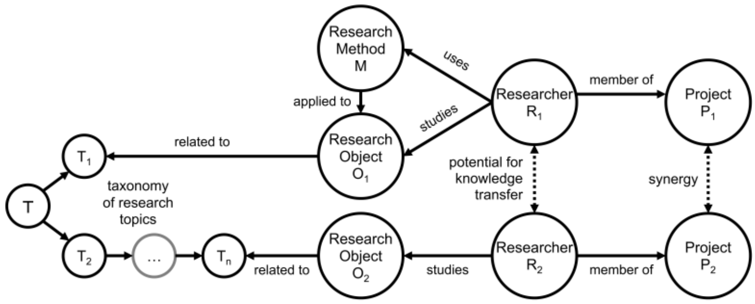


Figure 2. Example of a potential for knowledge transfer between two researchers

An example of a reasoning mechanism is illustrated in Figure 2. Two researchers, R_1 and R_2, study research objects, O_1 (e.g. *Crustacean*[2]) and O_2 (e.g. *Scleractinia*[3]), respectively. The objects are part of a taxonomy of research topics $T_1 \ldots T_n$ (e.g. the *World Register of Marine Species*). One of the researchers used a research method M (e.g. the *MNCR methodology*, an expert diver survey technique of marine biology). Since this research method could benefit the other researcher in the study of O_2, we can infer a new relationship between the two researchers.

Central to this architecture is a new organisational role, the Knowledge Transfer Manager. The KTM acts as both a gatekeeper and boundary spanner [36] in the sociotechnical system. The former is realised by reviewing and judging the inferred relationships between the knowledge workers provided by the reasoner. Only the validated relationships pass the gate and are accessible by other members of the organisation. The latter is realized by acting as a facilitator of knowledge transfer and as a manager of knowledge relationships. Based on the information inferred, the KTM may take steps to facilitate the communication between the researchers matched. We argue that the information inferred is a valuable support for the KTM's day-to-day work. The information enhances the KTM's capability of recognising actionable opportunities for the intra-organisational exchange of knowledge. The system proposed will support the KTM in 1) matching knowledge workers and facilitating advice-seeking to improve the knowledge utilization at the organisation, 2) catalysing collaboration at the organisation, 3) designing, managing and cultivating a knowledge exchange community and 4) fostering a culture of open innovation.

The validated knowledge transfer potentials will be presented in an adaptable and interactive visualisation accessible by all employees of the organisation. Through this visualisation, the system additionally supports the self-organised networking and competence development of the researchers.

[2] For example, crabs, lobsters, crayfish, shrimp, etc.
[3] Scleractinia are a biological order of corals.

5 Conclusions

We reported on a research in progress project that aims to increase the knowledge transfer within a research organisation. We introduce a new organisational role, the Knowledge Transfer Manager, which is supported by a proposed sociotechnical system. The system materialises opportunities for intra-organisational knowledge transfer. We aim to make implicit potentials for internal knowledge transfer actionable by integrating the manager into the system.

Acknowledgement: This work is supported by the German Federal Ministry of Education and Research, grant 03IO1633 ("IKON – Visualizing the potential for knowledge transfer in research museums").

References

1. Conner, K.R., Prahalad, C.K.: A Resource-Based Theory of the Firm: Knowledge Versus Opportunism. Organ. Sci., 7(5), 477–501 (1996)
2. Argote, L., Ingram, P.: Knowledge Transfer: A Basis for Competitive Advantage in Firms. Organ. Behav. Hum. Dec. Proc., 82(1), 150–169 (2000)
3. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana and Chicago (1963)
4. Szulanski, G.: Exploring Internal Stickiness: Impediments to the Transfer of Best Practice Within the Firm. Strategic Manag. J., 17(S2), 27–43 (1996)
5. Macho-Stadler, I., Pérez-Castrillo, D., Veugelers, R.: Licensing of University Inventions: The Role of a Technology Transfer Office. Int. J. Ind. Organ., 25(3), 483–510 (2007)
6. Chugh, R.: Workplace Dimensions: Tacit Knowledge Sharing in Universities. J. Adv. Manag. Sci., 1(1), 24–28 (2013)
7. Ackerman, M.S., Dachtera, J., Pipek, V., Wulf, V.: Sharing Knowledge and Expertise: The CSCW View of Knowledge Management. Comput. Support. Coop. Work, 22(4–6), 531–573 (2013)
8. Gottschalk, P.: Strategic Knowledge Management Technology. Idea Group Publishing, Hershey et al. (2005)
9. Argote, L., Miron-Spektor, E.: Organizational Learning: From Experience to Knowledge. Organ. Sci., 22(5), 1123–1137 (2011)
10. Porter, M.E.: Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, New York (1985)
11. Hansen, M.T., Nohria, N., Tierney, T.: What's Your Strategy for Managing Knowledge? Harv. Bus. Rev., 77(2), 106–116 (1999)
12. Nonaka, I.: A Dynamic Theory of Organizational Knowledge Creation. Organ. Sci., 5(1), 14–37 (1994)
13. Schleyer, T., Butler, B.S., Song, M., Spallek, H.: Conceptualizing and Advancing Research Networking Systems. ACM Trans. Comput. Hum. Interact., 19(1), Article 2 (2012)
14. Spallek, H., Schleyer, T., Butler, B.S.: Good Partners Are Hard To Find: The Search for and Selection of Collaborators in the Health Sciences. Proc. Int. Conf. on eScience, IEEE Computer Society, Indianapolis, IN, 462–467 (2008)

15. Kahlon, M., Yuan, L., Daigre, J., Meeks, E., Nelson, K., Piontkowski, C., Reuter, K., Sak, R., Turner, B., Weber, G.M., Chatterjee, A.: The Use and Significance of a Research Networking System. J. Med. Internet Res., 16(2), e46 (2014)
16. Börner, K., Conlon, M., Corson-Rikert, J., Ding, Y.: VIVO: A Semantic Approach to Scholarly Networking and Discovery. Morgan & Claypool Publishers (2012)
17. Bentley, P.J., Gulbrandsen, M., Kyvik, S.: The Relationship between Basic and Applied Research in Universities. High. Educ., 70(4), 689–709 (2015)
18. Bozeman, B.: Technology Transfer and Public Policy: A Review of Research and Theory. Res. Policy, 29(4), 627–655 (2000)
19. Fullwood, R., Rowley, J. Delbridge, R.: Knowledge Sharing amongst Academics in UK Universities. J. Knowl. Manag., 17(1), 123–136 (2013)
20. Gold, A.H., Malhotra, A., Segars, A.H.: Knowledge Management: An Organizational Capabilities Perspective. J. Manag. Inf. Syst., 18(1), 185–214 (2001)
21. Ajmal, M., Helo, P., Kekäle, T.: Critical Factors for Knowledge Management in Project Business. J. Knowl. Manag., 14(1), 156–168 (2010)
22. Asrar-ul-Haq, M., Anwar, S.: A Systematic Review of Knowledge Management and Knowledge Sharing: Trends, Issues, and Challenges. Cogent Business & Management, 3(1), 1127744, 17 pages (2016)
23. O'Kane, C., Mangematin, V., Geoghegan, W., Fitzgerald, C.: University Technology Transfer Offices: The Search for Identity to Build Legitimacy. Res. Policy, 44(2), 421–437 (2015)
24. Argote, L., Fahrenkopf, E.: Knowledge Transfer in Organizations: The Roles of Members, Tasks, Tools and Networks. Organ. Behav. Hum. Dec. Proc., 136(C), 146–159 (2016)
25. McDermott, R., O'Dell, C.: Overcoming Cultural Barriers to Sharing Knowledge. J. Knowl. Manag. Pract., 5(1), 76–85 (2001)
26. De Long, D.W., Fahey, L.: Diagnosing Cultural Barriers to Knowledge Management. Acad. Manag. Exec., 14(4), 113–127 (2000)
27. Landry, R., Amara, N., Ouimet, M.: Determinants of Knowledge Transfer: Evidence from Canadian University Researchers in Natural Sciences and Engineering. J. Technol. Transf., 32(6), 561–592 (2007)
28. Connelly, C.E., Zweig, D., Webster, J., Trougakos, J.P.: Knowledge Hiding in Organizations. J. Organ. Behav., 33(1), 64–88 (2012)
29. Bock, G.W., Zmud, R.W., Kim, Y.G., Lee, J.N.: Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate. MIS Quarterly, 29(1), 87–111 (2005)
30. Wieringa, R.J.: Design Science Methodology for Information Systems and Software Engineering. Springer, Berlin and Heidelberg (2014)
31. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly, 28(1), 75–105 (2004)
32. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. J. Manag. Inf. Syst., 24(3), 45–77 (2007)
33. Schreiber, G., Raimond, Y. (eds.): RDF 1.1 Primer. W3C Working Group Note, 24 June (2014)
34. Biesenbender, S., Hornbostel, S.: The Research Core Dataset for the German Science System: Developing Standards for an Integrated Management of Research Information. Scientometrics, 108(1), 401–412 (2016)
35. Gruber, T.R.: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Int. J. Hum. Comp. Stud., 43(5–6), 907–928 (1995)
36. Allen, T.J.: Managing the Flow of Technology. MIT Press, Cambridge, MA (1977)

A Conceptual Model of Benchmarking Data and its Implications for Data Mapping in the Data Economy

Markus Spiekermann[1], Sven Wenzel[1], Boris Otto[1,2]

[1]Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany
{markus.spiekermann,sven.wenzel,boris.otto}@isst.fraunhofer.de
[2]TU Dortmund, Chair of Supply Net Order Management, Dortmund, Germany
{boris.otto}@tu-dortmund.de

Abstract. Digitalization of the economy requires enterprises from all industries to revisit their current business models and prepare their organizations for the digital age. One task is the (re-)design of hybrid and digital products and services. The foundation builds the improved interchangeability of data and the availability of external data sources through data markets and platforms. This leads to the requirement of a structured decision-making while mapping data sources to digital products. In order to successfully transform their business and develop valuable new products, companies require methodological help. This paper proposes a high-level conceptual model for the assessment of data sources value. It consists of an approach for comparing data sources based on a common description of data and individual metrics definition enable a benchmark process. The development of the model and its practicability has been validated in a case study with an industrial partner.

Keywords: Data Economy, Data Assets, Data Valuation, Digital Transformation, Data Mapping

1 Introduction

The digital transformation challenges companies to review their business models and consider their digitalization by developing hybrid or purely digital products. Thereby, the foundational resource is data. While data, as a corporate resource seems to grow to infinity from the perspective of volume, variety and velocity, companies face challenges on the technical as well as on the organizational side. Data becomes a product itself [1, 2], and since the improved possibilities to exchange data in a secure and sovereign way, the amount of external data sources extends the internal available pool of master data, operational transaction data and other self-created data sources. On the organizational side, challenges are reflected by the digital transformation of business models and business strategies [3–5] which are strongly related to the topics of data valuation and monetization [1, 2, 6]. On the technological side integration of data sources, improving and controlling the data quality, provide traceability and provide access to possible data citizens within the company presents the greatest

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018,Lüneburg, Germany

challenges to data management [5, 7–9]. To solve these challenges, a tight integration of IT and business is required before companies can benefit from the vast amount of data. Structures that welcome data-driven business engineering and mapping the right data to their digital products have to be created. This paper deals with the research question on how to evaluate the most economical data source for a given business model in times of digital economy.

Research in the fields of digital transformation and the assessment of data or information has existed separately so far. This paper suggested a "data benchmark" approach to provide a concept for data mapping activities within the creation of digital business models, combine both topics, and improve practicability.

The goal is a theoretically sound and practically useful approach that may help companies to address the challenges of mapping data sources to their digital products and structure the decision-making by providing an applicable model.

Methodologically, the paper belongs to the domain of design science research (DSR), which strives at developing "IT artifacts intended to solve identified organizational problems" [10]. The model is developed based on a review of scientific literature and on ongoing action design research. The contribution of the proposed conceptual model for the academic community lies in getting deeper insights of challenges in the future business engineering and the derivation of methods to manage data resources. For practitioners, the model helps to implement a structured process of data assessment for mapping and selection of data sources while the company's digital transformation.

The remainder of the paper is structured as follows. Section 2 establishes the necessary background on research in the fields of digital economy and the changing role of data and its valuation. Section 3 presents the current version of the data-benchmark methodology and the corresponding conceptual approach. Section 4 then shows a real-life case to demonstrate the practicability and the evaluation of the method. Section 5 discusses the theoretical and managerial implications and limitations of the research. Section 6 concludes the paper and gives an outlook on future research.

2 Background

The theoretical background for this paper is given by two fields of research. First, the paper gives the necessary background for the topic of digital business and especially the activity of data mapping in the data economy. Second, the paper describes the changing role of data and the current situation of determining an economic value.

2.1 Digital Business

Digital Business is experiencing a renaissance at present. Initially coined in the 1990s, it is today used in broader context [3]. Traditionally the understanding of digital business was very much influenced by the debate around treating information as an enterprise asset [11–13]. This perspective on digital assets acknowledged the growing importance that data plays for enterprises business. Digitalization as the term of today's transformation and advancement of company's business model considers the opportunities and challenges, data and digital goods bring to business models. Research and practice broadly discuss the new Digital Business and the effects on business strategies and models. Examples are the MIT Center for Digital Business [14], the Digital Business Transformation at the University of St. Gallen [15] and the Industrial Data Space as consortium of Fraunhofer Society and the Industrial Data Space Association to develop an architecture for a digital ecosystem [16]. One design principle that is fundamental of digital business models is consumer-centricity and individualization, which is closely related to the independence of time and place by providing multi-channel integration [15]. Business models in the digital economy [3, 5, 14] are characterized by an evolvement of products into hybrid or purely digital services. The close integration of IT and physical products in combination with the vast amount of available data enables smart and data-driven service offerings for traditional products as well as new innovations to gain more added value around the core product [17].

One key action within the design of data-driven business is the data mapping, "making sure that the business objects required in the end-to-end customer process are transparent and that corresponding data objects are identified and described" [3]. Since the vast amount of available internal as well as external data sources, an economical selection of data is business critical for digital services. The economical selection of the right data sources gets harder with the growing volume and indeed cannot be done intuitive without a structured process in the age of digitalization and data economy [18, 19].

2.2 Data Assets and their Valuation

Information as a resource has been researched over a longer period. First approaches were made in the concepts of Data Resource Management, where influencing factors for information resources are described [20]. Literature in the 1990s suggested that information should be viewed as intangible assets [11, 21, 22], motivated by the development of information becoming an important corporate resource. While in the 1960s and 1970s data was the result of business processes, it enables a company-wide business process management in the 1980s and 1990s. Since the turn of the millennium, data plays a key role in enabling products and (digital) services, up to the emerging of data as a stand-alone product [16].

The management of these resources is still insufficient compared to other intangible as well as tangible resources [6, 23]. A direct transfer of known and established methods for managing other tangible as well as intangible assets to the data domain is problematic, because of differences in the characteristics of data. As an example, compared to other assets, data has no abrasion, is infinitely shareable and the value increases with its use [6, 12, 13]. While available research considers information assets, we propose to distinguish between information and data, and consider data as the building blocks of information. Consequently, this paper suggests to value data as the raw material for creating digital business instead of information. We assume that data plays an important role as an input for different business processes and has an important impact on the success of businesses [24], especially in the data economy where exchange of data resources between different companies gets more important to create products and provide better services. This makes the product data an important asset of the future company.

Valuation approaches were researched by considering the use value, the cost-of production value and the exchange value of information [25–28].

Despite of the research done in the past years, there is still a lack of practical concepts and methodologies for valuing information assets [2, 29]. Nevertheless current research shows the general possibility to value data as assets [2, 29]. Authors propose to apply existing valuation concepts (e.g. the cost-based approach, the market-based approach, and the income-based approach), without elaborate on new or modified approaches for the data valuation [29].

The cost-based approach quantifies either the cost of sourcing, exchange or reproducing data and managing it through the entire data lifecycle, by consider costs for data quality management and maintenance [2, 26, 28]. The approach of determine a use-based value quantifies the reduced uncertainty in decision-making [25, 27]. The basic idea is to determine the value of the business opportunities a company gains from using data in different business processes [29].

In the market-based approach, data assets are measured by observing the market value on competitive, active markets. The value of the data is then derived from the existing market demand. While the market-based approach revealed strong limitations for being used in the context of data valuation due to missing markets, this changes within the data economy, where platforms and marketplaces being established and can be used to compare different data products [16, 30]. . Comparable products mean that the value has to be adjusted concerning differences in for example data quality, volume, or content. A practical approach for the valuation of data sources would enable an economic data mapping for digital businesses [3]. In addition, it could provide transparency for the trade in data goods on arising data marketplaces and platforms [30, 31] and therefore implicates sourcing of external data sources to meet their business requirements.

In order to find the most economical data source, companies need to make a structured decision about which data they want to buy, based on key indicators and their metrics.

3 Research Approach

This research aims at developing a conceptual model that outlines a high-level process to structure the selection and mapping of data sources to digital business models from an economical perspective. A conceptual model therefore specifies the generally valid elements that are representative for a real system. The conceptual model can serve as a reference for designing company-specific models [32]. We designed the resulting artifact in a research project with a large pharmaceutical company since September 2016, following the Action Design Research (ADR) methodology. ADR combines design science (DS) and action research (AR) and constitutes “a research method for generating prescriptive design knowledge through building and evaluating ensemble IT artifacts in an organizational setting" [33].

After an initial discussion in September 2016 with industry partners about the current lack of a concept for an economic data mapping in digital business, the first phase considers the corresponding requirements for such an approach (*Problem Formulation*). Experienced data managers, that identified the need of action, initiated the research activities. Although these data managers are aware of the value, data provides to their business users, they lacking guidance for the data-related challenge of determining a specific economic value for a KPI-based decision-making. As suggested by Sein et al. [33], the first BIE cycle aims at developing an early alpha version of the artifact, which will be instantiated, repeatedly tested, and continuously refined in a second cycle. In the first cycle, the emerging artifact based on an implementation of a proof of concept, comparing a list of attributes of data sources with a scoring model, presented to a focus group end of September 2016 as depicted in **Table 1**.

Table 1. Participants of initial focus group

Date	*Location*	*Focus Topics*	*Participants*
2016-09	Duisburg (D)	Business in the data economy, value of data, innovation and digital transformation	9 participants from two companies including experienced data managers, architects and scientists

The discussion results, described in more detail in section 4.1, formed the basis for further adjustments of the artifact. The concept was repeatedly discussed and refined (*Building, Intervention, and Evaluation*) in a case study between November 2016 and July 2017. The first and second session focused on the general structure and procedure, while a further session considered practicability and visualization of the conceptual model.

In addition to the focus group discussions, we conducted a software-based evaluation started in March 2017, where we implemented the alpha version of the conceptual model as a usable software-tool for practitioners. The results of the evaluation are presented in section 5. In future BIE cycles, we will instantiate the artifact in selected companies. Based on the interventions and evaluations from these cases, a beta version of the artifact will be developed (*Reflection and Learning*). The research team has initiated the activities of this cycle and we are currently applying the conceptual model in several companies to ensure that learning from our company-specific instantiations are further developed and documented (*Formalization of Learning*) as a general solution.

4 The Data Benchmark Model – Conceptual Model for Data Valuation in the Digital Economy

4.1 Purpose and Guiding Requirements

The conceptual model of data benchmarking aims at structuring the decision-making for mapping data sources to data-driven products and digital business models. It addresses the challenges of the data economy in terms of purchasing and economic selection of external data. External data extends the existing internal data pool and has to be purchased from data marketplaces or platforms to improve company's products and services. Identifying these data needs of business requires – apart from the technical capabilities – close alignment between data management, sourcing and the data citizens as the consumer of data driven insights. Therefore, a concept of determine a value for data sources, that enables a structured decision-making and mapping the most economic data sources, is required. The concept represents the real system on a high-level model guided by identified key requirements and the mentioned research question. We derived these requirements and elements from the initial focus group discussion and a subsequent literature research following Webster and Watson [34]. We searched for "data asset", "information asset", "data assessment", and "data valuation" on Scopus, as well as Google Scholar database to retrieve our information. The key elements of the pursued conceptual model are summarized in Table 2.

Within the data economy, data is business critical and has to be managed as strategic asset within the companies that has an economic value (1) [1, 2, 6, 11]. Furthermore, data assets have attributes that can describe the characteristics of data or data sources (2) [20, 35, 36]. Attributes of data can be categorized and structured in a metadata model to unify a description (3) [36]. With economic, qualitative, technologic, and competitive we identified four categories of attributes. While economic attributes e.g. refer to the price, license and ownership of the data source, qualitative attributes refer to the completeness, timeliness, accuracy, or integrity. The technologic category sums up attributes like format, type, or availability, while the category of competitive attributes refer to accessibility, novelty, and publicity level. On base of the identified attributes, metrics can be defined that enable an assessment and valuation (4). Metrics have to be individualizable by requirements of the specific use cases (5) [29]. Benchmarking methodology for data sources with a peer group leads to a structured decision-support for data mapping (6).

Based on the identified requirements and taking into account the surrounding environment (technology and organization), we designed the alpha version of our artifact.

Table 2. Requirements for data mapping using a data benchmark approach

No.	*Requirement*	*Description*
(1)	Data is a corporate asset managed as a stand-alone product.	The role of data has changed and data is traded as a stand-alone commodity on data marketplaces and data spaces.
(2)	Data is described by its attributes.	Data has attributes and characteristics that describes the data. Attributes are related to data sources, data sets or to single values.
(3)	Attributes can be used to form a metadata model for common description.	The overall attributes and characteristics of data sources can be organized with a metadata model that serves as a standard description template and group attributes in different categories (e.g. economical, technological, qualitative, competitive)
(4)	Metrics can be defined to determine an economic value.	By creating metrics and indicators on base of the data's attributes, an economic value can be determined.
(5)	Specific use cases influence the value based on their requirements	Specific business use cases influence the requirements placed on the data, which has to be considered when selecting a data source to a targeted business model.
(6)	Metrics enable a comparison of data with peers	The combination of a unified description and the defined metrics enable a benchmark of a data source within its peer group.

4.2 Design Process

Given the understanding of data as a strategic resource for the digital economy, the structure of the conceptual model builds on established procedures of benchmarking.

Benchmarking is as primarily a tool for improvement, achieved through comparison. It identifies the highest standards of excellence for products, services, or processes [37]. This paper proposes an adaption of the benchmarking wheel [38], which is well suited to the dynamic changes in the data ecosystem, to develop a product benchmarking concept for data sources. It follows the structure of the five steps proposed by the benchmarking wheel to enable a comparison between data sources:

1. *Plan*: Determine what to benchmark
2. *Find*: Identify benchmarking partners
3. *Collect*: Understand and document the benchmarking partners' information and attributes
4. *Analyze*: Analyze benchmarking information due metrics and indicators

5. *Improve*: Use results to take action and improve for example the decision-making

To design a reusable artifact regarding the ADR methodology we took the results of our focus group discussions and literature review. We adapt the identified requirements of the considered use case of data mapping in digital business to develop a conceptual model. Therefore, we based our approach on the identified steps of a benchmarking process and then assign the necessary actions regarding a data source comparison as depicted in **Figure 1**.

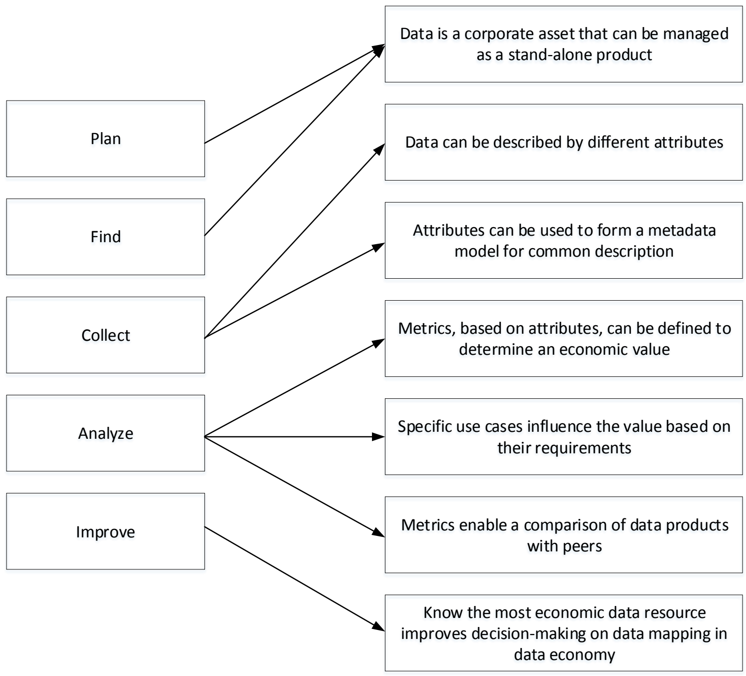


Figure 1. Assignment of data mapping requirements to the benchmarking process

Our conceptual model describes the benchmarking of data sources, resulting in the data sources to be the main object of consideration in the *"Plan"* phase of the benchmarking process. Since data becomes a stand-alone commodity in the digital ecosystem, data sources, as our peers for the benchmark, can be found on data marketplaces and (open) data platforms, which refers to the "*Find*" phase. To enable a benchmarking, available information of data sources have to be *"Collected"* and stored in a common information model. As our metadata model, we used the Metadata Model for Data Goods (M4DG)[1].

The phase *"Analyze"* in our approach is based on individualizable metrics that can be defined out of the attributes of the data sources regarding the information model. Regarding the identified requirements of the focus group discussions, we take into

[1] Metadata Model for Data Goods (M4DG) is a research project of the Fraunhofer ISST to develop a standardized model for data assets.

account that the specific business needs can influence the metrics while mapping data sources. The final *"Improve"* phase is about a comparison of data sources using the defined metrics. This finally leads to a determination of a data sources value and supports a structured decision-making, which data source to select.

As a final step in our design process, we concentrated on the graphical visualization to depict the key elements, relations and process flow of our conceptual model. The activities of the design process result in the conceptual model, depicted in **Figure 2**. The model is company-unspecific and intended to be reused in a design process of an information system.

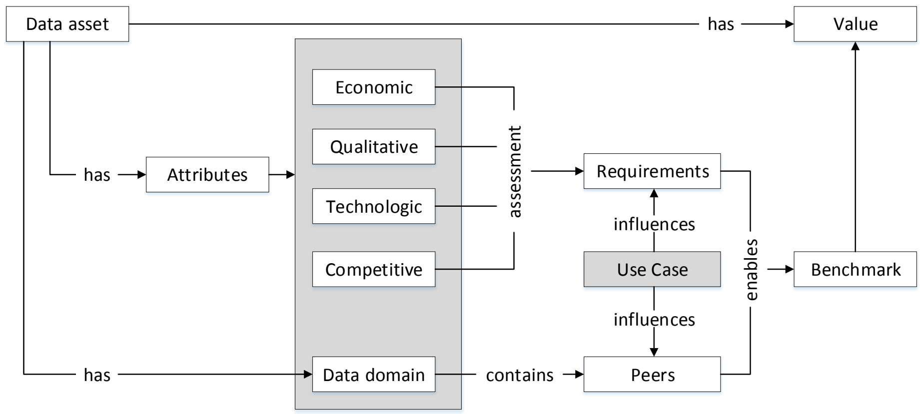


Figure 2. Conceptual model for data benchmarking

5 Evaluation

"Action research is the application and testing of ideas developed in an academic environment in real world situations under participation of the researching individuals" [39]. Therefore, we tested the validity of the designed conceptual model, the benchmarking approach, and the key elements in a case study with our industrial partner. We implemented the designed concept in a practical information system and presented the application to data management experts and data citizens, who need to map data to their digital business, for an evaluation. The information system, a web-based application, allows collecting data sources, searching for data sources, evaluating quality attributes, and defining metrics to be used for a comparison.

After a presentation of the conceptual model and the information system to different business areas, participants were asked to evaluate the structure (i.e. the completeness, simplicity, or clarity), the adaptability (i.e. robustness, practicability), and the environmental fit (i.e. utility, personal and organizational fit) of our approach.

We got a positive feedback for the chosen structure of our model. The idea behind our approach was straightforward and clearly understandable. Participants notice, that defining roles and responsibilities would make the model more complete in a next step. Regarding the environmental fit, the feedback was also very positive. The participants think it is very good that the topic of data mapping is examined and solutions are sought. Even when the problem is not yet critical for many business users, the view of future necessity in data economy is shared for personal as well as organizational interest. The problem relates to every business unit we presented our approach and our solution was applicable in all considered cases, since all plan or actually do purchase external data sources for their business processes. This leads to the assumption of a positive adaptability of our concept having in mind, that a high effort of collect the required information and connect to platforms and marketplaces is necessary. That is why we get the feedback that the collection of data sources and its information could be time consuming and should be automatized as much as possible, in terms of the practicability of our approach.

6 Conclusion and Outlook

Within the data economy and the requirement to purchase external data this topic gets even more important to the practitioners. This paper presents a conceptual model for a data benchmark to meet the challenges in mapping the most economic data sources while designing digital business models. The emerging model was systematically developed following the four steps proposed by Action Design Research. The conceptual model reflects the key elements, assumptions and requirements related to the identified real life problems. These requirements were collected through discussions with experienced data architects, data managers, and data citizens. Because we see data as a stand-alone commodity of future business, we then designed a conceptual model based on the process of a product benchmarking reflecting data sources as products within the data economy. Using a common description for data sources and the definition of metrics related to the data sources, our concept enables a comparison between a data source and its peers to determine a value.

The evaluation of the alpha version of the artifact as well as a first prototype implementation have demonstrated its utility and organizational fit. Hereby limitations stem from the research context as only industrial partner was involved while the first evaluation phase. Despite being a conceptual model that strives for a transferable and company-unspecific degree as a scientific artifact [40], we will need more empirical testing to evaluate and complement our concept. The idea behind our concept was well received by the data managers as well as data citizens we presented our approach within the case study.

Further research activities are currently ongoing to instantiate the artifact in further companies and refine the concept in another BIE cycle. Planned results of these research activities include instantiations as well as using the conceptual model as basis to develop a framework for an economic valuation of data as corporate resource.

7 Acknowledgments

This work has been funded by the German Federal Ministry of Education and Research (BMBF) (project InDaSpace, grant no. 01IS15054).

References

1. Wang, R. Y., Lee, Y., Pipino, L., Strong, D.: Managing your information as a product. Sloan Manag. Rev. 95–106 (1998).
2. Otto, B.: Quality and Value of the Data Resource in Large Enterprises. Inf. Syst. Manag. 32, 234–251 (2015).
3. Otto, B., Bärenfänger, R., Steinbuß, S.: Digital Business Engineering: Methodological Foundations and First Experiences from the Field. Proc. 28th Bled eConference. 58–76 (2015).
4. Brown, B., Sikes, J.: Minding your digital business: McKinsey Global Survey results. McKinsey Bus. Technol. 27, 24–31 (2012).
5. Zimmermann, H.: Understanding the Digital Economy : Challenges for new Business Models. AMCIS 2000 Proc. Paper 402, 729–732 (2000).
6. Moody, D., Walsh, P.: Measuring The Value Of Information: An Asset Valuation Approach. Seventh Eur. Conf. Inf. Syst. 1–17 (1999).
7. Agrawal, D., Bernstein, P., Bertino, E.: Challenges and Opportunities with Big Data. Proc. VLDB Endow. 5, 1–15 (2012).
8. Labrinidis, A., Jagadish, H. V.: Challenges and opportunities with big data. Proc. VLDB Endow. 5, 2032–2033 (2012).
9. Ammu, N., Irfanuddin, M.: Big Data Challenges. Int. J. Adv. Trends Comput. Sci. Eng. 2, 613–615 (2013).
10. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q. 28, 75–105 (2004).
11. Horne, N.W.: Information as an asset—The board agenda. Comput. Audit Updat. (1995).
12. Oppenheim, C., Stenson, J., Wilson, R.M.S.: The attributes of information as an asset. New Libr. World. 102, 458–464 (2001).
13. Stenson, J.: The Attributes of Information As an Asset, (2006).
14. The MIT Center for Digital Business: MIT initiative on the digital economy, http://ide.mit.edu/.
15. Leimeister, J.M., Österle, H., Alter, S.: Digital services for consumers. Electron. Mark. 24, 255–258 (2014).
16. Otto, B., Auer, S., Cirullies, J., Jürjens, J., Menz, N., Schon, J., Wenzel, S.: Industrial Data Space: Digital Sovereignty over Data. (2016).
17. Yoo, Y., Henfridsson, O., Lyytinen, K.: Research Commentary—The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. Inf. Syst. Res. 21, 724–735 (2010).
18. Newman, D.: How to Plan, Participate and Prosper in the Data Economy. Gartner, Stamford, CT. (2011).
19. Dennedy, M.F., Fox, J., Finneran, T.R.: Value and Metrics for Data Assets. In: The Privacy Engineer's Manifesto: Getting from Policy to Code to QA to Value. pp. 279–298. Apress, Berkeley, CA (2014).
20. Goodhue, D.L., Quillard, J. a, Rockart, J.F.: Managing the Data Resource: A Contingency Perspective. MIS Q. 12, 373–392 (1988).

21. Glazer, R.: Measuring the Value of Information: The Information-Intensive Organization. IBM Syst. J. 32, 99–110 (1993).
22. Fisher, T.: The Data Asset: How Smart Companies Govern Their Data for Business Success. John Wiley & Sons, Inc., Hoboken, New Jersey (2009).
23. Oppenheim, C., Stenson, J.: Studies on information as an asset I: defnitions. J. Inf. Sci. 29, 159–166 (2003).
24. Wilson, R.M.S., Stenson, J.A.: Valuation of information assets on the balance sheet. Bus. Inf. Rev. 25, 167–182 (2008).
25. Hirshleifer, J.: Where Are We in the Theory on Information? Am. Econ. Rev. 63, 31–39 (1973).
26. Hilton, R.W.: The Determinants of Information Value: Synthesizing Some General Results. Manage. Sci. 27, 57–64 (1981).
27. Stigler, G.J.: The Economics of Information, (1961).
28. Repo, A.J.: The value of information: Approaches in economics, accounting, and management science. J. Am. Soc. Inf. Sci. 40, 68–85 (1989).
29. Zechmann, A.: Assessing the Economic Value of Data Assets. Work Rep. (2016).
30. Schomm, F., Stahl, F., Vossen, G.: Marketplaces for Data: An Initial Survey. SIGMOD Rec. 42, 15–26 (2013).
31. Heckman, J., Peters, E., Kurup, N.G., Boehmer, E., Davaloo, M.: A Pricing Model for Data Markets. iConference 2015 Proc. (2015).
32. Brocke, J., Buddendick, C.: Reusable Conceptual Models – Requirements Based on the Design Science Research Paradigm. Int. Conf. Des. Sci. Res. Inf. Syst. Technol. - DESTRIST 2006. 576–604 (2006).
33. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.I.: Action design research. MIS Q. 30, 611–642 (2011).
34. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future : Writing a Literature Review. 26, (2002).
35. Ahituv, N.: Assessing the Value of Information: Problems and Approaches. ICIS 1989 Proc. Paper 45 (1989).
36. Garifova, L.F.: Infonomics and the Value of Information in the Digital Economy. 2nd Glob. Conf. BUSINESS, Econ. Manag. Tour. 23, 738–743 (2014).
37. Bhutta, K., Huq, F.: Benchmarking–best practices: an integrated approach. Benchmarking An Int. J. Vol. 6, 254–268 (1999).
38. Camp, R.C.: Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance. (1989).
39. Recker, J.: Conceptual model evaluation. Towards more paradigmatic rigor. In: CEUR Workshop Proceedings. pp. 183–194 (2005).
40. Vom Brocke, J.: Design Principles for Reference Modeling: Reusing Information Models by Means of Aggregation, Specialisation, Instantiation, and Analogy. In: Innovations in Information Systems Modeling: Methods and Best Practices. pp. 47–76 (2009).

A Metadata Model for Data Goods

Markus Spiekermann[1], Daniel Tebernum[1], Sven Wenzel[1], and Boris Otto[1,2]

[1]Fraunhofer ISST, Digitization in Service Industries, Dortmund, Germany
{firstname.surname}@isst.fraunhofer.de
[2]TU Dortmund, Chair of Supply Net Order Management, Dortmund, Germany
{firstname.surname}@tu-dortmund.de

Abstract. Digital transformation implies the development of data-driven business models and thus the management of data goods. While marketplaces for data are being established and platforms for the exchange of data are being created, companies have to adapt their data management to the increasing requirements. One central question can be deduced: How can data goods be described in a standardized way? This paper describes the development of the metadata model for data goods M4DG. The M4DG, based on an analysis of existing data marketplaces and metadata models of related topics, makes it possible to describe data sources with defined properties. This creates a unified understanding of the properties of data goods to facilitate selection and trading. We are convinced that the M4DG will contribute to the practical design of data management.

Keywords: data management, metadata model, data goods, data economy

1 Introduction

The role of data has changed from an entrepreneurial resource, used in the individual processes of a company, to an independent product. For example, platforms such as OpenWeatherMap[1] not only use the data from connected weather stations as their business resource. In addition to free access to weather information, they also offer more detailed and finely granular data in exchange for money. Among other things, this development can be attributed to the emergence of platforms and marketplaces on which data of different categories, different characteristics and under different conditions are offered. On the other hand, a general change in companies' business models can be seen. As a result of the expansion of digital transformation in companies, there is a shift from standardized and material-based business models to highly individual, hybrid or purely digital products [1]. In contrast to material goods, however, data has special characteristics that prevent the simple transfer of established processes and rules of material trade [2, 3]. This has an impact on the supply chain and data exchange between companies.

[1] https://openweathermap.org, accessed 10.11.2017

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

For the emerging trade of data goods and their cross-company exchange, a uniform understanding of the properties of the data offered or purchased is necessary. A standardized description of data sources can improve interoperability and processes and rules for data management can be defined. This would pave the way for a data supply chain and a sovereign exchange of data, which is aimed at by many current initiatives [4, 5]. This raises the research question of whether and how data sources can be modeled in the sense of a classical economic good, a so-called data good. This paper proposes a model that allows data goods to be managed like corporate assets.

This paper describes the qualitative approach to the creation of a model for the standardized mapping of data goods. It shows how a suitable model was developed from existing approaches of creating metadata models in cooperation with companies and how a first version of a possible metadata model for data goods was developed and tested. The remainder of the paper is structured as follows: First, the relevant approaches, related topics and technologies within the thematic area are examined. The following is an explanation of the scientific methodology used. Section 4 describes the corresponding procedure for implementing the model design by explaining the qualitative data collection and the resulting decisions for the model development and an evaluation. In the following, the use of the model in the data backend of a software component is described in a practical application case in section 5. Section 6 summarizes the results of the report and gives an outlook on further research.

2 Background

2.1 Data Goods

The importance of data for business success has been growing steadily since the introduction of electronic data processing and the automation of production processes in companies. Although research and industry have been discussing the role of data as an asset since the 1980s, digitization has for the first time led to the use of the term "data-driven business" [6]. Digitization is not possible without data. The individual development stages in the role of data are not disjoint, but can be found in parallel in companies. This results in the so-called "data paradox". On the one hand, data is a consequence of the above-mentioned developments such as digitization and industry 4.0 (because machines, smart services, etc. produce an increasing amount of "big data"). On the other hand, it is a resource for the production of services or even the product itself. We consider data as an independent product that can be traded between companies. For this reason, we call data in this context "data goods", analogous to economic goods. When considering data as a product, it is essential to consider the special properties of the product. This is evident, for example, in the economic analysis of a price determination, a qualitative evaluation or the definition of user authorizations. The challenges of considering data as an economic asset are not a new topic in scientific literature [6–10]. Moody and Walsh describe the seven "laws" of information by highlighting the differences between data and material goods [3]. Despite the fact that the topic has been dealt with, there is still no practical approach to the topic. While other intangible assets such as patents, trademarks or human capital are shown in the balance sheet, this is currently not possible for data. As part of the process of digital transformation, however, the scientific community is once again increasingly concerned with the economic consideration of the resource for digital business processes and its value added. Data trading can also represent an independent business model [11].

2.2 Data Marketplaces and Platforms

We observed several independent data marketplaces where sellers can offer their data (see Table 1). The marketplaces allow the providers to make their data available to interested parties [12]. Various price models have been developed in which data can be made freely available or monetized via models such as costs per call-off, costs per volume or flat rates [11, 13]. Data marketplaces can therefore be defined as a platform that allows you to offer data for sale and to specify the rules and prices for the data [14]. Data marketplaces, as a platform for trading data, can have different architectures. One possibility is the creation of a central instance (cloud computing approach), where the data is offered by different vendors via a central point. The advantages of this approach are the high traceability through a central search for products, a uniform interface for the provision of data goods and the possible specification of formats and distribution channels.

Another platform architecture is the decentralized approach [15]. In this case, the data items are offered decentrally, i.e. independently by the sales staff. This gives the seller control over his data and enables him to regulate the use of the data. This means that data to be sold is not offered to the public but only in certain partnerships, such as a supply chain. A major disadvantage of this architecture is the large number of different interfaces, formats, price models and the difficulty of finding interesting data.

This variety of possible descriptions and interfaces is a well-known problem of data management, which is becoming increasingly important due to the increasing importance of cross-company data exchange. This illustrates the development of platforms for managing data and secure trade in digital goods between companies. One of these platforms is the internationally established Industrial Internet Consortium (IIC), which has been developing the architecture of a standardized platform for connecting human beings and machines since 2014 [5]. In Germany, a network of various institutions and associations is developing a reference model specifically for the context of Industry 4.0; the so-called Reference Architecture Model for Industry 4.0 (RAMI 4.0) [16]. The Industrial Data Space, a combination of research project and industry consortium, which describes a reference architecture and strives for a prototypical implementation, chooses a holistic approach with the focus on preserving data sovereignty when trading data goods [15, 17].

2.3 Metadata Models

A well-known metadata model is the Dublin Core Metadata Element Set (DCE) [18]. The DCE does not focus on the description of a specific application area, but on the general description of resources. However, DCE is limited to the basic properties such as *author*, *format*, *language*, *ID number* or *title*. In February 2009, the Dublin Core Metadata Initiative registered the DCE as ISO Standard 15836 [19]. For other models, DCE often serves as a basis for the extension of certain application contexts [20–23] due to its simplicity, comprehensibility and flexibility.

Based on the DCE, Maali et al. developed the Data Catalog Vocabulary (DCAT) [20]. DCAT is designed to increase interoperability between different data catalogs and simplify the search for data sources. For this purpose, properties of different metadata models from the field of data cataloguing have been compared and commonalities identified, which were then combined into a new model. Another metadata model based on the DCAT is used in the Comprehensive Knowledge Archive Network (CKAN). CKAN provides an open source software tool that allows data to be published on a platform as a self-hosted service. This is mainly used by public authorities or non-profit organizations [24]. Within the scope of Open Data initiatives, CKAN and its metadata model have become a de-facto standard. CKAN is also limited to the essential properties of the data sources to describe them [25].

Another metadata model in the context of describing digital assets is the Asset Description Metadata Schema (ADMS). It is a standard accepted by W3C and also an extension of existing models. ADMS is based on DCAT and defines itself as a special variant [26]. ADMS allows the description of assets within data catalogs. An asset can be a data source, vocabulary, or other metadata models. However, none of the metadata models found provides economic properties for a data source.

2.4 Metadata Models for Data Goods

Indicators supporting the decision to develop a new data model for data goods were identified in the previous background notes. Data is increasingly becoming the product itself. There are already metadata models for immaterial goods [20, 26, 27]. However, according to our research, there is no metadata model that reflects the requirements concerning data goods (see Chapter 2.1). The lack of a standard makes it difficult, for example, to develop federal data marketplaces that can easily exchange their offers among each other. Therefore, it is necessary to create a uniform understanding of the properties of data goods.

3 Methodological Approach

The creation of the desired metadata model for data goods followed the methodological approach of Design Research (DR) [28]. In particular, we chose the method of Action Design Research (ADR), which focusses on generalizing the results through an iterative approach to the creation of artifacts and application in several use cases. This prevents the results from being influenced too much by a single individual application. The iterative process of ADR goes through the phases “Problem Formulation”, “Building, Intervention, and Evaluation”, “Reflection and Learning” and “Formalization of Learning”. The procedure implies a close exchange between science and practice, in which both sides contribute their experience, knowledge and ideas to develop a solution to the problem [29]. The development of the model described in this paper was supported by practitioners in three workshops in the first half of 2017. The first partner is from the pharmaceutical industry. Employees from the area of data purchasing and data modeling were involved. The second partner is from the consulting sector. The Head of Business Design and the Head of IT supported us here. As a third partner, experts from the field of data modelling supported us. The goal of the partners is to provide a unified description of data sources, so that communication and exchange via automated interfaces can be simplified. From a scientific point of view, the main aim is to contribute to the gain of knowledge in the field of data management and to create a basis for further research.

In a first step, the scientists reviewed the current literature and conducted an analysis of existing data marketplaces (Table 1). Findings from related research projects that deal with data management have also been incorporated here. Afterwards, a first workshop with the members of the pharmaceutical company was held. They collected possible characteristics of data sources in a brainstorming session. The focus was on properties that support the idea of a company-wide inventory based on a standardized description and establishment of a transparent mapping of governance processes. Afterwards, a first metadata model for data goods was developed based on the research and workshop results. In a second workshop, the model was presented and discussed to the consulting company. This has yielded new findings that are incorporated into the next iteration of the development of the scientific artifact within the ADR process. The same procedure as in the second workshop was repeated in a third one with the data modelling experts.

4 Metadata Model for Data Goods

4.1 Initial Research

In order to determine the necessary properties for a description of data goods, we have carried out a literature analysis to include relevant work within the subject area. The literature analysis was conducted after Webster and Watson [30]. Springerlink, DBLP Computer Science Bibliography, Elsevier and ScienceDirect were used for the research. The keywords “Data Catalog”, “Data Source Description”, “Data Source Metadata”, “Data Asset”, “Data Inventory”, “Data Marketplace”, “Data Governance”, “Data Economy” and “Data Management” were used for the initial search. Further keywords were identified by screening the found literature. Only those works were considered which show in the abstract that they are positioned in the field of data and data source description or consider economic aspects of data. Literature that was not available in English or German was ignored. We found two existing vocabularies, Data Catalog Vocabulary (DCAT) and Asset Description Metadata Schema (ADMS), which had a strong relation to the desired result. While DCAT allows a rough description of data sources, the ADMS has been developed as a more specialized metamodel to describe vocabularies. In addition to literature research, we have looked at existing web-based data marketplaces. Due to the focus on data as a product, the analysis investigated those data marketplaces where data is offered for sale in order to monetize them. The data marketplaces considered represent internet platforms that allow a data exchange between a seller and a buyer. The data marketplaces considered in this report have been selected by means of a manual internet search for corresponding platforms. The list does not claim to be exhaustive, as new marketplaces are constantly being created, especially in this area, or access is only granted against payment. Table 1 shows which properties are represented by which data marketplace. By comparing the found properties of data sources from literature and the found properties of data marketplaces, we found a large consensus. We chose ADMS as a starting point for the further development of our model, as it has been identified during our initial research as the metamodel with the closest relation, the representation of vocabulary as a special class of data goods. Due to the aspired goal of a high level of standardization and generalization, the design decision was therefore made to adapt the ADMS [26], taking into account the different focus and specialization on data sources.

Table 1. Overview of the properties represented in data marketplaces

	dmi.io[2]	*Azure*[3]	*datastreamx*[4]	*Mashape*[5]	*bids4bytes*[6]
Title	X	X	X	X	X
Type				X	
Category	X	X	X	X	X
Dateadded	X	X			X
Provider	X	X	X	X	X
Description	X	X	X	X	X
Pricinginformation	X	X	X	X	X
Example	X	X	X	X	
Licensing information	X	X	X	X	X
User rating	X			X	
Structure	X		X	X	
URL (API)	X		X	X	
Numberofdatarecords		X			
Quality information		X			
Deliveryperiod		X			
Format		X	X	X	X
Deliverymethod		X	X		
Update cycle		X			X
Alternatives			X		

[2] https://dmi.io/, accessed 09.11.2017
[3] https://datamarket.azure.com/browse/data/, accessed 09.11.2017
[4] http://www.datastreamx.com/, accessed 10.11.2017
[5] https://market.mashape.com/, accessed 10.11.2017
[6] https://www.bids4bytes.de/, accessed 10.11.2017

4.2 Development of the M4DG

The following describes the development of the M4DG from a global perspective, so that a holistic picture is in the focus, and not the individual workshops. In order to determine the necessary variations and to map the special requirements in a new metamodel, we carried out a four-stage process. In the first stage, we formed so-called "building blocks" in order to enable a rough classification of the determined properties and to visualize them. The characteristics can be divided into the areas of *organization*, *economy*, *technology* and *compliance*. The building blocks allow a general understanding of the categories of the individual properties and their structuring.

In the second stage, we examined the primary concepts of ADMS. The ADMS divides the properties into 3 concepts, the *repository*, the *asset* and the *distribution*. This classification is also suitable for describing a data source:

- *Repository:* A data source has a storage location. This storage location is marked as a repository. It describes the data catalog of a company or business area.
- *Asset:* The concept of the asset describes the actual data source and its main properties. The characteristics of an asset are primarily organizational and economic properties that enable the management and retrieval of a data source.
- *Distribution:* The distribution of a data source describes the characteristics of data provision. This concept mainly describes the technical characteristics of the data source and is an important source of information for interactions such as data access.

During the third stage, we assigned the identified characteristics to the individual concepts. We merged thematically dependent or related properties into classes that were assigned to the concepts. At this point, even though the models build upon each other, the differences between the models and their specializations become clear. Georgiev et al. show a similar approach with the RAGE metamodel [27].

In the fourth stage, the development ends with the typing of the various attributes in order to achieve the highest possible degree of standardization.

For the description of data goods, we added the missing classes and properties identified from analyses of marketplaces and scientific literature to the ADMS model. Table 2 lists these properties added to the metamodel for data goods with a description.

Table 2. Properties added to the first concept level

Property	*Description*
Curator	In addition to the property of publisher, we propose another for curator information. The curator is the responsible person for the data source.
Subscriber	Applications, processes or individuals that use the data source.
Accessibility	Define access level (e.g. 'public', 'club', 'private') in competitive markets where public availability means less competitiveness.
Owner	For data sources, the publisher must not always be the owner, so we added the property to solve this issue.

Property	Description
Rating	Compared to material goods a recommendation play more and more an important role. Because of this, we added a rating property for the product data as well.
Alternatives	By cataloguing *company's* data sources, the alternative property allows to tag alternative sources for information retrieval.
Origin	For data goods, it is necessary to know the origin and add territorial notices.
DataSetCount	The dataset count describes how many datasets are included in the data source.
UpdateCycle	Specifies the frequency at which the data source is usually updated.
Encryption	Encryption property is used to describe the encryption type of the data source and the used algorithm.
AuthorizationType	Describes whether and if so, which logon method is used.
CompressionType	Describes whether and if so, how the data is compressed.
AccessType	Describes how the data can be retrieved technically.
Sample	Here an example data record can be stored, which allows a first insight into the data source.
Volume	Describes if the volume of the data source is static or dynamic.

In addition to adding properties to existing classes, new classes have also been added to the model, if they serve to describe data as a commodity (see Table 3). Especially economic attributes were added to the M4DG, as they have not received much attention in previous metadata models for data sources.

Table 3. Added classes to represent economic properties

Class	*Property*	*Example*
BillingModel	metric	Pricing model for a data source
BillingModel	price	Pricing information for the data source
BillingModel	discount	Information on discounts that play a role in the purchase of data goods
	payment	The method of paying for the data
LicenceDocument	SLA	Service Level Agreements between the publisher and the user of the data source. Especially in the B2B environment also an interesting factor for data.
LicenceDocument	restriction	Restrictions on the use of data
LicenceDocument	rights	Rights of use for the data
RightsStatement	rightsHolder	Who owns rights to the data
RightsStatement	actions	What can be done with the data
RightsStatement	accessRights	Who can access the data

Like DCAT and ADMS, we also use existing vocabularies, like *dcterms* or *foaf*, to typify the properties of the new classes [31, 32]. This allows a common way of typification and common understanding when transferring the model into an information system for practitioners.

4.3 Resulting Metadata Model

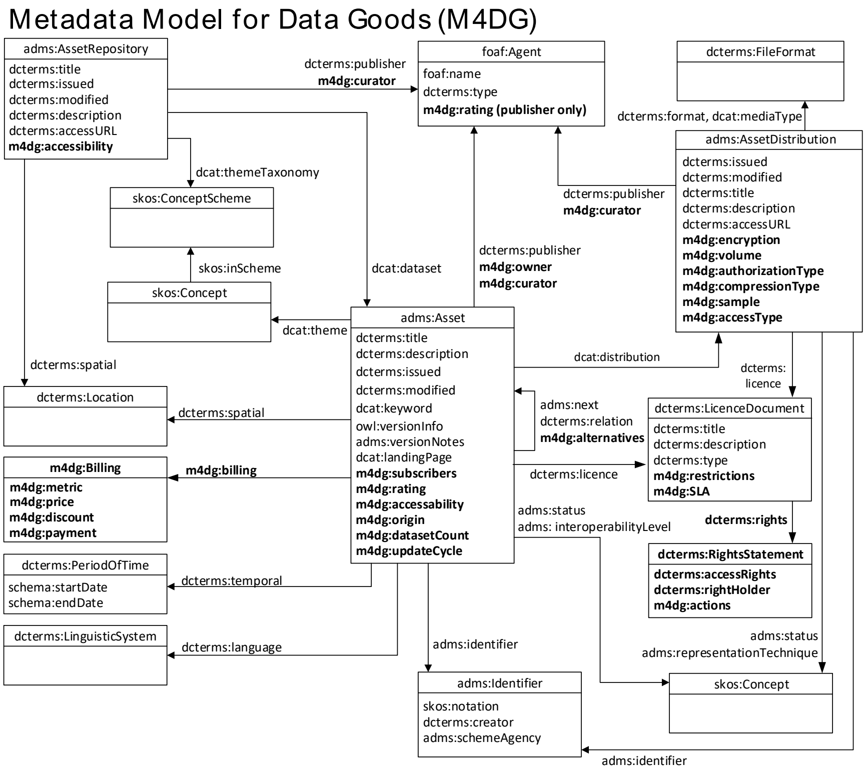


Figure 1. The metadata model for data goods (M4DG)

The result is a metadata model derived from the ADMS, which contains the necessary properties for managing data. The metadata model shown in Figure 1was developed by looking at existing marketplaces, discussions with companies and the analysis of existing literature. We added necessary properties during the design phase and we removed properties that were not related to the context under consideration. Wherever possible, we used existing vocabularies to define and unambiguously typify the properties of the model. By using the 3-levels *repository*, *asset* and *distribution* and the removal and collection of individual properties in classes and themes the data model remains extensible for future versions of the M4DG.

5 Evaluation of the Metadata Model

The goal of the metadata model is to describe data sources, and thus data goods, in a uniform way and consider them as a future resource in companies. For this reason, if data is considered as product, it should be inventoried like any other product. In collaboration with an worldwide active partner from the pharmaceutical industry, who purchases a lot of data from third parties (e.g. patient data or drug studies), we developed an inventory software that enables internal and external data goods to be managed and which supports the mapping of data governance processes, such as assigning responsibilities or evaluating data goods. We developed our software prototype as a web application that enables users to collect existing data sources, search for already registered data goods or display statistics and indicators on inventoried data sources.

The backend of this application uses the M4DG metadata model presented in this paper. Therefore, we implemented M4DG in this project as a mongoose[7] scheme. It enables storing and validating application data in the document-based database MongoDB[8].

Thus we can collect the data sources used by a company in a central repository and catalogue them according to a uniform model. The uniform description and management of the data sources enabled some potentials to be made visible, such as the avoidance of duplicates through the internal search function, the simplified determination of alternative data sources based on entities and simplified control options, including user evaluation. Furthermore, the inventory can be extended with external data sources such as from data marketplaces and industrial platforms by means of a uniform interface.

6 Summary and Outlook

Companies face new challenges in handling and managing data. This report introduces the Metadata Model for Data Goods (M4DG), a metadata model for describing data sources with a focus on using information in an economic context. For this purpose, we examined and compared descriptions of data goods on existing marketplaces on the internet. Subsequently, we identified related metadata models that could serve as the basis for a standardized design. Based on the Asset Description Metadata Schema (ADMS), we developed a model by adding properties and attributes and structuring them using the concept of classes. We examined the model critically from a practical point of view and adapted it to the knowledge gained from practical experience through workshops with companies. Finally, we were able to verify the practicability of the M4DG by using it as a database schema for the inventory of data sources for an international company which purchases a lot of data sources and has to manage all these data sources.

[7] http://mongoosejs.com/, accessed 01.09.2017
[8] https://www.mongodb.com/, accessed 01.09.2017

In the context of our future research, we want to use M4DG as the basis for a standardized description of data sources. We are looking forward to further evaluate the model in cooperation with our current partners and new companies from other industry sectors in order to improve the design and the general validity of the model through further iterations in the sense of the DSR. A first step in this direction could be the extension of the data basis by looking at data marketplaces that have been released in the meantime. Data marketplaces previously ignored due to a lack of availability of information could also be considered again. We are aware that the selection of marketplaces is a weak point and we would like to improve it in the following works. With this contribution we would like to encourage the community to carry out further research in this field.

7 Acknowledgements

This work was partly supported by Fraunhofer Center of Excellence for Logistics and IT.

References

1. Otto, B., Bärenfänger, R., Steinbuß, S.: Digital Business Engineering. Methodological Foundations and First Experiences from the Field (2015)
2. Choi, S.-Y., Stahl, D.O., Whinston, A.B.: The economics of electronic commerce. Macmillan Technical Publishing Indianapolis, IN (1997)
3. Moody, D.L., Walsh, P. (eds.): Measuring the Value Of Information-An Asset Valuation Approach (1999)
4. Otto, B., Jürjens, J., Schon, J., Auer, S., Menz, N., Wenzel, S., Cirullies, J.: Industrial Data Space. Digitale Souveränität über Daten. Whitepaper (2016)
5. Object Management Group, Inc.: Industrial Internet Consortium, http://www.iiconsortium.org/
6. Otto, B.: Quality and value of the data resource in large enterprises. Information Systems Management 32, 234–251 (2015)
7. Hirshleifer, J.: Where are we in the theory of information? The American Economic Review 63, 31–39 (1973)
8. Hilton, R.W.: The determinants of information value. Synthesizing some general results. Management Science 27, 57–64 (1981)
9. Stigler, G.J.: The economics of information. Journal of political economy 69, 213–225 (1961)
10. Repo, A.J.: The value of information. Approaches in economics, accounting, and management science. Journal of the American Society for Information Science (1986-1998) 40, 68 (1989)
11. Muschalle, A., Stahl, F., Löser, A., Vossen, G. (eds.): Pricing approaches for data markets. Springer (2012)
12. Schomm, F., Stahl, F., Vossen, G.: Marketplaces for data. An initial survey. ACM SIGMOD Record 42, 15–26 (2013)

13. Heckman, J.R., Boehmer, E.L., Peters, E.H., Davaloo, M., Kurup, N.G.: A Pricing Model for Data Markets. iConference 2015 Proceedings (2015)
14. Armstrong, A.A., Durfee, E.H. (eds.): Mixing and memory. Emergent cooperation in an information marketplace. IEEE (1998)
15. Otto, B., Auer, S., Cirullies, J., Jürjens, J., Menz, N., Schon, J., Wenzel, S.: Industrial data space. Digital souvereignity over data (2016)
16. Hankel, M., Rexroth, B.: The reference architectural model industrie 4.0 (rami 4.0). ZVEI (2015)
17. Otto, B., Lohmann, S., Auer, S., Brost, G., Cirullies, J., Eitel, A., Ernst, T., Haas, C., Huber, M., Jung, C.: Reference architecture model for the Industrial Data Space. Fraunhofer-Gesellschaft, Munich (2017)
18. Weibel, S., Kunze, J., Lagoze, C., Wolf, M.: Dublin core metadata for resource discovery (1998)
19. International Organization for Standardization: ISO 15836:2009 - Information and documentation -- The Dublin Core metadata element set, https://www.iso.org/standard/52142.html
20. Maali, F., Cyganiak, R., Peristeras, V. (eds.): Enabling Interoperability of Government Data Catalogues. Springer (2010)
21. Alexander, K., Cyganiak, R., Hausenblas, M., Zhao, J. (eds.): Describing Linked Datasets (2009)
22. Tambouris, E., Manouselis, N., Costopoulou, C.: Metadata for digital collections of e-government resources. The Electronic Library 25, 176–192 (2007)
23. Alasem, A.: An overview of e-government metadata standards and initiatives based on Dublin Core. Electronic Journal of e-Government 7, 1–10 (2009)
24. CKAN Association: CKAN instances around the world, https://ckan.org/about/instances/
25. Marienfeld, F., Schieferdecker, I., Lapi, E., Tcholtchev, N. (eds.): Metadata aggregation at govdata.de. An experience report. ACM (2013)
26. Shukair, G., Loutas, N., Peristeras, V., Sklarß, S.: Towards semantically interoperable metadata repositories. The asset description metadata schema. Computers in Industry 64, 10–18 (2013)
27. Georgiev, A., Grigorov, A., Bontchev, B., Boytchev, P., Stefanov, K., Bahreini, K., Nyamsuren, E., van der Vegt, W., Westera, W., Prada, R. (eds.): The RAGE software asset model and metadata model. Springer (2016)
28. Alan, R.H. von, March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS quarterly 28, 75–105 (2004)
29. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action design research. MIS quarterly, 37–56 (2011)
30. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future. Writing a literature review. MIS quarterly, xiii–xxiii (2002)
31. The Dublin Core Metadata Initiative: DCMI Metadata Terms, http://dublincore.org/documents/dcmi-terms/
32. Brickley, D. and Miller, L.: FOAF vocabulary specification 0.91. Technical report, ILRT (2007)

Towards an Understanding of Stakeholders and Dependencies in the EU GDPR

Dominik Huth[1], Anne Faber[1], and Florian Matthes[1]

[1] Technical University of Munich, Department of Informatics,
Garching bei München, Germany
{dominik.huth,anne.faber,matthes}@tum.de

Abstract. Personal data has evolved into an essential element of current business models, which pose new challenges to legislation and organizations. To address these challenges at a European level, the European Commission has passed the General Data Protection Regulation (GDPR). Using a data-driven approach, we identify the key stakeholders that are described in the GDPR, which are the data subject, the controller, the processor, the data protection officer and the supervisory authority. We provide a visual representation of how these entities are interrelated according to the corresponding GDPR articles and determine that companies acting as controllers have the largest need for future actions to achieve compliance with the GDPR.

Keywords: General Data Protection Regulation, GDPR compliance, stakeholders, dependencies, data privacy

1 Introduction

Personal data is the fuel for innovation. Recent technology advances have enabled a range of applications that were previously impossible to conduct. Sensor data from smartphones can be collected at virtually no cost, it can be transmitted via mobile networks, and stored and processed at volumes that were unattainable in prior times.

Since technological innovation is closely related to Business Models [1], the rapid technological developments have motivated substantial changes in Business Models as well. It has enabled a shift from product development towards information aggregation: Facebook creates no own content, Uber does not employ any drivers, German long distance bus company Flixbus does not own any buses, Airbnb does not own any real estate [2].

Some sources suggest that the extensive collection of personal data leads to an Orwellian society, where individuals lose their freedom through surveillance mechanisms [3]. [4], however, argues that not all collection activities interfere with such high values as freedom, but that the problem lies rather within the processing of information, because it changes the balance of power between individuals and institutions.

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Since *free of charge* is an appealing concept to consumers, the market for personal data has failed, just as economic markets can fail [5]. The European Union has recognized the need for unified regulation in this particular market and has passed the General Data Protection Regulation (EU GDPR) to address the challenges of privacy protection for its citizens that are induced by digitization [6]. The GDPR is a regulation that is released at the European Union level and, due to the special instrument as a regulation, is effective May 25, 2018 without further action from members of the Union [7].

The GDPR lays out an extensive set of rules to be followed and can be considered a goal definition in the state and abilities of an institution. We consider it our task in the Information Systems discipline to investigate the "appropriate technical and organizational measures" ([6], Art. 24, 28) to achieve compliance with the GDPR.

As an initial contribution, we regard it as essential to analyze who is involved and affected by the regulation. Thus, we define the research question *"Which are the major stakeholders mentioned in the GDPR and how are they related to each other?"*

The contributions of this paper are:

- A short survey of related work on the GDPR from the Information Systems discipline and in practice in section 2
- An analysis of the key stakeholders in the GDPR and a visualization of their interdependence in section 3

2 Related Work on GDPR Compliance

As the GDPR Regulation was initiated in 2012 and passed in 2016, related work on GDPR compliance is limited to this short timeframe. Work on the topic is investigated from either the perspective of jurisdiction, from information systems or from computer science. For practical advice, we identified guidelines by national supervisory authorities and whitepapers from commercial companies.

This discussion exclusively deals with work that specifically targets the GDPR. Results on compliance with prior regulation (e.g. Directive 95/46/EC, which will be repealed by the GDPR, [6]) will have to be subject of a detailed literature analysis. For the purposes of this work, we performed a literature analysis based on [8]. We searched for the term "EU GDPR" using the academic search engine Scopus in September 2017. The search resulted in 42 papers, with 38 being published in the year 2016 and after. In an analysis of the paper title, the results were reduced to 29, of which 18 were categorized as relevant after reading the abstracts. The available six works are presented in this section.

2.1 Academia

In [7], the implications of the GDPR from a legal perspective are discussed. While the underlying principles for the lawful processing are similar in prior legislation, penalties are increased dramatically with the GDPR. New rules include the transmission to third countries, the accountability of the data controller, and the extended role of supervisory authorities. For organizations, the author derives a need for a data privacy management system, which supports the identification of relevant processes, provides guidelines, and enables control of compliance and discovery of deficits.

[9] points out the importance of personal data for medical research. The accountability principle and data protection by design pose new challenges to medical researchers. Especially the secondary use of research data has to be restricted to usages that are compatible with the initial reasons for collection. Communication of data breaches to data subjects and the supervisory authority requires the establishment of new processes.

[10] discusses the impact of the GDPR on the design and development of smart factories. The authors propose explicit guidelines for technical implementation of consent, the representation of data flows and data expiry as embodiment of the principles of data minimization.

A detailed two-step questionnaire for a data privacy impact assessment is given in [11]. The authors propose to use this artifact to identify risks in projects using cloud technologies. Unfortunately, the presented work lacks an evaluation.

Data compliance and data privacy is stated as one of the key requirements in data management in [12]. A reference model to guide practitioners in designing a data strategy is presented, considering processes & methods, roles & responsibilities, performance management, data architecture, data applications and the data lifecycle. The model addresses the recent developments in the Digital Economy.

[13] presents the architecture of a tool that focuses on the goal of data traceability. It incorporates the definition of customer records as XML files for data portability, central collection and distribution modules and a traceability module that implements an algorithm to discover all entities who received a copy of an individual's data.

The alignment of customer's privacy and security preferences with a service provider's system design is analyzed in [14]. They formalize Privacy Level Agreements (PLA) and develop an extensive metamodel to guide this analysis.

2.2 Practice

The United Kingdom's Information Commissioner's Office (ICO) represents the supervisory authority of the UK. [15] provides 12 steps for companies to take, such as assessing current processing and establishing processes for answering data subject requests or communicating data breaches. Due to the nature of the GDPR, the target state for implementation concurs in all EU countries, but the gap from previous national legislation differs.

The European Union Agency for Network and Information Security (ENISA) published guidelines for SMEs to achieve compliance with the GDPR [16]. The document gives detailed questionnaires for assessing risks and identifying need for action. Related topics of ISO 27001:2013 are given with the assessment questions.

Other practical advice is published by commercial organizations, usually with references to offerings related to achieving compliance with the GDPR. IBM [17] presents five key GDPR duties (rights of EU data subjects, security personal data, lawfulness and consent, accountability of compliance, and data protection by design and by default) and proposes to start with pragmatic steps. Symantec [18] presents survey data that reinforces the need to act and provides four areas of action and corresponding products and implementation partners. Oracle [19] identifies nine core actors and four main areas of work: assess security risks, prevent attacks, monitor and detect breaches, and ensure the quality of protection. These areas are analyzed in the context of database technology.

3 Key Stakeholders in the EU GDPR

3.1 Research Approach

We adopted a two-step research approach analyzing the GDPR documents. First, we coded all stakeholders and relationships. Second, we consolidated the findings and visualized the key stakeholders to be considered for further evaluation in future work.

Following a sequence for a *structuring content analysis* [20], we defined the direction of analysis as the entities or authorities that are mentioned within the 99 GDPR articles [6], as well as statements or prescriptions on interactions between two such entities. We did not cover the means to fulfil these rules yet, because we regard it as essential to establish an understanding of the involved parties first before any other analysis can be conducted. It has to be clear who reports to whom before we can address technical or organizational measures. We defined the coding rule as unambiguous sentences with the structure *subject – predicate – object*, where we applied tags to both entities and the connecting relationship in the full regulation document. This data was extracted and collected in a relational table. We identified 17 unique entities and 64 bilateral relationships, each with their corresponding reference to an Article in the GDPR.

In the second step, we reduced the set of relationships from 64 specific relationships, such as *requests data from* or *consents to collection by* to 33 simple relationships of type *interacts with*. As criteria for a key stakeholder in the GDPR, we defined those entities that have at least three relationships with other entities (active or passive). We graphed these key entities as nodes and relationships as edges and specified the articles that define the corresponding connection.

3.2 Results

Out of the 17 entities we encountered in the document, 12 have a relationship with at most two other entities. The following five entities are involved in at least three relationships:

- *Data subject*: “an identifiable natural person (data subject) is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;" (Art. 4 (1))

- "*Controller* means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data" (Art. 4 (7))
- "*Processor* means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller;" (Art. 4 (8))
- *Data Protection Officer* (Art. 37-39) is designated by the controller or processor and has the tasks to advise and monitor the controller or processor and serve as a contact point for the supervisory authority
- *Supervisory Authority* (Art. 51): an independent public authority that is responsible for monitoring the application of the GDPR

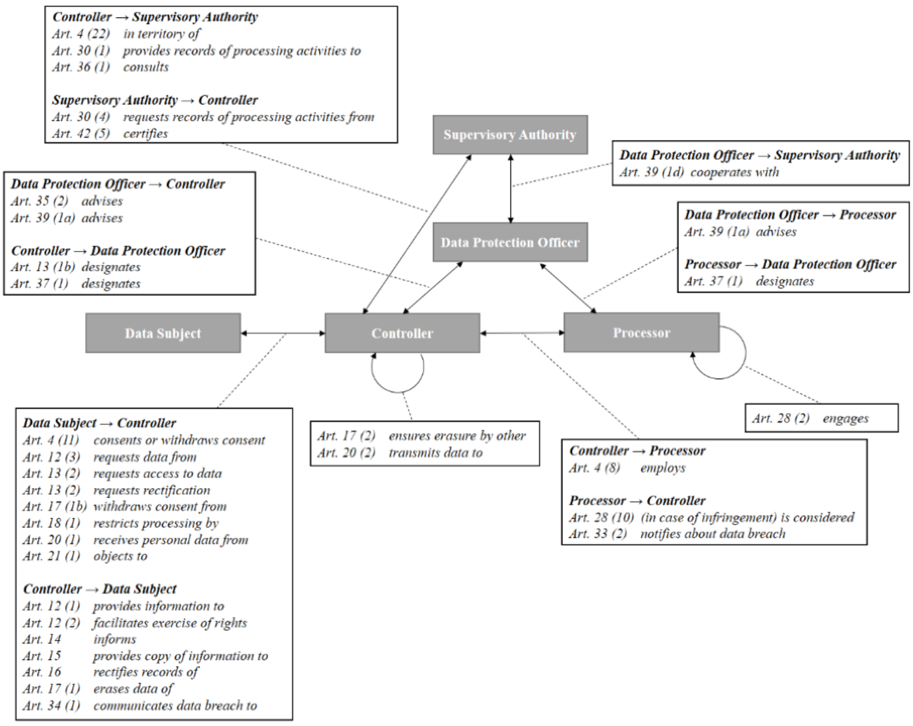


Figure 1: Network of most important stakeholders and their relationships within the GDPR

This list differs from the 9 core actors described by [19], who additionally mention personal data, recipient, enterprise and third party. We do not regard personal data itself as an actor, but rather the subject matter. Recipient, enterprise and third party do refer to entities, but not to a specific role described in the GDPR, i.e. a recipient could be both a controller and a processor.

Figure **1** shows a network of these most relevant stakeholders and their relationships. It becomes evident that the center of activity in this regulation revolves around the data subject and the controller with 15 explicit relationships, as well as between the controller and the supervisory authority with 5 explicit relationships. Thus, we derive that the main actor in the GDPR is the data controller.

This implies a large set of necessary actions. From a technical perspective, systems need to be able to provide options for storing and revoking consent, as well as to restrict processing on a fine-grained level (Art. 4 (1), Art. 21(1)), Art. 18 (1)). The ability to deliver complete and coherent data to data subjects or transfer it to competitors has to be implemented (Art. 20 (2), Art. 20 (1)). The right to data rectification or deletion (Art. 16, Art. 17 (2), and Art. 17 (1)) pose further challenges, especially for tamper-proof systems.

From an organizational perspective, controllers must define processes for the timely communication of data breaches to data subjects and the supervisory authority (Art. 34 (1)). The role of the data protection officer, who reports directly to top management, has to be established (Art. 13 (1b), Art. 37 (1)).

4 Conclusion and Outlook

We have used a structured, data-driven approach to extract the most relevant stakeholders in the GDPR regulation: the data subject, whose personal data is collected and whose rights are reinforced by the GDPR; the controller, who is the single entity that is accountable for lawful data processing; the processor, who is not involved in direct communication with the data subject; the data protection officer as an entity within any processing or controlling entity; and the national supervisory authority. The majority of the articles refer to the interaction between the controller and the data subject and the controller and the supervisory authority.

An understanding of the stakeholders is one element in a complete picture of the GDPR, its implications and possible ways to act upon them. In future work, we will analyze the literature more thoroughly, especially with respect to analogous problems of achieving compliance with prior privacy regulation. Further, we plan to analyze industry efforts of becoming compliant before and after the GDPR enters into force on 25 May 2018. The goal is to identify patterns of dealing with single aspects of the regulation, e.g. the right to transparency, and investigate the effectiveness of the employed methods.

5 Acknowledgements

This work is part of the TUM Living Lab Connected Mobility (TUM LLCM) project and has been funded by the Bavarian Ministry of Economic Affairs and Media, Energy and Technology (StMWi) through the Center Digitisation.Bavaria, an initiative of the Bavarian State Government.

References

1. Baden-Fuller, C., Haefliger, S.: Business Models and Technological Innovation. Long Range

Plann. 46, 419–426 (2013)
2. Goodwin, T.: The Battle Is For The Customer Interface, https://techcrunch.com/2015/03/03/in-the-age-of-disintermediation-the-battle-is-all-for-the-customer-interface/ (Accessed: 29.09.2017)
3. Schneier, B.: Data and Goliath: The hidden battles to collect your data and control your world. WW Norton & Company (2015)
4. Solove, D.J.: I've got nothing to hide and other misunderstandings of privacy. San Diego l. Rev. 44, 745 (2007)
5. Stiglitz, J.E.: The price of inequality: How today's divided society endangers our future. WW Norton & Company (2012)
6. European Union: Regulation 2016/679 of the European parliament and the Council of the European Union (2016)
7. Hamann, C.: Europäische Datenschutz-Grundverordnung - neue Organisationspflichten für Unternehmen. Betriebs-Berater. 1090–1097 (2017)
8. Kitchenham, B.: Procedures for performing systematic reviews. Keele, UK, Keele Univ. 33, 1–26 (2004)
9. Chassang, G.: The impact of the EU general data protection regulation on scientific research. Ecancermedicalscience. 11, 1–12 (2017)
10. Preuveneers, D., Joosen, W., Ilie-Zudor, E.: Data protection compliance regulations and implications for smart factories of the future. Proc. - 12th Int. Conf. Intell. Environ. IE 2016. 40–47 (2016)
11. Alnemr, R., Cayirci, E., Dalla Corte, L., Garaga, A., Leenes, R., Mhungu, R., Pearson, S., Reed, C., de Oliveira, A.S., Stefanatou, D.: A data protection impact assessment methodology for cloud. In: Annual Privacy Forum. pp. 60–92. Springer (2015)
12. Pentek, T., Legner, C., Otto, B.: Towards a Reference Model for Data Management in the Digital Economy. Des. Digit. Transform. DESRIST 2017 Res. Prog. Proc. 12th Int. Conf. Des. Sci. Res. Inf. Syst. Technol. 51–66 (2017)
13. Gjermundrød, H., Dionysiou, I., Costa, K.: privacyTracker: A Privacy-by-Design GDPR-Compliant Framework with Verifiable Data Traceability Controls BT - Current Trends in Web Engineering: ICWE 2016 International Workshops, DUI, TELERISE, SoWeMine, and Liquid Web, Lugano, Switzerland, June 6-9, 201 (2016)
14. Ahmadian, A.S., Jurjens, J.: Supporting model-based privacy analysis by exploiting privacy level agreements. Proc. Int. Conf. Cloud Comput. Technol. Sci. CloudCom. 360–365 (2017)
15. Information Commissioner's Office: Preparing for the General Data Protection Regulation (GDPR) 12 steps to take now. 11 (2016)
16. ENISA: Guidelines for SMEs on the security of personal data processing. (2016)
17. IBM: Planning for the General Data Protection Regulation
18. Symantec: Compliance, the " Privacy By Design " Approach To Protect. (2017)
19. Rajasekharan, D.: Accelerate Your Response to the EU General Data Protection Regulation (GDPR). (2017)
20. Mayring, P., Brunner, E.: Qualitative Inhaltsanalyse. (2009)

Integration von Social-Media-Daten in analytische CRM-Systeme

Milad Mirbabaie[1], Tobias Potthoff[1], Carina Lange[1], and Stefan Stieglitz[1]

[1] Universität Duisburg-Essen, Abteilung Informatik und
Angewandte Kognitionswissenschaft, Duisburg, Deutschland
{milad.mirbabaie,tobias.potthoff,stefan.stieglitz}
@uni-due.de;carina.lange@stud.uni-due.de

Abstract. Im Rahmen des vorliegenden Artikels werden Einsatzmöglichkeiten und Implikationen für die Integration von Social-Media-Daten (SMD) in das Customer Relationship Management (CRM) von Unternehmen untersucht. Der Fokus liegt hierbei auf deutschen Großunternehmen, die erste Schritte zur Integration bereits getätigt haben. Mittels Experteninterviews wird ein empirischer Einblick in das Themenfeld ermöglicht. Die Interviews weisen auf, dass die Integration der SMD in das CRM-System derzeit noch einen geringen Reifegrad hat und verschiedene Barrieren existieren wie das Matchen mehrerer Quellen und der Datenschutz bzw. die Nutzungsbedingungen der Social-Media-Plattformen. Die Unternehmen erkennen aber, dass die SMD einen Mehrwert für die Kundensegmentierung mit sich bringen und die Unternehmen mittelfristig Lösungen für die automatische Integration finden müssen, um wettbewerbsfähig zu bleiben. Als Ziele der Integration von SMD in den CRM-Systeme werden eine bessere Kundensegmentierung, höhere Konversionsraten, aber auch eine kundenorientierte Produktentwicklung genannt.

Keywords: Customer Relationship Management, Social Media, aCRM

1 Einleitung

In der neuen digitalen Weltwirtschaft herrschen immer kürzere Produktlebenszyklen, globale Märkte und eine schwindende Kundenloyalität [1]. Auch der wachsende Konkurrenzdruck weist die Unternehmen darauf hin, dass eine unpersönliche Kundenansprache heute nicht mehr wirksam ist [2]. Daher haben sich in Unternehmen in den letzten Jahrzehnten Customer-Relationship-Management-Systeme (CRM) durchgesetzt, die es ermöglichen, Kundendaten und -beziehungen besser zu verwalten.

Zum anderen haben die letzten Jahre gezeigt, dass Menschen dazu neigen, von der Offline- in die Online-Welt zu wechseln und „omnipräsent, vernetzt und always-on sein“ wollen [3]. Viele Internetnutzer wollen sich im Netz selbst präsentieren, anderen Nutzern dabei zusehen, wie diese sich präsentieren, oder auch miteinander Wissen und Informationen teilen. So bildet sich eine digitale Gemeinschaft, die virtuell neue Inhalte schafft. Oftmals sind die Plattformen, auf denen Nutzer Inhalte veröffentlichen können,

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

in den Basisausführungen entgeltfrei, sodass das bislang knappste und teuerste Gut im Internet ebenfalls ohne Kosten für die Unternehmen produziert wird: der Inhalt.

Unternehmen haben in den letzten Jahren festgestellt, dass insbesondere der Inhalt, der auf den Social-Media-Plattformen erstellt wird, gewinnbringenden Nutzen für die Unternehmen darstellt [4]. Social Media wird genutzt, um Kundenservices zu verbessern und um neue Kunden zu akquirieren. Es wird genutzt, um die Kunden zu motivieren mit ihnen in Kontakt zu treten, um Feedback oder auch verschiedene Wettbewerbe in die Produktentwicklung mit einzubeziehen. Insgesamt zeigte sich bereits in einer Bitkom/KPMG-Studie aus dem Jahr 2016 eine Steigerung der Kundenzufriedenheit und der Kundenbindung zum Unternehmen um 20 Prozent [5].

Daten von Social-Media-Plattformen können bspw. genutzt werden, um Kundenmeinungen über Produkte zu analysieren [6, 7] oder Bedarfe zu identifizieren [8]. In den letzten Jahren hat sich gezeigt, dass Unternehmen immer mehr versuchen, wertvolle Inhalte der Nutzer auf den Social-Media-Plattformen zu gewinnen und sinnvoll in ihre Kundendatenbank zu integrieren [9]. Die Kundendatenbank im CRM beinhaltet nach Möglichkeit alle Offline- als auch Onlinekundendaten. Im CRM werden Daten aus unterschiedlichen Datenquellen gesammelt. Diese werden – typischerweise im analytischen CRM (aCRM) – analysiert und anschließend gezielt anhand der erstellten Kundensegmente für Marketingmaßnahmen genutzt.

Die Aufgabe, die Social-Media-Daten (SMD) in das CRM-System zu integrieren, stellt Unternehmen jedoch vor eine große Herausforderung, da die benötigten Informationen für eine differenzierte Kundenansprache nur unter großen Mühen gewonnen werden können und auch nicht immer rechtzeitig zur Verfügung stehen, da die Kundendaten häufig über eine Vielzahl von Social-Media-Plattformen verteilt liegen. Sowohl für die Praxis als auch für die Wissenschaft ist es dabei notwendig zu erfahren, welche Vor- und Nachteile die Integration der SMD mit sich bringt und was bei der Integration allgemein zu beachten ist. Daher befasst sich das vorliegende Manuskript mit den folgenden Forschungsfragen:

1) *Wie ist der Status quo der Integration von Social-Media-Daten in das analytische CRM bei deutschen Großunternehmen?*
2) *Welche Implikationen gehen für Unternehmen aus der Integration von Social-Media-Daten in das analytische CRM hervor – insbesondere im Hinblick auf die Kundensegmentierung?*

Die Untersuchung dieses Manuskripts soll erstmalig einen empirischen Einblick in die Einsatzmöglichkeiten und Implikationen sowohl für Unternehmen als auch Kunden geben. Dadurch lässt sich ein wissenschaftlicher Beitrag leisten, welche SMD relevant sind, um die Einsatzgebiete und Implikationen zu verbessern, und welche Analysen mit den gegebenen Daten durchgeführt werden können.

2 Forschungsstand

Die veränderte hohe Bereitschaft der Kunden bestehende Geschäftsbeziehungen aufzulösen und Anbieter zu wechseln, zwingt die Unternehmen, dieser Entwicklung entgegenzuwirken. Durch die Abwanderung mancher Kunden, müssen Unternehmen zum einen versuchen, diese zu ersetzen bzw. neue Kunden zu gewinnen. Andererseits bedeuten das Halten existierender Kunden und langanhaltende Kundenbeziehung höhere Profite [10]. Es für Unternehmen schlichtweg kostengünstiger ist, den bestehenden Kundenstamm zu halten, anstatt neue Kunden zu akquirieren. Daher wächst der Druck, den Kunden nachhaltig an das Unternehmen zu binden [1, 11]. Es empfiehlt sich für Unternehmen, vor allem Konzepte wie Kundenmanagement, Beziehungsmarketing und auch die Kundenbindung zu stärken [12].

Die Basis hierfür stellt die kundenbezogene individuelle Analyse dar [13]. Diese ermöglicht eine spezifische Kundenselektion, um so auf ihre individuellen und aktuellen Bedürfnisse eingehen zu können und die Bindung zum Unternehmen aufrechtzuerhalten [2]. Um ein aussagekräftiges, vollumfängliches Kundenprofil zu erstellen, haben Unternehmen die Möglichkeit, aus verschiedenen Quellen verschiedene Daten zu nutzen [14, 15]. Ein Kundenprofil wird hierbei abhängig von Zweck, Inhalt und Situation erstellt. Zudem ist vorab ebenfalls zu beachten, dass eine Kundenprofilerstellung auch stark branchenabhängig ist [2].

Die *Digitale Gesellschaft* wird durch die vielfältigen und weit verbreiteten Nutzungsmöglichkeiten geprägt [16–19]. Social-Media-Anwendungen werden von Kunden intensiv genutzt, um Informationen zu Produkten oder Dienstleistungen zu erhalten [20]. Es wird auf den Social-Media-Plattformen nach den öffentlichen Seiten von Unternehmen gesucht, um sich über verschiedene Produkte oder Dienstleistungen zu informieren. Potenzielle Kunden lesen sich das Feedback zu Produkten, Dienstleistungen und den Kundenservice durch, da die Kunden auf die Empfehlungen anderer Nutzer in sozialen Netzwerken sehr stark vertrauen [21]. Feedback zu Produkten bzw. Dienstleistungen werden dabei oft öffentlich auf der entsprechenden Unternehmensseite platziert oder in Kommentaren unter einem Produkt bzw. einer Dienstleistung geschrieben. Zudem wird gerne der Kundenservice auf den Unternehmensseiten von Kunden beurteilt. Der Kundenservice in Social Media steht dabei unter besonderem Druck, da er meistens öffentlich agiert und lange Antwortzeiten oder inkompetente Antworten für andere Kunden sichtbar werden. Daraus hat sich auch der Kundenanspruch entwickelt, einen schnellen, guten und hilfsbereiten Kundenservice über Social Media vorfinden zu wollen [22, 23].

Durch die Nutzung der Social-Media-Kanäle kommen weitere Datenquellen zum CRM-System hinzu, um das gesamte Kundenprofil abzudecken [24, 25]. Nicht zuletzt ist die immer stärker werdende Akzeptanz und Nutzung von Social-Media-Anwendungen daran beteiligt, dass das Datenaufkommen steigt [26]. Der Stellenwert von Social Media nimmt nicht zuletzt auch aufgrund einer abnehmenden Wirksamkeit klassischer Werbung und PR zu. Social-Media-Marketing kann vor diesem Hintergrund als ein neuer Ansatz betrachtet werden, die Effektivität der Onlinekommunikation zu erhöhen sowie weitere Anknüpfungspunkte für ein erfolgreiches Marketing zu bieten [27].

Um das Kundenerlebnis über alle Marketing- und Vertriebskanäle hinweg zu verbessern, werden somit auch SMD als Datenquellen sinnvoll mit dem bestehenden CRM-System verbunden [9]. Diese Erweiterung durch die Integration verschiedener Social-Media-Kanäle in das CRM stellt das Social CRM dar [28–31]. Social CRM wird von Greve 2011 [21] wie folgt definiert: „*Social CRM ist eine Strategie zur Beteiligung* *des Unternehmens an kollaborativen Konversationen von Kunden und Interessiertes in Social-Media-Kanälen, um für beide Seiten Vorteile aus der kollektiven Wissensgenerierung zu realisieren. Es ist die Antwort des Unternehmens auf die Vorherrschaft des Nutzers von Social Media über die Kommunikationsinhalte, -orte und -zeitpunkte. Ziel ist es, Beziehungen zu Interessierten und Kunden über Social Media zu intensivieren, um daraus für die Kundeninteraktion zu lernen.*“

Zwar liegen Forschungsarbeiten vor, die eine Integration von SMD in die Wertschöpfung vorschlagen. Untersuchungen der konkreten Ansätze von Unternehmen und die verbundenen Mehrwerte aber auch Barrieren sind allerdings kaum zu finden.

3 Methode

Zur Beantwortung der Forschungsfragen wurden zehn halbstrukturierte Experteninterviews und eine qualitative Inhaltsanalyse durchgeführt. Dabei wurden verschiedene Branchen einbezogen, um Unterschiede identifizieren und branchenübergreifende Schlüsse ziehen zu können. Diese Studie konzentrierte sich dabei auf Großunternehmen (GU), die bereits SMD in ihr CRM-System integrieren bzw. gerade dabei sind, diese Integration umzusetzen. Die Recherche der Unternehmen wurde über das Internet sowie über persönliche Kontakte durchgeführt. Die Unternehmen wurden entweder über Kontaktformulare der Unternehmenswebseite oder über Facebook angeschrieben. Darüber hinaus wurden zuständige Ansprechpartner über Xing identifiziert und kontaktiert.

Die Interviewpartner stammten aus acht verschiedenen Unternehmen und unterschiedlichen Branchen (siehe Tabelle 1). Trotz dieser Unterschiede sind alle Interviewpartner Experten für das aCRM oder für den Social-Media-Bereich innerhalb des Unternehmens. Die ähnlichen Funktionen aller Interviewpartner sorgt daher für eine vergleichbare Perspektive.

Tabelle 1. Branche, Umsatz und Mitarbeiterzahl der befragten Unternehmen

Unternehmen	*Branche*	*Umsatz in €*	*Mitarbeiter*
A	Medien/Nachrichten	230 Mio.	700
B	Telekommunikation	7,5 Mrd.	120.000
C	Medien/Verlag	260 Mio.	1.350
D	Automobil	1 Mrd.	5.000
E	Handel	51 Mrd.	328.000
F	Telekommunikation	2 Mrd.	2.600
G	Maschinenbau, Haushaltsgeräte	9 Mrd.	41.500
H	Automobil	22 Mrd.	27.600

Die Interviews wurde zwischen April und Mai 2017 durchgeführt. Insgesamt wurden acht Interviews mit entsprechenden Experten aus den Unternehmen durchgeführt. Zwei Interviews wurden dabei mit zwei Experten gleichzeitig aus demselben Unternehmen durchgeführt, da diese jeweils den Schwerpunkt auf CRM oder Social Media hatten. Drei der Interviewten waren weiblich, sieben männlich. Der Ablauf der Interviews verlief nach dem folgenden Schema: Zunächst bekamen die Experten einen Auszug aus dem Leitfaden. Das diente dazu, dass die Experten sich bereits auf verschiedene Fragen vorbereiten und eventuell einen zweiten Experten aus einem anderen Bereich konsultieren konnten, falls die vorhandene Fachkenntnis nicht ausreichte. Daraufhin wurde das Interview auf Grundlage des Leitfadens entweder mit einem Experten oder maximal mit zwei Experten aus verschiedenen Abteilungen durchgeführt, als Audiodatei aufgezeichnet und anschließend vollständig transkribiert. Aufgrund der Verteilung der Interviewpartner in ganz Deutschland, fanden die Interviews per Telefon statt. Die Länge der Interviews reichte von 13 bis 57 Min.

Der Leitfadenaufbau wurde anhand gesichteter Literatur und der formulierten Forschungsfragen gewählt. Die ersten Fragen bezogen sich allgemein auf Social Media und CRM in dem jeweiligen Unternehmen. Es wurde nach Social-Media-Plattformen gefragt, die vom Unternehmen als Kommunikationskanäle nach Außen angeboten werden. Auch wurde nach typischen Fragestellungen/Analysen in der jeweiligen Branche gefragt. Zudem wurden die Position des Interviewpartners und sein Aufgabenbereich abgefragt.

Im weiteren Verlauf wurden Fragen zu den Einführungsgründen und -zielen der Social-Media-Integration gestellt. Zudem wurden verschiedene Fragen zur Integration aus technischer und analytischer Sicht gestellt. Darunter fielen auch Fragen zu den jeweiligen Social-Media-Plattformen, deren Auswirkungen auf eine veränderte Kundensegmentierung, das Integrationsvorgehen und aufgetretene Konflikte. Im Fokus der Entwicklungsbetrachtung standen die neuen Analysemöglichkeiten im aCRM. Als Abschluss enthielt der Leitfaden Fragen zur Integration der SMD in das CRM-System sowie einen Ausblick auf die zukünftige Verwendung.

Das gewählte Messverfahren dieser Untersuchung ist die qualitative Inhaltsanalyse für teilstandardisierte Interviews nach Mayring [32, 33]. Dabei wurde eine induktive Herangehensweise verfolgt [33, 34]. Bei der induktiven Kategorienbildung werden diese auf Grundlage der Fragestellung und dem theoretischen Hintergrund definiert. Die gebildeten Kategorien legen dann fest, welche Informationen aus den geführten Interviews bei der Textanalyse berücksichtig werden sollen. Nachdem das Material vollständig gesichtet wurde, wurden die angepassten Kategorien revidiert [32, 33].

Insgesamt wurden mithilfe der induktiven Kategorienbildung anhand des vorhandenen Materials sieben Kategorien inklusive Sub-Kategorien entwickelt.

Tabelle 2. Hauptkategorien der Inhaltsanalyse

#	*Kategorien*
1	Einführungsgründe und -ziele
2	Ausgangszustand im analytischen CRM
3	Veränderung im analytischen CRM
4	Integrationsaspekte
5	Steuerungsmechanismen
6	Wahrnehmung und Bedeutung der Social-Media-Daten für das Unternehmen
7	Optimierungsziele

4 Ergebnisse

Anhand der Literatur und des daraus abgeleiteten Interviewleitfadens wurden die Forschungsfragen in den Interviews offen bearbeitet und mit dem Kategoriensystem analysiert. Die Ergebnisse werden im Folgenden anhand vier übergeordneter Aspekte dargestellt, wobei 4.1 insbesondere die erste Forschungsfrage (Status quo) adressiert und anschließend Implikationen aus technischer, analytischer und ökonomischer Perspektive mit Zitatbeispielen dargestellt werden.

4.1 Status Quo und Ziele der Integration

Die Integration der SMD in CRM-Systeme ist bei allen befragten Unternehmen noch in der Anfangsphase. Hier sei die Wichtigkeit von Social Media in den letzten Jahren „sehr stark unterschätzt […worden]. Da ist nun ein Wandel […] zu erkennen“ (E). Die Hälfte der Unternehmen begann ungefähr im Jahr 2015, die anderen vier Unternehmen erst im vergangenen Jahr. Bei nur zwei Unternehmen wurde die Social-Media-Integration Top-Down vom Management vorangetrieben. Bei den anderen sechs Unternehmen war hingegen der Fachbereich maßgeblich treibend.

Die interviewten Unternehmen gaben an, SMD vor allem in Form von Feedbackangaben, Beschwerden oder auch Kommentaren zu den jeweiligen Produkten bzw. Dienstleistungen zu integrieren. Damit soll eine gezieltere Kundenansprache durchgeführt werden können. Drei der acht Unternehmen gaben ebenfalls an, dass mit den SMD die vorgelagerten Analysen verbessert werden sollen, um Kampagnen zu optimieren, Zielgruppen besser zu clustern und letztlich die Kundenbindung zu stärken: „Für mich ist es wichtig, die Daten zu haben, um ein perfektes Kundenerlebnis sicherzustellen“ (H). So könnten Kunden, die eventuell zur Konkurrenz wechseln, frühzeitiger erkannt und Maßnahmen rechtzeitig und gezielt gestartet werden. Nur zwei Unternehmensvertreter sagten, dass die Dienstleistung bzw. das Produkt mit den SMD (z. B. Feedback, Fehlerberichte) verbessert werden kann.

Als Datenquelle sind derzeit verschiedene Social-Media-Kanäle im Einsatz. Facebook und Twitter werden von allen Unternehmen aktiv mit Inhalten des Unternehmens (z. B. über neue Produkte oder Aktionen) versorgt. Aber auch Instagram und YouTube werden von sechs Unternehmen genutzt bzw. beobachtet. Diese werden aber eher genutzt für „das ganz klassische Content-Management. Bezüglich des 1:1-Kundendialogs fokussieren wir uns im ersten Schritt auf Facebook und Twitter.“ (H). Die Analysen im aCRM bei den befragten Unternehmen zielen hauptsächlich auf die Identifizierung von Kundenbedürfnissen ab. Diese umfassen Wahrscheinlichkeiten eines Kaufs, Abwanderungsprognosen oder die Zielgruppenoptimierung, „damit wir keinem Kunden eine Kampagne schicken über Produkte, die sie schon haben oder wo kein Interesse besteht“ (F). Darüber hinaus sollen die Analysen auch zur Optimierung der Dienstleistung bzw. des Produkts dienen: „Probleme an Geräten, Ersatzteilbestellungen, Informationen zu Betriebsdaten“ (G), „wie wird unsere Dienstleistung bewertet im Vergleich zu anderen“ (D), „wie oft kam denn Lob zu welcher Sendung oder zu welchem Thema“ (A). Auch wenn nun die Positivbeispiele herausgegriffen wurden, um den Status Quo zu beschreiben, ist der Grundtenor dennoch zurückhaltend. Vielfach werden die Social-Media-Plattformen nur mit Neuigkeiten versorgt, ohne das Feedback oder weitere SMD systematisch zu erfassen oder in das aCRM zu überführen.

4.2 Integrationsaspekte

Da die Integration von SMD in CRM-Systeme derzeit noch zögerlich betrieben wird, sind Integrationsaspekte von wissenschaftlichem und praktischem Interesse (s. zweite Forschungsfrage). Die Interviews zeigen, dass die meisten Unternehmen ihre Daten aus dem persönlichen Kundendialog in Social Media gewinnen, was oft beim Beschwerdemanagement angesiedelt ist. Zwei Unternehmen integrieren darüber hinaus Daten, die der Kunde z. B. über einen Facebook-Login bereitstellt oder in öffentlichen Kommentaren schreibt. Unternehmen A, das anhand eines Pilotprojekts SMD integriert hatte, interessierte sich zunächst nur für die Kommentare und Feedbacks.

Vier Unternehmen nutzen verschiedene kommerzielle Tools, um SMD zu gewinnen (z. B. BIG CONNECT). Diese Tools haben eigene Zugänge zu Social Media und filtern wichtige Daten heraus. Zudem setzen zwei Unternehmen den Facebook Login ein, mit dem ein Nutzer eine neue Registrierung z. B. für einen Service-Account spart, gleichzeitig aber Zugriff auf einige Facebook-Daten gewährt. Zwei Unternehmen streben eine umfängliche Integration an und nutzen die Programmierschnittstelle von Facebook: „also über die API können Sie eigentlich alles rausziehen, sofern die User das ganze öffentlich posten“ (G). Dennoch fehlt zwei Unternehmen auch noch die Vorstellung über die technische Realisierung einer SMD-Integration.

Die Übertragung der SMD in das CRM-System geschieht letztlich bei der Mehrzahl der Unternehmen händisch. Wenige Unternehmen verwenden dafür alleinstehende oder integrierte Matching-Tools (z. B. Microsoft Social Engagement für MS Dynamics CRM), welche nur eine einmalige Zuordnung verlangen, danach aber automatisch Daten den richtigen Kundenprofilen zuordnen können. Daten aus Kundendialogen in Social Media werden über eindeutige Daten zusammengeführt, wie z. B. E-Mail-Adresse, Kunden- oder Produktnummer.

Problematisch im Rahmen der Integration wird insbesondere der Datenschutz in Deutschland gesehen. „In der Tat ist der große Blocker derzeit das rechtliche Thema “, berichtet Unternehmen F. Durch diese ist es „teilweise etwas schwierig, alles auszuwerten, wie man es eigentlich könnte“ (G). Auch werden nicht alle beobachteten Kanäle übertragen, da die Plattformen unterschiedliche Nutzungsbedingungen haben. Dennoch scheint es möglich, Lösungen für die rechtliche Problematik zu finden („das war relativ hart, bis man das gelöst hatte auf einer sicheren Basis“, G). Zwei Unternehmen nannten fehlende Ressourcen als hemmenden Faktor für die Integration, was letztlich auch an dem Aufwand des händischen Matchings liegt. Generell fällt es zwei der acht Unternehmen sehr schwer, das Matching an sich durchzuführen. Unternehmen H beschreibt das Matching als „die größte Schwierigkeit“. Zudem steht ein Unternehmen vor der Herausforderung, die Plattformen mit unterschiedlichen Bewertungsschemata und Detailtiefe auf ein Kriterium bzw. wenige Kriterien für das CRM-System zusammenzufassen.

4.3 Veränderung im analytischen CRM

Die Integration von SMD hat bei vielen Unternehmen auch Einfluss auf das operative CRM, da hierüber die Kundenkommunikation z. B. bei einem Defekt abgewickelt wird. Hier stellen Social Media in der Regel nur eine Kanalerweiterung dar. In Bezug auf das analytische CRM berichten die befragten Unternehmen von teils starken Veränderungen. Durch die neuen Daten sind „ganz neue Arten von Analysen entstanden […] ganz andere Aspekte von Kundenanalysen und Kundenbetrachtungen in den Mittelpunkt gekommen“ (D). Durch die Verknüpfung verschiedener Informationen wie „[p]ersonenbezogene Informationen, demografische, geografische Informationen [...verbunden mit] Kaufverhalten, Wiederkehrung, Frequenz, auch bestimmte Präferenzen von Kunden“ habe sich viel in der Analytik geändert (D). Konkret wirken sich die SMD positiv auf die Kundensegmentierung aus, welche durch z. B. obengenannte Variablen angereichert werden kann. Ohne die Integration der SMD „würde uns sehr viel eben an Informationen verloren gehen“ (A). Auch Unternehmen, die SMD bisher nicht integrieren, erwarten einen Mehrwert: „Ich glaube, dass wir viel besser unsere Kundensegmente schnüren können, sehr viel genauer sein können, auch in der Kundenansprache und auch nochmal die gegebenen Kundensegmente challengen und gucken, ob wir denn da auch richtig gelegen haben“ (E).

4.4 Erfolg der Integration und dessen Messbarkeit

Die wenn auch anfängliche Integration der SMD wird von den Befragten als große Bereicherung gesehen, auf die nicht verzichtet werden kann: „Wir haben erkannt, da passiert was, was per E-Mail immer mehr zurückgeht oder stagniert […] und wir müssen uns Social Media hinsichtlich des Feedbacks ansehen“ (A). In den Interviews wird auch deutlich, dass es Themen gibt, die nur in Social Media diskutiert werden, welche nun verfolgt werden können. „Die Nutzung und auch die Wichtigkeit haben sich immens nach oben bewegt und […] da sind wir auch noch lange nicht am Ende“ (D). Es werden wenig konkrete Erfolge genannt, vielmehr wird die Social-Media-Integration als Notwendigkeit für die Wettbewerbsfähigkeit angesehen. Social Media „hat sich schon signifikant in Richtung Unternehmenszielerreichung [...] gewandelt“ (F). Gleichzeitig gibt es aber auch Stimmen, die derzeit keinen akuten Nachteil sehen, wenn Social Media noch nicht integriert würden. „Aber langfristig verlieren wir natürlich dadurch einen wichtigen Kanal“ (H). Informationen über Social Media werden zwar als hochwertig und ergänzend, das Kosten-Nutzen-Verhältnis jedoch teilweise als unzureichend bewertet. Dennoch bemisst ein Unternehmen den Erfolg der Integration daran, dass „Kunden, die man [...] vorher – aus ökonomischer Perspektive – als interessant betrachtet hat, [...] später weniger Interessant geworden sind“ (D).

Feste Kennzahlen, die den Wert der Integration von SMD quantitativ bemessen (RoI), sind noch nicht etabliert. Teilweise werden steigende Konversionsraten und sinkende Marketingkosten genannt. Andererseits wurde der Sinn von RoI-Berechnungen von einem Unternehmen infrage gestellt, „weil das aus unserer Sicht eigentlich eine riesen Zeitverschwendung ist“ (G). Das Management wisse eigentlich um den Wert der Daten und deren Bedeutung in der Digitalisierung, weshalb „die da gar nicht so genau fragen, was [...] der Return on Investment“ ist (B). Eine andere Perspektive auf den Erfolg ist die erfüllte Erwartungserhaltung der Kunden über Social Media mit Unternehmen kommunizieren zu können: „Wir können diesen Kanal nicht ignorieren, weil da der Großteil unserer Kunden sich bewegt und weil die Erwartungshaltung der Kunden ist, dass man dort auch präsent ist“ (C).

5 Diskussion

Die Interviews haben ergeben, dass Unternehmen weiterhin eine umfassende Integration von SMD in ihre CRM-Systeme scheuen, was insbesondere auf Datenschutzbedenken zurückzuführen ist. Ein aufschlussreicher Hinweis auf die einfachere Handhabung in den USA (Unternehmen G) sollte zu denken geben, ob zumindest global das deutsche bzw. europäische Datenschutzgesetz nicht auch einen Wettbewerbsnachteil bedeuten kann. Zwar wurde erwähnt, dass sich die Unternehmen nicht besonders unter Druck gesetzt fühlen, SMD sofort zu integrieren. Gleichzeitig sehen sie in der langfristigen Entwicklung, der weiteren Durchdringung der sozialen Medien und der damit verbundenen Abnahme konventioneller Marketingkanäle eine gewisse Bedrohung und Handlungspflicht.

Übergeordnetes Ziel der SMD-Integration ist die Ergänzung relevanter Informationen in Kundenprofilen. Das betrifft sowohl die Historie des und Interaktion mit dem Kundenkontakt, da vermehrt der Kundenservice über Social Media abgewickelt wird und Kunden dies auch einfordern, als auch die Anreicherung mit bspw. demographischen Daten oder allgemeinen Interessen des (potenziellen) Kunden. Die Integration der Kundeninteraktionen in Social Media steht im Grunde nicht zur Disposition [22] und stellt datenschutzrechtlich auch keine Herausforderung dar, wenn der Kunde den Kanal und die preisgegebenen Daten selbst wählt. Hier äußern sich die Unternehmen eindeutig, dass bei solchen Kundeninteraktionen eine automatisierte Integration von SMD in die CRM-Systeme nötig ist. Die Integration dieser Daten bedeutet dann keine neuen Herausforderungen für die aCRM-Analysen, da die Informationen in der Regel dieselben sind wie über konventionelle Kanäle. Schwierigkeiten bereitet allerdings die automatisierte Integration. Bislang werden teilweise noch manuelle Übertragungen in das CRM-System vorgenommen.

Anders sieht es bei SMD aus, die nicht direkt aus einer Kundeninteraktion hervorgehen, sondern durch Beobachten und Verfolgen von öffentlichen Social Media herausgefunden oder mitgelesen werden. Diese Informationen stellen laut den befragten Unternehmen eine gänzlich neue Informationsquelle dar, die über bisherige Ansätze hinausgeht und Analysen sowie Marketingkampagnen signifikant beeinflusst. Bei einem Unternehmen zeigte sich, dass ein Kundensegment bei Berücksichtigung weiterer Informationen aus SMD als weniger profitabel bewertet wurde. Hier haben die neuen Daten folglich im direkten Vergleich zu einer Neubewertung geführt, die das Kundensegment niedriger priorisiert. Kundensegmente sollten möglichst homogen sortiert werden. Unternehmen profitieren von der Optimierung durch SMD, da gezieltere Marketingmaßnahmen zu höherem Erfolg bzw. Konversionen führen und Budget gespart werden kann [9]. Für die Unternehmen scheint aber ebenfalls das

Feedback der Kunden wichtig zu sein. Zum einen zur Produktentwicklung, zum anderen aber auch, um bei Beschwerden Kundenbindungsprogramme zu initiieren [24]. Die Befragten sehen daher in den sozialen Netzwerken eine Informationsquelle, die es bislang nicht so gab. Bei dieser vom Benutzer oder Kunden nicht zugestimmten Verwendung der SMD kommen die Unternehmen schnell in rechtliche Bedrängnis. Es bleibt nach wie vor Aufgabe der Forschung und der Praxis und der Plattformbetreiber, Wege zu finden, die allen Beteiligten einen Nutzen bieten. Vielfach wünschen sich Kunden eine bessere Ansprache und gezieltere Werbekampagnen, insbesondere wenn die Alternative unpassende und störende Werbemaßnahmen bedeutet.

Die Bandbreite an Social-Media-Plattformen und -typen stellt die Unternehmen vor die Herausforderung, die richtigen auszuwählen. Es ist verständlich, dass die Unternehmen, die die SMD-Integration erst planen, zunächst an eine möglichst vollständige Integration aller Plattformen denken. Hier zeigt aber die Erfahrung der anderen Unternehmen, dass es zum einen komplex ist, eine Vielzahl zu integrieren und Informationen zu matchen, manche Netzwerke aber auch keine oder kaum relevante Information liefern (z. B. LinkedIn). In der Praxis scheinen insbesondere Daten aus dem Netzwerk Facebook interessant zu sein, was aber letztlich auch an der weiten Verbreitung liegen mag und sich mittelfristig ändern kann.

Neben den datenschutzrechtlichen Bedenken existieren auch immer noch technische Herausforderungen auf Informationen zuzugreifen bzw. diese automatisiert zusammenzufügen (matching). Es wird deutlich, dass verschiedene Tools bzw. Erweiterungen zu CRM-Systemen verwendet werden. Eigenständige Tools wie BIG CONNECT werden dabei als sehr leistungsfähig bewertet, lassen sich oft aber nicht gut in bestehende Systeme integrieren.

6 Fazit

Die Integration von Social-Media-Daten in CRM-Systeme wurde in der Vergangenheit mit Vorteilen wie einer besseren Kundensegmentierung, geringeren Marketingkosten verbunden und als Datenquelle für neue Produkt- oder Dienstleistungsentwicklungen gesehen. Dennoch scheuen selbst Großkonzerne die vollumfängliche Einbeziehung von Social Media. In dieser Studie wurden acht deutsche Unternehmen über den Stand ihrer SMD-Integration befragt, um eine detaillierte Bestandsaufnahme zu erstellen. Es zeigt sich, dass das Datenschutzrecht eine umfänglichere Nutzung hemmt, aber auch legale Wege gefunden wurden, die Daten aus sozialen Netzwerken zu nutzen. Die Unternehmen beschränken sich dabei auf wenige Plattformen wie Facebook oder Twitter, wenngleich sie auf weiteren im externen Content-Marketing aktiv sind. Professionelle Tools unterstützen die Kundeninteraktion über Social Media schon weitestgehend, dennoch geschieht die Überführung in das CRM-System oft noch manuell, was bei einem Unternehmen auch zu einem zu schlechten Kosten-Nutzen-Verhältnis und einem Stopp des Pilotprojekts geführt hat.

Der Artikel trägt zum weiteren Verständnis der zögerlichen Integration von SMD bei. Unternehmen können ihren eigenen Reifegrad mit den hier untersuchten Unternehmen vergleichen. Gleichzeitig zeigt die Studie, dass weitere Forschung notwendig ist, welche gerade die Wirtschaftsinformatik anspricht, da sowohl rechtliche, ökonomische als auch technische Aspekte eine Rolle spielen.

Die befragten Unternehmen sind nicht repräsentativ für die deutsche Wirtschaft und lassen daher keine Verallgemeinerung zu. Der Fokus lag auf Großunternehmen, da von diesen ein höherer Reifegrad der SMD-Integration angenommen wurde. Die konkrete Auswahl der Interviewpartner konnte aufgrund der zurückhaltenden Bereitschaft nicht gänzlich systematisch erfolgen. Dennoch erlauben die reichhaltigen qualitativen Daten, Ziele, Ansätze oder auch Hindernisse besser zu verstehen. Weitere Forschung in diesem Feld sollte ausgewählte Aspekte weiter fokussieren (z. B. Lösung der Datenschutzthematik) oder auch quantitativ eine größere Menge an Unternehmen zu den hervorgehobenen Herausforderungen befragen.

Literaturverzeichnis

1. Brezina, R.: Analytisches Customer Relationship Management. Controlling. 13, 219–226 (2001).
2. Neckel, P., Knobloch, B.: Customer Relationship Analytics - Praktische Anwendung des Data Mining im CRM. dpunkt.verlag GmbH (2015).
3. Hamidian, K., Kraijo, C.: DigITalisierung -- Status quo. In: Keuper, F., Hamidian, K., Verwaayen, E., Kalinowski, T., and Kraijo, C. (eds.) Digitalisierung und Innovation: Planung - Entstehung - Entwicklungsperspektiven. pp. 1–23. Springer, Wiesbaden (2013).
4. Hoffman, L.D., Fodor, M.: Can You Measure the ROI of Your Social Media Marketing? MIT Sloan Manag. Rev. 52, 41–49 (2010).
5. Pols, A., Heidkamp, P., Erwin, T.: Pressekonferenz: Mit Daten Werte Schaffen 2016.
6. Ding, X., Street, S.M., Liu, B., Street, S.M., Yu, P.S., Street, S.M.: A Holistic Lexicon-Based Approach to Opinion Mining. In: WSDM'08. pp. 231–239 (2008).
7. Hu, M., Liu, B.: Mining Opinion Features in Customer Reviews. In: American Association for Artificial Intelligence. pp. 755–760 (2004).
8. Kuehl, N., Scheurenbrand, J., Satzger, G.: NEEDMINING : IDENTIFYING MICRO BLOG DATA CONTAINING CUSTOMER. In: ECIS (2016).
9. Rosemann, M., Eggert, M., Voigt, M., Beverungen, D.: LEVERAGING SOCIAL NETWORK DATA FOR ANALYTICAL CRM STRATEGIES - THE INTRODUCTION OF SOCIAL BI. In: ECIS (2012).
10. Stauss, B.: Rückgewinnungsmanagement: Verlorene Kunden als Zielgruppe. In: Bruhn, M. and Stauss, B. (eds.) Dienstleistungsmanagement Jahrbuch 2000: Kundenbeziehungen im Dienstleistungsbereich. pp. 449–471. Gabler Verlag, Wiesbaden (2000).
11. Meyer, M.: CRM-Systeme mit EAI: Konzeption, Implementierung und Evaluation. Vieweg+Teubner Verlag (2002).
12. Mullins, R.R., Ahearne, M., Lam, S.K., Hall, Z.R., Boichuk, J.P.: Know Your Customer: How Salesperson Perceptions of Customer Relationship Quality Form and Influence Account Profitability. J. Mark. 78, 38–58 (2014).
13. Miranda, M.D., Sassi, R.J.: Using Sentiment Analysis to Assess Customer Satisfaction in an Online Job Search Company. In: Abramowicz, W. and Kokkinaki, A. (eds.) BIS 2014 International Workshops, Larnaca, Cyprus, May 22-23, 2014, pp. 17–27.

14. Chang, Y.-C., Ku, C.-H., Chen, C.-H.: Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. Int. J. Inf. Manage. (2017).
15. Stieglitz, S., Dang-Xuan, L., Bruns, A., Neuberger, C.: Social media analytics: An interdisciplinary approach and its implications for information systems. Bus. Inf. Syst. Eng. 6, 89–96 (2014).
16. Gabriel, R., Röhrs, H.-P.: Social Media - Potenziale, Trends, Chancen und Risiken. Gabler Verlag (2017).
17. Stieglitz, S., Bunker, D., Mirbabaie, M., Ehnis, C.: Sense-Making in Social Media During Extreme Events. J. Contingencies Cris. Manag. 1–17 (2017).
18. Mirbabaie, M., Zapatka, E.: Sensemaking in Social Media Crisis Communication - A Case Study on the Brussels Bombings in 2016. In: 25th ECIS. pp. 2169–2186 (2017).
19. Stieglitz, S., Mirbabaie, M., Schwenner, L., Marx, J., Lehr, J., Brünker, F.: Sensemaking and Communication Roles in Social Media Crisis Communication. In: Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen. pp. 1333–1347 (2017).
20. Whiting, A., Williams, D.: Why people use social media: a uses and gratifications approach. Qual. Mark. Res. An Int. J. 16, 362–369 (2013).
21. Greve, G.: Social CRM -- Zielgruppenorientiertes Kundenmanagement mit Social Media. In: Bauer, C., Greve, G., and Hopf, G. (eds.) Online Targeting und Controlling: Grundlagen -- Anwendungsfelder -- Praxisbeispiele. pp. 261–285. Gabler, Wiesbaden (2011).
22. Ashley, C., Tuten, T.: Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. Psychol. Mark. 32, 15–27 (2015).
23. Fan, W., Gordon, M.D.: The Power of Social Media Analytics. Commun. ACM. 57, 74–81 (2014).
24. Ang, L.: Community relationship management and social media. J. Database Mark. Cust. Strateg. Manag. 18, 31–38 (2011).
25. Baird, C.H., Parasnis, G., Baird, C.H., Parasnis, G.: From social media to social customer relationship management. (2015).
26. Rudolph, T., Linzmajer, M.: Big Data im Handel. Mark. Rev. St. Gall. 31, 12–25 (2014).
27. Hettler, U.: Social Media Marketing. Marketing mit Blogs, Sozialen Netzwerken und weiteren Anwendungen des Web 2.0. De Gruyter Oldenbourg (2010).
28. Trainor, K.J., Andzulis, J. (Mick), Rapp, A., Agnihotri, R.: Social media technology usage and customer relationship performance. J. Bus. Res. 67, 1201–1208 (2014).
29. Rodriguez, M., Peterson, R.M., Ajjan, H.: Crm/Social Media Technology: Impact on Customer Orientation Process and Organizational Sales Performance. In: Kubacki, K. (ed.) Ideas in Marketing: Finding the New and Polishing the Old: Proceedings of the 2013 AMS Annual Conference. pp. 636–638. Springer International Publishing, Cham (2015).
30. Malthouse, E.C., Haenlein, M., Skiera, B., Wege, E., Zhang, M.: Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House. J. Interact. Mark. 27, 270–280 (2013).
31. Choudhary, A., Hendrix, W., Lee, K., Palsetia, D., Liao, W.-K.: Social media evolution of the Egyptian revolution. Commun. ACM. 55, 74 (2012).
32. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. (2008).
33. Mayring, P.: Qualitative Inhaltsanalyse. Forum Qual. Sozialforsch. 1, 1–10 (2000).
34. Gläser, J., Laudel, G.: Experteninterviews und qualitative Inhaltsanalyse. VS Verlag für Sozialwissenschaften, Wiesbaden (2010).

Entwicklung eines Beschreibungsschemas für Workflow Privacy Patterns

Alexandra Janssen[1], Franziska Plate[1], Dominik Schneider[1], Erik Buchmann[2] und Jürgen Anke[2]

[1] Detecon International GmbH, Sternengasse 14-16, 50676 Köln
{Alexandra.Janssen, Franziska.Plate, Dominik.Schneider}@detecon.com

[2] Hochschule für Telekommunikation Leipzig, Gustav-Freytag-Straße 43-45, 04277 Leipzig
{buchmann, anke}@hft-leipzig.de

Abstract. Aufgrund der Digitalisierung von Geschäftsprozessen werden immer mehr personenbezogene Daten von computergestützten Informationssystemen erfasst, gespeichert und verarbeitet. Deshalb steigen die Anforderungen an Unternehmen bezüglich der Umsetzung von Datenschutzregelungen. Workflow Privacy Patterns (WPP) haben das Potenzial, Unternehmen bei dieser Aufgabe zu unterstützen. WPPs sind abstrakte, validierte Prozessmuster, die Prozessentwicklern, Datenschutzbeauftragten oder Auditoren bei der Modellierung, Implementierung oder Prüfung von Prozessen in Hinblick auf wiederkehrende Datenschutzprobleme helfen sollen. In diesem Beitrag wird ein Beschreibungsschema für WPPs vorgeschlagen, das als Grundlage für WPP-Modellierungen und zum Vergleich von WPPs genutzt werden kann. Weiterhin wird die Anwendbarkeit des Beschreibungsschemas anhand eines konkreten WPPs veranschaulicht, diskutiert und bewertet.

Keywords: Workflow Privacy Pattern, Beschreibungsschema, Datenschutz in Geschäftsprozessen, Prozessmodellierung, EU-DSGVO.

1 Einleitung

Für große Unternehmen wird es immer aufwändiger, die Anforderungen von Datenschutzgesetzen wie dem Bundesdatenschutzgesetz (BDSG) [1], der neuen EU-Datenschutzgrundverordnung (EU-DSGVO) [2] oder unternehmensbezogenen Compliance-Vorschriften zu erfüllen. Das liegt unter anderem daran, dass rasante Entwicklungen im Bereich der Informations- und Kommunikationstechnologien neue Möglichkeiten zur Digitalisierung von Geschäftsprozessen bieten [3].

Unternehmen, die personenbezogene Daten elektronisch erfassen, speichern und verarbeiten, müssen sicherstellen, dass die Vorgaben aus Datenschutzgesetzen und Datenschutzrichtlinien in Geschäftsprozessen eingehalten werden. Dafür entwickelte Werkzeuge wie Privacy Management-Systeme [4] helfen Unternehmen bei der Überprüfung und Realisierung des Datenschutzes bzw. der Datenschutzrichtlinien [5]. Sie setzen jedoch zumeist erst nach der Umsetzung von Datenschutz-kritischen

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Anwendungen oder Prozessen an. Die Entwicklung von Geschäftsprozessen unter Berücksichtigung des Datenschutzes ist weiterhin mit viel manueller Arbeit für Datenschutzbeauftragte, Prozessentwickler oder Auditoren verbunden.

Unser zentrales Anliegen besteht darin, dass wir den Prozessentwickler dabei unterstützen wollen, bestimmte Datenschutzkonzepte in Geschäftsprozessen einfach umzusetzen und den Analysten dazu befähigen wollen, existierende Geschäftsprozesse auf die Umsetzung dieser Datenschutzkonzepte zu prüfen. In einer Vorarbeit haben wir dazu Workflow Privacy Patterns (WPP) [6] vorgestellt. Analog zu den Design Patterns im Software Engineering sollen WPPs die Modellierung von wiederkehrenden Datenschutzproblemen in Form von standardisierten, validierbaren Prozessmustern erlauben. Ein Beispiel ist das Pattern „Anonyme Dienstnutzung“ (Abbildung 1). Hierbei kann der Aufruf eines bestimmten Services (z.B. Startseite eines Nachrichtenportals oder Hotelbuchungssystems) entweder personalisiert oder anonym angezeigt werden. Im personalisierten Fall werden personenbezogene Daten erhoben und für die Vorbereitung der Diensterbringung verwendet.

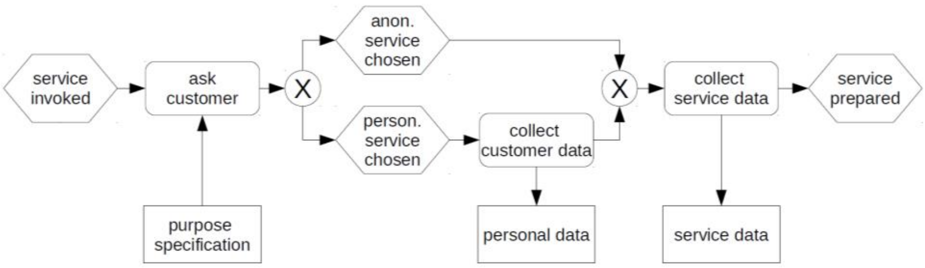


Abbildung 1: WPP für die optional anonyme Nutzung eines Service [6]

Das WPP in Abbildung 1 zeigt, dass für eine datenschutzkonforme Modellierung die Reihenfolge eine wesentliche Rolle spielt, z.B. die Bekanntgabe des Erhebungszwecks vor der Entscheidung für oder gegen eine personalisierte Dienstnutzung oder die Erfassung von personenbezogene Daten erst nach der Entscheidung für eine personalisierte Nutzung. Weitere offensichtlich wichtige Modellbestandteile sind das Vorhandensein von speziellen Datenobjekten, die gesetzliche Eigenschaften erfüllen und an ganz bestimmten Stellen im Prozess verankert sein müssen, wie z. B. eine Datenschutzerklärung. Damit ein WPP Datenschutzkonzepte aus Gesetzestexten, Datenschutz-Richtlinien oder Compliance-Vorschriften in Geschäftsprozesse umsetzen kann, darf es darum nicht nur Kontroll- oder Datenflüsse abbilden, sondern muss domänenspezifische Aspekte des Datenschutzes berücksichtigen.

In diesem Beitrag wird ein Beschreibungsschema entwickelt, das alle erforderlichen Attribute für die Abbildung von Datenschutzkonzepten auf WPPs enthält. Dies ist eine herausfordernde Aufgabe: WPPs müssen für unterschiedliche Kategorien von Datenschutzkonzepten einsetzbar sein, wie herkömmliche Workflow Pattern auch generische Eigenschaften wie Wiederverwendbarkeit oder Verständlichkeit erfüllen, und darüber hinaus domänenspezifische Eigenschaften des Datenschutzes abbilden können. Wir haben daher ausgehend vom aktuellen Stand der Forschung grundsätzliche

Eigenschaften für WPPs zusammengestellt und auf der Basis von Gesetzestexten Attribute aus der Datenschutzdomäne identifiziert. Daraus ergibt sich ein umfassendes Bild über Eigenschaften und Anforderungen von WPPs, die sich mittels Kategorisierung in ein Beschreibungsschema überführen lassen.

Der Beitrag gliedert sich in fünf Abschnitte. Im nächsten Abschnitt wird der aktuelle Stand der Forschung beschrieben. In Abschnitt 3 werden die Workflow Pattern (WP)-Kategorien als Grundlage für das Beschreibungsschema erläutert. Auf Basis der WP-Kategorien und weiterer Forschung wird in Abschnitt 4 ein Beschreibungsschema vorgeschlagen. Der Beitrag schließt mit einem Fazit und Ausblick.

2 Stand der Forschung

Nachfolgend werden verwandte Arbeiten aus vier Bereichen vorgestellt: (1) Die Umsetzung von Datenschutzkonzepten auf der Ebene der Geschäftsprozess-modellierung, (2) Vorarbeiten zu Workflow Patterns, (3) Vorarbeiten zur Spezifikation von Workflow Privacy Patterns sowie (4) allgemeine Ansätze zur Ableitung von domänenspezifischen Workflow Patterns.

2.1 Berücksichtigung von Datenschutz in Geschäftsprozessen

Datenschutzaspekte in Geschäftsprozessen werden bereits seit mehr als zehn Jahren in der Literatur diskutiert. So schlägt Lange ein Vorgehensmodell für die Gestaltung vertrauenswürdiger Informationssysteme vor, das sowohl rechtliche als auch unternehmerische Ziele berücksichtigt. Dabei wird der Zusammenhang zwischen Daten-, Regel- und Akteursmodell konzeptionell formuliert [7]. Zur Datenschutzbewertung in kleinen und mittleren Unternehmen stellen Rodeck et al. einen pattern-basierten Ansatz vor, mit dem häufige Prozessmuster als BPMN (Business Process Model and Notation)-Modell gespeichert werden. Diese werden zur Identifikation von Schwachstellen sowie deren Dokumentation in einem Tool genutzt [8].

Ein Ansatz von Karjoth widmet sich der Frage von maschinenlesbaren Datenschutzrichtlinien zur Sicherstellung von Datenschutzkonformität in Geschäftsprozessen. Dazu werden Prozessmodelle mit datenschutzrelevanten Aspekten annotiert. Zentral ist hier die Verknüpfung von Zwecken und Aufgaben („Tasks“). Durch die formale Beschreibung von Datenschutzrichtlinien können sie automatisch mit den Eigenschaften des Prozesses verglichen werden [9]. Ein weiteres Konzept für die Formulierung und Durchsetzung von Datenschutzanforderungen sind „purpose-aware policies“ [10]. Diese beschreiben formal den Zusammenhang von Zwecken, verantwortlichen Stellen, Daten sowie den zulässigen Aktionen auf diesen. Der Vorteil der formalen Beschreibung liegt in der Prüfbarkeit der Einhaltung (bei Verfügbarkeit entsprechender Daten).

2.2 Workflow Patterns

Patterns beschreiben generische Lösungen für teilweise wiederkehrende, nichttriviale Probleme [11], die in einem spezifischen Entwicklungskontext entstehen [12]. Bei Workflow Patterns bilden die betrieblichen Abläufe [13] diesen Kontext.

Allgemein werden Patterns beschrieben durch Bedingungen für die Anwendbarkeit des Patterns, konkrete Anwendungsszenarien, ggf. Probleme bei der Instanziierung des Patterns in existierenden Modellierungssprachen sowie die implementierbaren Lösungen selbst [11]. Aufgrund der verschiedenen Anwendungsbereiche werden WPs in vier Perspektiven eingeteilt. So sind nach [14] die Perspektiven Daten, Ressourcen, Kontrollfluss und Exception-Handling (Behandlung von Fehlerzuständen) definiert.

Kontrollfluss-Patterns beschreiben alle Aspekte und Abhängigkeiten von Flussobjekten. Flussobjekte bezeichnen in der BPMN-Notation Aktivitäten eines Prozesses bzw. Prozessschritte. Ein Beispiel für diese Kategorie ist ein WP, das regelt, wie eine Parallelisierung von Prozessschritten sowie deren Synchronisation modelliert werden. Daten-Patterns fokussieren sich auf die Beschreibung von Informationsaustauschen und den dazugehörigen Daten, die in Geschäftsprozessen vorkommen [15]. Daten treten in Geschäftsprozessen auf unterschiedliche Weise auf. Sie können unter anderem den Kontrollfluss verändern, von Komponenten im Prozessablauf verarbeitet werden oder einen Datenaustausch zwischen Komponenten ermöglichen [14]. Mithilfe von Ressourcen-Patterns kann die Aufgabenverteilung in Geschäftsprozessen geregelt werden [15]. Dazu werden Ressourcen-Objekte, z. B. Arbeitskräfte oder Organisationseinheiten, zu Prozess-Komponenten zugeordnet. Exception-Handling-Patterns basieren auf den drei anderen Perspektiven, während Kontrollfluss-, Daten- und Ressourcen-Patterns kaum zusammenhängen [15]. Exception-Handling-Patterns definieren, wie verschiedene Ausnahmesituationen in den einzelnen Perspektiven gehandhabt werden und welche Maßnahmen darauf erfolgen müssen. Sie geben zum Beispiel Auskunft darüber, welche Maßnahmen nach einem Ausnahmefall eingeleitet werden müssen, um diesen zu beheben [15].

2.3 Workflow Privacy Patterns

Workflow Privacy Patterns [6], wie das in Abbildung 1 dargestellte, spezifizieren eine generische, wiederverwendbare Lösung zu einem mehrfach auftretenden Problem in der Datenschutz-Domäne in Form eines Workflow Patterns.

WPPs lassen sich in drei Kategorien einteilen, die sich nach der Art und Weise der Datenschutz-Umsetzung in Geschäftsprozessen unterscheiden. „Privacy Patterns" beschreiben Bausteine für Datenschutz-Prozesse, z.B. die Implementierung eines Auskunftsersuchens [6]. „Crosscutting Privacy Patterns" definieren Patterns, die innerhalb der Geschäftsprozesse eines Unternehmens übergreifend Anwendung finden, beispielsweise um zu validieren, ob eine bestimmte Datennutzung vom Betroffenen autorisiert wurde. „Meta Privacy Patterns" bilden übergeordnete Konzepte wie „Separation of Duties" oder „Separation of Concerns" auf Patterns ab.

WPPs finden bei Prozessmodellierern nur dann Akzeptanz, wenn sie einen klaren Vorteil gegenüber der manuellen Prozessmodellierung bieten. In [6] werden dafür drei wesentliche Anforderungen an WPPs definiert:

A1: *WPPs sollen keine Modifizierungen der Modellierungssprache erfordern.*
A2: *WPPs sollen für Prozessmodellierung als auch für Implementierung nützlich sein.*
A3: *WPPs sollen separat vom eigentlichen Geschäftsprozess entwickelt und gepflegt werden können.*

2.4 Ableitung von domänenspezifischen Patterns

Das Erstellen von WPs für konkrete Anwendungsdomänen ist schwierig, da an Patterns hohe Erwartungen gestellt werden: Sie sollen nützlich, verständlich, wiederverwendbar, in unterschiedliche Sprachen transformierbar, verifizierbar usw. sein.

Ein Bottom-Up-Ansatz [16] zur Entwicklung von Patterns besteht darin, dass aus einer Grundgesamtheit von Workflow-Instanzen jeweils wiederkehrende Muster in einem manuellen Prozess herausextrahiert werden, beginnend bei einfachen Patterns. Die Korrektheit der Patterns hängt hier von der Korrektheit der Workflow-Instanzen ab, die Zahl der identifizierten Patterns von der Erfahrung des Prozessentwicklers.

Alternativ ist auch ein Top-Down-Ansatz [13] zur Gewinnung von Patterns möglich. Dabei wird zunächst ein abstraktes Referenzmodell aufgestellt, in dem die Prozessschritte als generische Aktivitäten repräsentiert werden. In einem weiteren Schritt werden dann domänenspezifische Aspekte, Zustandsübergänge etc. zu den Aktivitäten hinzugefügt. Zuletzt werden die generischen Aktivitäten des abstrakten Referenzmodells im gewünschten Detaillierungsgrad verfeinert. Dabei lassen sich auch Design-Alternativen als unterschiedliche Workflow Patterns in die Verfeinerungen integrieren. Die gewonnenen Patterns sind korrekt, wenn das finale Referenzmodell, das diese Patterns enthält, korrekt ist.

In Bezug auf die Entwicklung von Workflow Privacy Patterns hat die unmittelbare Anwendung dieser beiden Ansätze den Nachteil, dass der wesentliche Zweck der Patterns nicht explizit berücksichtigt wird: Die korrekte Umsetzung von gesetzlichen Vorgaben in Geschäftsprozessen oder die Validierung, dass bestimmte Geschäftsprozesse gesetzeskonform umgesetzt wurden. Wir haben daher analysiert, welche domänenspezifischen Attribute in der Datenschutzgesetzgebung Anwendung finden, und aus diesen ein datenschutz-spezifisches Beschreibungsschema für Workflow Privacy Patterns entwickelt, das sich sowohl Top-Down als auch Bottom-Up zur Definition von Patterns eignet.

3 Entwicklung des Schemas zur Beschreibung von WPPs

Die Entwicklung unseres Beschreibungsschemas für WPPs folgt dem gestaltungsorientierten Forschungsparadigma [17]: (1) Beschreibung der Problemdomäne Datenschutz in Geschäftsprozessen, (2) Entwicklung und Beschreibung des Beschreibungsschemas und (3) Einsatz des Beschreibungsschemas anhand der Entwicklung eines konkreten WPPs mit anschließender Untersuchung von dessen Anwendbarkeit.

Wir beginnen mit der Identifizierung relevanter Attribute zum Zweck der Unterstützung von Entwicklern bei der Implementierung von Datenschutz in Geschäftsprozessen. Dabei ist es wichtig, diejenigen Attribute zu identifizieren, die für eine eigenständige Konzipierung eines WPPs notwendig sind, um Anforderung A3 zu erfüllen. Dazu gehört auch, dass durch die Nutzung des Beschreibungsschemas durch den Entwickler keine anderen Quellen mehr benötigt werden, um das WPP zu implementieren. Das Beschreibungsschema soll als Grundlage für eine Modellierung dienen, wobei die Auswahl der Modellierungssprache einen nachgeordneten Schritt darstellt. Insofern wird Anforderung A1 an dieser Stelle nicht weiter berücksichtigt.

Aus der Themensicht ergeben sich im Hinblick auf den Zweck und die Nutzung des Beschreibungsschemas zwei Aspekte: der Prozess-Aspekt, der sich mit den Eigenschaften eines komplexen WPs beschäftigt, und der Datenschutz-Aspekt, in den Datenschutz-relevante Zusatzinformationen fallen. Ein Überblick über sämtliche im Folgenden erläuterte Attribute ist der Tabelle 1 zu entnehmen.

Tabelle 1. Attributsammlung für WPPs nach Themensicht

Prozess-Aspekt		*Datenschutz-Aspekt*	
1	Daten	9	Paragraph
2	Ressourcen	10	WPP-Kategorie
3	Kontrollfluss	11	Nutzungsvorschrift
4	Exception-Handling	12	Auswirkungen auf andere WPPs
5	Anwendungsvoraussetzungen	13	Auswirkungen durch andere WPPs
6	Betroffene Prozesse	14	Technische Voraussetzungen
7	Einsatzstelle	15	Optionale Bestandteile
8	Interaktion mit anderen Prozessen	16	Ausnahmeregelungen
		17	Modifikationsbedarf

Für den Prozess-Aspekt lassen sich die folgenden Attribute eines WPPs aus den Vorarbeiten zu WPs identifizieren: Zum einen gibt es die vier WP-Kategorien (1-4), auf denen WPPs u.a. basieren. Weiterhin sind Situationen oder Handlungen (5) zur Durchführung von WPPs relevant, die gegeben sein müssen, damit das WPP gestartet werden kann. Ausgehend davon werden Informationen darüber benötigt, in welche Prozesse (6) das Pattern eingesetzt werden kann. Im Zusammenhang mit (5) und (6) kann dann die konkrete Einsatzstelle (7) spezifiziert werden, an der das WPP in einen Prozess integriert werden kann. Ein weiteres Attribut beschreibt die Interaktion (8) mit anderen (Sub-) Prozessen oder WPPs.

Für den Datenschutz-Aspekt wurden Paragraphen aus dem Bundesdatenschutzgesetz [1] und der EU-Datenschutzgrundverordnung [2] analysiert, um daraus Attribute zu identifizieren. So finden sich im Aspekt Datenschutz Attribute wie die dem WPP zugrundeliegenden Paragraphen (9) des Gesetzestextes. Ebenfalls ist eine Einordnung des WPPs in eine der drei bereits genannten WPP-Perspektiven (10) hilfreich. Eine weitere Information ist, wann die betreffende Regelung zwingend anzuwenden ist (11). Diese Information kann relevant sein, da bestehende Gesetze zum Datenschutz durch andere ersetzt oder ergänzt werden können (z.B. Anwendung der EU-DSGVO ab 25. Mai 2018). Unter Umständen gibt es andere Prozesse oder WPPs, die sich auf das vorliegende WPP auswirken oder es einschränken (12). Diese Auswirkungen können ebenso umgekehrt (13) vorhanden sein. Ein weiteres WPP-Attribut betrifft die technischen Anforderungen (14), die erfüllt sein müssen, um das WPP ausführen zu können. Ein WPP kann abhängig vom Inhalt des betreffenden Paragraphen optionale Bestandteile (15) haben. Der Anwender besitzt dadurch die Entscheidungsfreiheit, ob er das WPP mit oder ohne diese optionalen Bestandteile umsetzt. Es ist wichtig, Ausnahmeregelungen mitaufzuführen, in denen das Pattern nicht oder anders angewendet wird (16). Der letzte Punkt ist der Modifikationsbedarf (17). Er dokumentiert den Fall, dass das WPP eine abstrakte Rechtsvorschrift ggf. in modifizierter Form abbilden muss, um abstrakte Rechtsbegriffe in konkrete Modellierungen zu überführen.

Anhand der Themensicht lassen sich die einzelnen Attribute zwar gut identifizieren und festhalten, jedoch wird dabei Anforderung 2 für WPPs nicht berücksichtigt. Die jeweils notwendigen Informationen für die Implementierung und die Modellierung sind als solche nicht klar erkennbar oder zuzuordnen. Deshalb empfiehlt sich für das Beschreibungsschema eine Anwendungssicht, die die Attribute Aspekten zuordnet, die für jeweils unterschiedliche Anwendungen eines WPPs sinnvoll sind.

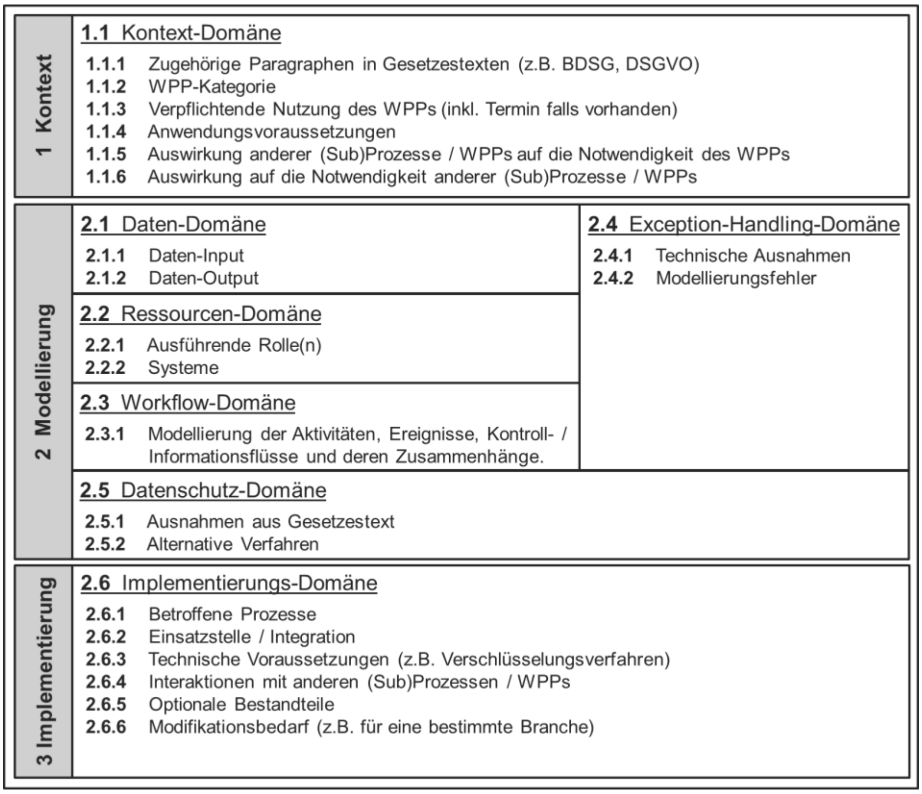


Abbildung 2. Beschreibungsschema für Workflow Privacy Patterns

Um ein WPP als solches identifizieren und darauf referenzieren zu können sind verschiedene Kontextinformationen (5, 9-13) erforderlich. Daraus ergibt sich die Kontext-Domäne, die dem Kontext-Aspekt zuzuordnen ist. Neben diesem Aspekt wird weiterhin eine Aufteilung der WPP-Attribute in zwei weitere Aspekte benötigt, die sich auf die Modellierung und die Implementierung fokussieren. Zum Aspekt „Modellierung" gehören die vier WP-Perspektiven (1-4) als eigenständige Domänen mit ihren Attributen, sowie eine spezielle Datenschutzdomäne, die sich mit den Ausnahmeregelungen (16) beschäftigt. Im Aspekt „Implementierung" befindet sich die Implementierungs-Domäne, die sämtliche Attribute (6-8, 14-15, 17) beinhaltet, um das WPP Ende-zu-Ende in einen Prozess integrieren zu können. Daraus ergibt sich ein umfassendes Schema, das in Abbildung 2 dargestellt ist. Über die aufgeführten Bereiche hinaus sind weitere Attribute aus anderen Domänen denkbar, die das Bild eines WPP vervollständigen. In diesem Beitrag wurde jedoch der Fokus auf die Anforderungen eines Entwicklers gelegt um ihm mit dem Beschreibungsschema ein nützliches Werkzeug zur Umsetzung von Datenschutz in Geschäftsprozessen zu bieten. Aus diesem Grund werden Domänen bzw. Attribute, die über diesen Zweck hinausgehen, an dieser Stelle nicht berücksichtigt.

4 Einsatz des Beschreibungsschemas

4.1 Anwendung des Beschreibungsschemas auf § 33 BDSG

Das Beschreibungsschema wurde in Tabelle 2 überführt und auf § 33 BDSG *[18]* angewendet, um die Anwendbarkeit des Schemas zu demonstrieren. § 33 BDSG beinhaltet Regelungen zur Benachrichtigung eines Betroffenen bei der Speicherung personenbezogener Daten und entspricht einem Crosscutting Privacy Pattern:

„Werden erstmals personenbezogene Daten für eigene Zwecke ohne Kenntnis des Betroffenen gespeichert, ist der Betroffene von der Speicherung, der Art der Daten, der Zweckbestimmung der Erhebung, Verarbeitung oder Nutzung und der Identität der verantwortlichen Stelle zu benachrichtigen. Werden personenbezogene Daten geschäftsmäßig zum Zweck der Übermittlung ohne Kenntnis des Betroffenen gespeichert, ist der Betroffene von der erstmaligen Übermittlung und der Art der übermittelten Daten zu benachrichtigen. Der Betroffene ist in den Fällen der Sätze 1 und 2 auch über die Kategorien von Empfängern zu unterrichten, soweit er nach den Umständen des Einzelfalles nicht mit der Übermittlung an diese rechnen muss." [18]

Zunächst wurde die Tabelle anhand des Schemas erstellt und mit den identifizierten und abgeleiteten Informationen aus § 33 BDSG befüllt. Dafür wurde der Detaillierungsgrad des BDSG übernommen. Wie in der Tabelle 2 zu sehen ist, konnte nicht jeder Prüfaspekt des Beschreibungsschemas ausgefüllt werden. Tabelle 2 hilft dem Modellierer dabei, einen Überblick über das WPP zu erhalten und alle notwendigen Informationen zu identifizieren und zu strukturieren. Anhand der Attribute können die Informationen aus den zugrundeliegenden Paragraphen abgeleitet werden. Hier wurde ein Paragraph ausgewählt und entlang der Domänen die dazugehörigen WPP-Eigenschaften ermittelt. Dadurch wurde ersichtlich, ob andere Paragraphen Einfluss auf das WPP haben. Diese Angaben sind für die Modellierung dieser Rechtsnorm ausreichend.

Tabelle 2. Anwendung des Schemas auf das WPP „Benachrichtigung des Betroffenen"

	Domäne	*Attribut*	*Anwendung auf „Benachrichtigung des Betroffenen"*
1 Kontext	1.1	1.1.1	§ 33 BDSG
		1.1.2	Crosscutting Privacy Pattern
		1.1.3	Die Nutzung des WPPs ist verpflichtend, sobald personenbezogene Daten gespeichert werden.
		1.1.4	Speicherung personenbezogener Daten
		1.1.5	§ 29 BDSG
		1.1.6	Nicht definiert
2 Modellierung	2.1	2.1.2	Nicht definiert
		2.1.2	Benachrichtigung an den Betroffenen
	2.2	2.2.1	Die Organisationseinheit, die verantwortlich für die Speicherung der personenbezogenen Daten des Betroffenen ist: "Verantwortliche Stelle" (§ 33 BDSG).
		2.2.2	Nicht definiert
	2.3	2.3.1	Aktivitäten / Subprozesse: Erstmaligkeit der Speicherung prüfen, Ausnahmeregelungen prüfen, Zweck der Speicherung ermitteln, Betroffenen benachrichtigen. Gateways/Konnektoren: XOR bei Anti- oder Kontravalenz (positives oder negatives Prüfergebnis), OR bei Disjunktion je nach Ermittlungsergebnis: Eigene Zwecke (A), Zweck der Übermittlung (B). Ereignisse: Personenbezogene Daten werden gespeichert (Start), Betroffener ist benachrichtigt (Ende), Pflicht zur Benachrichtigung besteht nicht (Ende). Kontrollflüsse: Zwischen allen Aktivitäten bzw. Subprozessen, Ereignissen und Gateways.
	2.3	2.3.1	Anmerkungen zu "Betroffenen benachrichtigen" bei Ermittlungsergebnis A: von der Speicherung, Art der Daten, Erhebungszweck, Verarbeitung/-Nutzung, Identität der verantwortlichen Stelle, ggf. Empfängerkategorien. Anmerkungen zu "Betroffenen benachrichtigen" bei Ermittlungsergebnis B: von der erstmaligen Übermittlung, Art der Daten, ggf. Empfängerkategorien.
	2.4	2.4.1	Nicht definiert
		2.4.2	Nicht definiert
	2.5	2.5.1	§ 33 Absatz 2 Artikel 1 bis 9
		2.5.2	Der Eintritt einer der Ausnahmen § 33 Absatz 2 Artikel 1 bis 9 führt zum Ende des WPPs, also zur Abbruch-

			bedingung "Keine Benachrichtigung des Betroffenen notwendig".
3 Implementierung	3.1	3.1.1	Betroffen ist jeder Prozess, in dem personenbezogene Daten gespeichert werden.
		3.1.2	Vorhergehende Aktivität im Prozess: Speicherung personenbezogener Daten
		3.1.3	Nicht definiert
		3.1.4	Nicht definiert
		3.1.5	Nicht definiert
		3.1.6	Kenntnis des Betroffenen prüfen, Allgemeine Ausnahmefallregelungen prüfen, Zweck der Speicherung prüfen, Ausnahmefallregelungen für Übermittlungszwecke prüfen, Ausnahmefallregelung für eigene Zwecke prüfen, Betroffenen benachrichtigen

4.2 Bewertung der Anwendbarkeit des Beschreibungsschemas

In dem Beschreibungsschema existiert eine Sammlung von WPP-Attributen, die nach Kontext, Modellierung und Implementierung strukturiert sind. Wir haben festgestellt, dass hier ein Leitfaden zur Strukturierung von Gesetzestexten bzw. Paragraphen notwendig werden, um eine einheitliche Beschreibung zu gewährleisten.

Wenn Informationen im zugrundeliegenden Paragraphen fehlen oder kontextbedingt irrelevant sind, fehlen Angaben im Beschreibungsschema. Hier scheint eine Kennzeichnung der obligatorischen oder optionalen Attribute sinnvoll.

Durch die Demonstration des Beschreibungsschemas konnte noch keine Aussage über die Generalisierbarkeit des Beschreibungsschemas getroffen werden. Dazu müssen weitere Paragraphen anhand des Beschreibungsschemas strukturiert werden. Insbesondere die Überprüfung für die anderen beiden WPP-Kategorien ist relevant. Ein Beispiel dafür ist die Kategorie der Privacy Processes/Patterns. Bei diesen sind die Attribute 3.1.1 und 3.1.2 per se nicht vorhanden, sodass eine Evaluierung der Handhabung im Rahmen der Nutzung des Beschreibungsschemas notwendig wird.

5 Fazit und Ausblick

Aufgrund der steigenden Anforderungen an die Umsetzung des Datenschutzes in Geschäftsprozessen, wurden WPPs als Unterstützung für Prozessentwickler und Prozessmodellierer identifiziert. Im vorliegenden Beitrag wurde ein Beschreibungsschema erarbeitet, das zur Entwicklung und Vergleichbarkeit von WPPs eingesetzt werden kann. Für eine standardisierte Entwicklung und Modellierung von WPPs bedarf es jedoch noch weiterführender Forschung. Anhand der erkannten Herausforderungen, die sich aus dem bisherigen Entwicklungsstand des Beschreibungsschemas ergeben, lassen sich dafür folgende Themen ableiten:

- *Entwicklung eines objektiven Leitfadens:* Wie können Gesetztestexte standardisiert, strukturiert und in ein Beschreibungsschema überführt werden?
- *Ermittlung der Notwendigkeit aller Attribute:* Welche Attribute sind für welche WPP-Kategorien obligatorisch, optional oder irrelevant?
- *Validierung des Beschreibungsschemas:* Muss das Beschreibungsschema für die Anwendung auf alle Paragraphen aus allen drei WPP-Kategorien modifiziert werden und falls ja - wie?
- *Weiterentwicklung des Beschreibungsschemas:* Welche Domänen sind über die bisher berücksichtigten Domänen hinaus für WPPs relevant und um welche zusätzlichen Attribute sollte das Beschreibungsschema ergänzt werden?

Basierend auf der Entwicklung des Beschreibungsschemas für WPPs und der Diskussion dessen Anwendbarkeit anhand eines Beispiels können zudem folgende Forschungsthemen in Ausblick gestellt werden:

- *Erstellung eines Vorgehensmodells:* Wie können WPPs basierend auf Gesetzestexten und Datenschutzprinzipien nach einem einheitlichen Vorgehen entwickelt und modelliert werden?
- *Erstellung eines Modells zur Integration und Implementierung:* Wie können WPPs in Unternehmen bzw. deren Geschäftsprozesse nach einem einheitlichen Vorgehen integriert und implementiert werden?
- *Detaillierungsgrad:* Welcher Detaillierungsgrad eignet sich für WPPs?
- *Ausnahmeregelungen:* Wie können Subprozesse für die Handhabung von Ausnahmereglungen gestaltet werden und welchen Grad der Detaillierung sollten diese aufweisen?
- *Modellierungssprache:* Welche Modellierungssprache eignet sich am besten zur Modellierung von WPPs?

Damit WPPs erfolgreich in Unternehmen implementiert und genutzt werden können, kann die Bearbeitung der aufgeführten Themen hinsichtlich der Optimierung des Beschreibungsschemas und der weiteren Anwendung wichtige Beiträge liefern.

Literaturverzeichnis

1. Bundesdatenschutzgesetz in der Fassung der Bekanntmachung vom 14. Januar 2003. BDSG. Bundesgesetzblatt I, 66–88 (2003)
2. Europäisches Parlament: Verordnung (EU) 2016/679 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG. EU-DSGVO (2016)
3. Anke, J., Berning, W., Schmidt, J., Zinke, C.: IT-gestützte Methodik zum Management von Datenschutzanforderungen. HMD 54, 67–83 (2016)
4. Mont, M.C., Pearson, S.: An Adaptive Privacy Management System for Data Repositories. In: Katsikas, S.K. (ed.) Trust, privacy, and security in digital business. Second international conference TrustBus 2005, Copenhagen, Denmark, pp. 236–245. Springer, Berlin (2005)

5. Karjoth, G., Schunter, M., Waidner, M.: Platform for Enterprise Privacy Practices. Privacy-Enabled Management of Customer Data. In: Goos, G., Hartmanis, J., van Leeuwen, J., Dingledine, R., Syverson, P. (eds.) Privacy Enhancing Technologies, 2482, pp. 69–84. Springer, Berlin, Heidelberg (2003)
6. Buchmann, E., Anke, J.: Privacy Patterns in Business Processes. In: Eibl, Maximilian, Gaedke, Martin (eds.) INFORMATIK 2017, pp. 793–798. Ges. für Informatik, Bonn
7. Lange, J.A.: Sicherheit und Datenschutz als notwendige Eigenschaften von computergestützten Informationssystemen. Ein integrierender Gestaltungsansatz für vertrauenswürdige computergestützte Informationssysteme (2005)
8. Rodeck, M., Voigt, C., Schnütgen, A., Schiering, I., Decker, R.: Toolgestützte Assessments zu Datenschutz und Datensicherheit in kleinen und mittelständischen Unternehmen. In: Plödereder, E. (ed.) INFORMATIK 2014, pp. 575–586. Ges. für Informatik, Bonn (2014)
9. Karjoth, G.: Datenschutzkonforme Geschäftsprozesse. In: Barton, T., Erdlenbruch, B., Herrmann, F., Marfurt, K., Müller, C., Seel, C. (eds.) Prozesse, Technologie, Anwendungen, Systeme und Management. Angewandte Forschung in der Wirtschaftsinformatik, pp. 20–30. mana-Buch, Heide (2015)
10. Masellis, R. de, Ghidini, C., Ranise, S.: A Declarative Framework for Specifying and Enforcing Purpose-Aware Policies. In: Foresti, S. (ed.) Security and Trust Management, pp. 55–71. Springer, Cham (2015)
11. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Abstraction and Reuse of Object-Oriented Design. In: European Conference on Object-Oriented Programming. Springer (1993)
12. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M.: Pattern-Oriented Software Architecture, A System of Patterns. Wiley, s.l. (2013)
13. van Dijk, A.: Contracting Workflows and Protocol Patterns. In: Goos, G., Hartmanis, J., van Leeuwen, J., ter Hofstede, A., van der Aalst, W.M.P., Weske, M. (eds.) Business Process Management, pp. 152–167. Springer, Berlin, Heidelberg (2003)
14. Russel, N., ter Hofstede, A.H.M., Edmond, D. and van der Aalst, W.M.P.: Workflow Resource Patterns
15. Gündogdu, F.: Analyse zur Verwendung der Workflow Pattern und der Business Process Modelling and Notation bei der Modellierung von Prozessen. Dissertation, http://dbis.eprints.uni-ulm.de/1101/ (2014)
16. van der Aalst, W.M.P., ter Hofstede, A.H.M., Kiepuszewski, B., Barros, A.P.: Workflow Patterns. Distributed and Parallel Databases 14, 5–51 (2003)
17. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75–105 (2004)
18. § 33 Benachrichtigung des Betroffenen. In: Bundesdatenschutzgesetz in der Fassung der Bekanntmachung vom 14. Januar 2003. BGBl. I (2003)

ALADDIN – Vorschlag eines Analyse- und Berechnungsmodells zur Investitionsbewertung für ein unternehmensweites Datenqualitätsmanagement

Thomas Schäffer[1], Christian Leyh[2] und Helmut Beckmann[1]

[1] Hochschule Heilbronn, Studiengang Wirtschaftsinformatik, Heilbronn, Deutschland
{thomas.schaeffer,helmut.beckmann}@hs-heilbronn.de

[2] Technische Universität Dresden, Lehrstuhl für Wirtschaftsinformatik, insb. Informationssysteme in Industrie und Handel, Dresden, Deutschland
christian.leyh@tu-dresden.de

Abstract. Stammdaten müssen heutzutage mehr denn je als Grundlage der digitalen Wirtschaft angesehen werden. Dies resultiert vor allem daraus, dass sich durch die zunehmende Digitalisierung von Unternehmen eine Transformation ursprünglich starrer Wertschöpfungsketten hin zu dynamischen Wertschöpfungsnetzwerken vollzieht. Die Bereitstellung einer angemessenen Stammdatenqualität ist dabei eine entscheidende Voraussetzung für vernetzte Geschäftsprozesse in und vor allem zwischen Unternehmen. Zahlreiche Studien zeigen jedoch, dass viele Unternehmen Probleme bezogen auf ihre Stammdatenqualität haben, obwohl Unternehmen mittlerweile den Zusammenhang zwischen Datenqualität und Rentabilität erkannt haben. Jedoch fehlt es oftmals an konkretem Nutzenpotential, um eine entsprechende Datenqualitäts-Unternehmenskultur in Form eines Datenqualitätsmanagements aufzubauen und mit entsprechenden Ressourcen auszustatten. Daraus leitet sich ein hoher Bedarf für eine quantifizierbare Kosten-Nutzen-Berechnung für Investitionsvorhaben zur Verbesserung der Datenqualität ab. Dies aufgreifend beschreibt der vorliegende Beitrag die Entwicklung eines Analyse- und Berechnungsmodells zur Bewertung von Investitionen im Bereich des Datenqualitätsmanagements und gibt Impulse zur Ermittlung des Wertbeitrags.

Keywords: Datenmanagement, Datenqualität, Design Science, Investitionsrechnung, Stammdaten, Wertbeitrag

1 Einleitung

Von dem starken und vor allem schnellen Voranschreiten der Digitalisierung der alltäglichen Lebensbereiche bleibt auch die Arbeits- und Geschäftswelt nicht unbeeinflusst. Technologien wie Cloud Computing, die zunehmende Automatisierung, z. B. unter dem Schlagwort Industrie 4.0, oder der Einsatz mobiler Endgeräte sind nur einige Beispiele der Digitalisierung, die den Unternehmen völlig neue Möglichkeiten im Geschäftsalltag eröffnen [1–3]. Dabei sehen sich Unternehmen aktuell mit

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Fragestellungen konfrontiert, wie und in welchem Umfang die Digitalisierung das Unternehmen beeinflusst und verändert. Im Zuge der Digitalisierung verschwimmt die Abgrenzung wertschöpfender und unterstützender Prozesse, wodurch eine Transformation von ursprünglich nachgelagerten Prozessen hin zu ganzheitlichen Wertschöpfungsnetzwerken erfolgt [4].

Den Potenzialen der Digitalisierung stehen aber auch verschiedene Bedenken bzw. Risiken gegenüber. Speziell die zunehmende Vernetzung der Systeme wird von Unternehmen oftmals als Problem wahrgenommen [2, 5]. Doch gerade die Vernetzung innerhalb des Unternehmens und auch über die Unternehmensgrenzen hinweg wird auch in Zukunft stetig wachsen (siehe z. B. [4]).

Dies aufgreifend sind vor allem Stammdaten und deren Qualität für den Informationsaustausch innerhalb von Wertschöpfungsnetzwerken von essentieller Bedeutung, welche im Rahmen der voranschreitenden Digitalisierung immer weiter zunehmen. Somit bilden Stammdaten heutzutage mehr denn je die Grundlage der digitalen Wirtschaft [3, 6]. Das Erreichen und die Sicherstellung einer angemessenen Stammdatenqualität ist eine essentielle Voraussetzung für eine effiziente und effektive unternehmensinterne und vor allem unternehmensübergreifende Zusammenarbeit [7–12].

Zahlreiche Studien zeigen jedoch, dass viele Unternehmen Probleme bezogen auf ihre Stammdatenqualität sehen [3, 9, 13–17]. Auch ergaben diese Studien, dass, um das Verbesserungspotenzial der Stammdatenqualität auszuschöpfen, ein adäquates unternehmensweites Datenqualitätsmanagement (DQM) etabliert werden sollte. Laut [15] sind beispielsweise 79 % der Unternehmen der Meinung, dass eine schlechte Stammdatenqualität sich stark bis sehr stark negativ auf die Rentabilität im Unternehmen auswirkt. Jedoch ist eine hohe Qualität in Verbindung mit einer umfangreichen Bewirtschaftung der Stammdaten selten gegeben. So sehen 84 % der Unternehmen einen hohen bis sehr hohen Aufwand zur Sicherstellung der Stammdatenqualität und lediglich 11 % der Unternehmen geben an, dass das Budget für die Umsetzung von Qualitätsmaßnahmen in ausreichendem Maße zur Verfügung steht. Überraschenderweise konstatieren viele Geschäftsführer den direkten Zusammenhang zwischen Qualität und Rentabilität, jedoch die entsprechenden Investitionen werden nicht getätigt. Als wesentlicher Grund wird eine fehlende Wirtschaftlichkeitsbetrachtung angeführt, die das konkrete Nutzenpotential eines derartigen Investitionsvorhabens zur Verbesserung der Stammdatenqualität quantifiziert und somit auch „rechtfertigt“ [16, 17].

An dieser Stelle setzt der vorliegende Beitrag an. Es soll ein Instrumentarium (konkret ein Analyse- und Berechnungsmodell) konzipiert werden, mit dem Investitionsvorhaben zur Verbesserung der Stammdatenqualität bewertet werden können. Ferner soll das Instrumentarium dazu beitragen, den Wertbeitrag eines Datenqualitätsmanagements zum Unternehmenserfolg zu ermitteln. Daraus ergeben sich die zwei zentralen Forschungsfragen für diesen Beitrag:

1. Wie muss ein Analyse- und Berechnungsmodell zur Investitionsbewertung eines unternehmensweiten Datenqualitätsmanagements gestaltet sein?

2. Welche bereits existierenden Verfahren der Investitionsrechnung können

angewendet bzw. adaptiert werden, um die Investitionsbewertung eines unternehmensweiten Datenqualitätsmanagements zu ermöglichen?

Um diese Fragen zu beantworten und zur Beschreibung eines ersten Vorschlags für ein derartiges Modell, ist der Beitrag wie folgt aufgebaut: sich der Einleitung anschließend, werden kurz die begrifflichen Grundlagen dargelegt. In Kapitel 3 wird die Methodik zur Entwicklung des Analyse- und Berechnungsmodells erläutert. Anschließend wird unser Analyse- und Berechnungsmodell ALADDIN (**A**na**L**ysis and C**A**lculation Mo**D**el for the Assessment of **D**ata Quality Management **IN**vestments) zur Bewertung von Investitionsvorhaben im Bereich des Datenqualitätsmanagements beschrieben und dessen Anwendung anhand eines konkreten Fallbeispiels dargelegt. Wir schließen den Beitrag mit Implikationen für Praktiker und Wissenschaftler sowie mit einem Ausblick auf weiteren Forschungsbedarf.

2 Begriffliche Grundlagen

2.1 Stammdaten und Datenqualitätsmanagement

Stammdaten beschreiben kritische Geschäftsobjekte eines Unternehmens und bezeichnen Produkte, Lieferanten, Kunden, Mitarbeiter bzw. ähnliche Gegenstände, die nur selten Änderungen erfahren [11].

Datenqualität ist ein „Maß für die Eignung der Daten für spezifische Anforderungen in Geschäftsprozessen, in denen sie verwendet werden. Die Datenqualität ist ein mehrdimensionales, kontextuelles Konzept, da es nicht mit einer einzigen Funktion beschrieben werden kann, sondern auf der Basis verschiedener Datenqualitätsdimensionen und Metriken“ [6]. Typische, häufig verwendete Qualitätsdimensionen sind Fehlerfreiheit, Korrektheit, Vollständigkeit, Relevanz, Konsistenz und Aktualität. Die Datenqualität wird deshalb oft mit dem Begriff “fitness for use” assoziiert [18].

Datenqualitätsmanagement ist somit das qualitätsorientierte Datenmanagement und umfasst die Modellierung, Erzeugung, Verarbeitung, Speicherung und Darstellung von Daten mit dem Ziel der Sicherstellung einer möglichst hohen Datenqualität.

2.2 Investitions- und Wirtschaftlichkeitsrechnung

Die Investitionsrechnung dient dazu, die absolute und relative Vorteilhaftigkeit von Investitionsvorhaben zu beurteilen sowie die optimale Nutzungsdauer und den optimalen Ersatzsatzpunkt von Investitionsobjekten zu bestimmen [19]. Investitionen haben einen einmaligen Charakter und übersteigen das Volumen der „üblichen Tagesausgaben“. Typischerweise sind die Einnahmen- und Ausgabenströme von Investitionsvorhaben unregelmäßig über die gesamte Investitions- bzw. Nutzungsdauer verteilt [20].

Zur Konkretisierung und Berechnung von Investitionsvorhaben dient ein Wirtschaftlichkeitsnachweis bzw. „Business Case“. Dieser umfasst alle entscheidungsrelevanten Aspekte des Vorhabens mit dem Ziel, die wirtschaftliche

Vorteilhaftigkeit und strategische Konformität des Gesamtprojekts aufzuzeigen und eine abschließende Entscheidung über dessen Ausführung zu ermöglichen.

3 Methodische Vorgehensweise

Unser Analyse- und Berechnungsmodell ALADDIN ist als Artefakt im Sinne der konstruktionsorientierten Forschung (Design Science) [21, 22] anzusehen, wodurch auch die Vorgehensweise zur Entwicklung von ALADDIN den Grundsätzen einer konstruktionsorientierten Forschung unterliegt. Dabei orientiere sich die Entwicklung an den Phasen von [23], welche diese Autoren im Sinne der Design Science für die Entwicklung von Reifegradmodellen vorschlagen. Die Entwicklungsschritte nach [23] folgen dabei im Allgemeinen den Phasen: Problemidentifizierung, Konstruktion und Evaluation. Auch wenn ALADDIN nicht als Reifegradmodell angesehen werden kann, wurden die Schritte für dessen Entwicklungsprozess auf Basis von [23] abgeleitet, da dies als passfähig im Sinne der allgemeinen Modellentwicklung angesehen wurde. Für den Entwicklungsprozess von Reifegradmodellen basierend auf den sieben Richtlinien (R) zur Durchführung von Design Science nach [21] (R1 Artefakte als Designergebnis; R2 Problemrelevanz; R3 Evaluation; R4 Forschungsbeitrag; R5 Stringenz der Forschungsmethode; R6 Design als Suchprozess; R7 Kommunikation der Forschungsergebnisse) definieren [23] acht Anforderungen (vgl. Tabelle 1).

Tabelle 1. Design Science als Basis der methodischen Vorgehensweise

#	Anforderungen zur Entwicklung von Reifegradmodellen nach [23]	Referenzierte Richtlinie Design Science nach [21]	Bezug zu unserer Vorgehensweise
A1	Vergleich mit existierenden Reifegradmodellen	R1, R4	Analyse
A2	Iteratives Vorgehen	R6	Entwurf / Evaluation
A3	Evaluation	R3, R6	Evaluation
A4	Multimethodisches Vorgehen	R5	Analyse / Entwurf
A5	Aufzeigen der Problemrelevanz	R2	Analyse
A6	Problemdefinition	R2	Analyse
A7	Adressatengerechte Ergebnisbereitstellung	R7	Entwurf / Evaluation
A8	Wissenschaftliche Dokumentation	R7	Entwurf / Evaluation

Dies aufgreifend erfolgte die Entwicklung von ALADDIN in drei Phasen (vgl. Abbildung 1), welche in Tabelle 1 den Anforderungen von [23] zugeordnet wurden. Die Analyse-Phase umfasste dabei eine initiale Literaturanalyse zu bestehenden Investitionsverfahren (auch bezogen auf konkrete Fallbeispiele in der Literatur) sowie zur Problemstellung des DQMs. Aufbauend auf der daraus abgeleiteten Praxisrelevanz der Fragestellung sowie auf den Wirtschaftlichkeitsfaktoren für IT-Investitionen im Allgemeinen und für das DQM im Speziellen, wurde ein erster Vorschlag unseres Analyse- und Berechnungsmodells in der Entwurfsphase konzipiert. Mit einer ersten

Evaluation auf Basis eines konkreten Fallbeispiels wurde anschließend die erste Iteration des Entwicklungsprozesses von ALADDIN abgeschlossen. Detaillierte Erläuterungen zu den einzelnen Phasen erfolgen im sich anschließenden Kapitel.

Abbildung 1. Vorgehensweise zur Entwicklung von ALADDIN

4 Entwicklung des Analyse- und Berechnungsmodells

4.1 Analysephase

In der Analysephase wurden zwei Gestaltungselemente betrachtet: primäre Wirtschaftlichkeitsfaktoren für ein Datenqualitätsmanagement (1a) und sekundäre Wirtschaftlichkeitsfaktoren bei der Vorgehensweise von IT-Investitionen (1b). Die Erkenntnisse wurden aus einer Literaturanalyse angelehnt an [24, 25] abgeleitet. Im Rahmen der Analyse wurden folgende Publikationsdatenbanken ausgewählt: Google Scholar, Science Direct und SpringerLink. Zusätzlich wurde in den Konferenzbänden folgender Konferenzen gesucht: ECIS, ICIS und WI. Die Suchabfrage wurde mit folgenden Suchbegriffen und Kombinationen durchgeführt: (Datenqualität ODER „Data Quality“ ODER Stammdaten ODER „Master Data“) UND (Investition ODER Investment ODER „Cost“ ODER „Data Asset“). Das Ergebnis der Literaturanalyse umfasste final 24 Beiträge, die als Grundlage zur Entwicklung von ALADDIN dienten. Die Liste aller analysierten Beiträge sowie die daraus resultierenden Funktionsanforderungen können bei den Autoren angefragt werden.

Primären Wirtschaftlichkeitsfaktoren (1a)
Die Leitfragen zur Untersuchung von Gestaltungselement 1a waren:

- Welche primären Faktoren sind für ein DQM-Vorhaben zu erheben?
- Wie lassen sich Kosten und Nutzen mangelhafter Datenqualität klassifizieren?

Die primären Wirtschaftlichkeitsfaktoren eines Datenqualitätsmanagements sind im Wesentlichen: Datenqualitätskosten und fachliche Nutzenpotentiale [26]. Dabei lassen sich die Kosten wie folgt aufteilen: Kosten schlechter Datenqualität (z. B. Kosten für Verifikation und Korrektur der Daten) bzw. Kosten für die Verbesserung der Datenqualität (Präventions-, Entdeckungs- und Korrekturkosten). Für die Nutzenpotentiale kann zwischen strategischem (Wettbewerbsvorteile, Kundenzufriedenheit, etc.) und operativem Nutzen (verkürzte Prozessdurchlaufzeiten, etc.) unterschieden werden. Ferner lassen sich indirekte Kosten infolge mangelhafter Stammdatenqualität (Umsatzeinbußen, Verschwendung von Budgets, Fehlentscheidungen, Imageverlust, gerichtliche Auseinandersetzungen, etc.) als Nutzenpotentiale ausfassen. Generell gilt auch für die Investitionsbewertung eines unternehmensweiten Datenqualitätsmanagements die Maßgabe „fitness for use“ nach [18]. Das bedeutet, dass die Wirtschaftlichkeitsfaktoren abhängig vom Geschäftsprozess, den zugehörigen Stammdatenobjekten und den zuständigen Funktionsbereichen im Unternehmen zu definieren und zu erheben sind.

Sekundäre Wirtschaftlichkeitsfaktoren (1b)

Die Leitfragen zur Untersuchung von Gestaltungselement 1b waren:

- Welche Merkmale charakterisieren typischerweise ein Investitionsvorhaben?
- Welche Faktoren sind bei einer Investitionsrechnung zu berücksichtigen?

Investitionsvorhaben lassen sich anhand von fünf Merkmalen charakterisieren [19, 27]: Ziel und Zweck; Investitionsarten; Verfahren zur Bewertung von Investitionen; Kalkulationszinssatz; Verfahren zur Risikoberücksichtigung. Eine Übersicht der entsprechenden Merkmalsausprägung findet sich in Abbildung 2.

	Merkmal	**Merkmalsausprägung**					
Ziel und Zweck	**Invesitionsziele**	Kapazitätserweiterung		**Rationalisierung**		Ersatzbeschaffung	
	Betriebswirtschaftlicher Zweck	**Einzelinvestition: Ist das zur Entscheidung anstehende Objekt vorteilhaft?**		Alternativenvergleich und Rangfolgeproblem: Welches Objekt ist das beste, zweitbeste, drittbeste usw.?		Nutzungsdauer- und Ersatzproblem: Soll die Altanl. ersetzt oder weiterbetrieben werden? Wie lange soll die Neuanl. genutzt werden?	
Investitionsarten	**Nach Objekten**	Sach- bzw. Realinvestitionen	Finanzinvestitionen	**Immaterielle Investitionen**			
	Nach Funktionsbreichen	Forschungsinvestitionen	Fertigungsinvestitionen	Absatzinvestitionen	**Verwaltungsinvestitionen**		
	Nach Zwecksetzung	Netto-Investitionen		**Re-Investitionen**			
		Errichtungs- bzw. Gründungsinvestitionen	Erweiterungs-investitionen	Ersatz-investitionen	Rationalisierungs-investitionen	Diversifikations-investitionen	Sozial- und Sicherheitsinvestitionen
	Nach Nutzungsdauer	Kurzfristige Investitionen	**Mittelfristige Investitionen**	Langfristige Investitionen			
	Nach Investitionsträger	Private Investitionen	**Unternehmersiche Investitionen**	Öffentliche Investitionen			
Verfahren zur Bewertung von Investitionen	**Klassifizierung**	**Einzelinvestitionen**			Investitionsprogramme		
		Eindimensionale Verfahren		Mehrdimensionale Verfahren	Investition und Finanzierung		
		Statische Verfahren	**Dynamische Verfahren**	Nutzwertanalyse			
	Statische Verfahren	Kostenvergleichs-rechnung	Gewinnvergleichs-rechnung	Rentabilitäts-rechnung (ROI)	**Amortisations-rechnung (Payback)**		
	Dynamische Verfahren	**Kapitalwert-methode (NPV)**	**Interne Zinsfuß-Methode (IRR)**	**Annuitäten-methode**	**Amortisations-rechnung (Payback)**	Baldwin-Zinsfuß-Methode	
	"Value"-orientierte Verfahren	Value at Stake (VaS)					
Kalkulations-zinssatz	**Finanzierung der Investition**	**Eigenfinanzierung**	Fremdfinanzierung	Mischfinanzierung			
	Feststellungsverfahren	**Nach Nutzwertanalyse**	Nach Opportunitätskosten	**Mittelwert aus der Praxis**	Kapitalmarktmodell CAPM	**Kapitalmarktmodell WACC**	
	Verwendung	Unternehmensweit	**Investitionsspezifisch**	Nach Organisationseinheiten	Nach In- und Ausland	Nach Zwecksetzung	
Verfahren zur Risiko-berücksichtigung bei Investitionen	**von der Praxis akzeptierte Methoden**	**Sensitivitätsanalyse**			**Amortisationszeit**	Korrekturverfahren	wahrscheinlichster Wert
		Dreifach-Rechnung	**Zielgrößen-Änderungsrechnung**	**Kritische-Werte-Rechnung**			
	von der Praxis nicht akzeptierte Methoden	Entscheidungs-baumverfahren	Risiko-analyse				

Abbildung 2. Morphologischer Kasten zur Charakterisierung einer Investition

Investitionen besitzen im Allgemeinen die Eigenschaften, dass sie das Volumen der

„üblichen Tagesausgaben“ übersteigen; dass sie geplant werden und einem festgelegten Ziel dienen; dass sie einen Nutzen haben; dass die Einnahmen- und Ausgabenströme von Investitionsvorhaben unregelmäßig über die gesamte Investitions- bzw. Nutzungsdauer verteilt sind. Die daraus abgeleiteten sekundären Wirtschaftlichkeitsfaktoren [20] sind: Zeithorizont, Terminierung, Steuersatz, Inflationsrate, Abschreibungsmethode, Abschreibungsdauer, Zahlungszeitpunkt und Kalkulationszinsfuß.

Aus der Analysephase ist zusammenfassend zu erkennen, dass es zwar zahlreiche Modelle für Investitionsberechnungen für IT-Projekte gibt, jedoch nur wenige adressieren das Themengebiet des DQMs (siehe z. B. [28]). Insbesondere fehlt es an einem „ease of use“ Berechnungsmodell für Praktiker, mit dessen Hilfe DQM-Investitionsvorhaben bewertet werden können. Zur Feststellung, ob das durch ALADDIN berechnete Investitionsvorhaben wirtschaftlich vorteilhaft, neutral oder nicht vorteilhaft sind, wird gemäß Abbildung 3 entschieden. Dabei werden die entsprechenden Bewertungskriterien mit spezifischer Berechnungsmethode ermittelt und zu einem gemeinsamen Bewertungsindex aggregiert.

	Methoden		Investition ist vorteilhaft	Investition ist eben noch vorteilhaft	Investition ist unvorteilhaft
Bewertungs-kriterien	nach der Kapitalwertmethode	(C_0= Kapitalwert)	$C_0 > 0$	$C_0 = 0$	$C_0 < 0$
	nach der Interne Zinsfuß-Methode	(r = Rendite; i = Kalkulationszinssatz)	$r > i$	$r = i$	$r < i$
	nach der Annuitätenmethode	(DJÜ = durchschnittliche jährlicher Überschuss)	DJÜ > 0	DJÜ = 0	DJÜ < 0
	nach der statischen Amortisationsrechnung	(t = Amortisationszeit)	$t < t_{max}$	$t = t_{max}$	$t > t_{max}$
	nach der dynamischen Amortisationsrechnung		$t_d < t_{max}$	$t_d = t_{max}$	$t_d > t_{max}$
Kapitalwert-methode	**Interpretationsmatrix**		$C_0 > 0$	$C_0 = 0$	$C_0 < 0$
	Der Investor gewinnt das eingesetztes Kapital zurück.		✔	✔	✘
	Der Investor erhält eine Verzinsung in Höhe seines Kalkulationszinssatzes auf die jeweils ausstehenden Beträge.		✔	✔	✘
	Der Investor gewinnt einen barwertigen Überschuss.		✔	✘	✘
	Der Investor erleidet einen barwertigen Verlust.		✘	✘	✔
Interne Zinsfuß-Methode	**Interpretationsmatrix**		$r > i$	$r = i$	$r < i$
	Der Investor gewinnt das eingesetzte Kapital zurück.		✔	✔	✘
	Der Investor erhält eine Verzinsung in Höhe seines Kalkulationszinssatzes auf die jeweils ausstehenden Beträge.		✔	✔	✘
	Der Investor erhält eine Extraverzinsung.		✔	✘	✘
	Der Investor erhält nicht die geforderte Mindestverzinsung.		✘	✘	✔

Abbildung 3. Übersicht von Kriterien zur Investitionsbewertung (in Anlehnung an [19])

4.2 Entwurfsphase

Aus der Analysephase konnten neben den Gestaltungselementen für die Entwurfsphase auch elf Funktionsanforderungen zur Konstruktion eines dynamischen Analyse- und Berechnungsinstruments erarbeitet werden. Im Wesentlichen zählen dazu eine Excel-basierte Software zur umfangreichen Auswertung, Implementierung von drei Investitionsrechenverfahren bzw. zwei Verfahren zur Risikobewertung, Berücksichtigung von Preissteigerungen und Steuern, grafische Aufbereitung der Zahlungsreihen und Ergebnisse, automatisierte Investitionsbewertung durch Entscheidungsregeln mit Gewichtungsfaktoren, flexible Parametereingabe von allen wichtigen Wirtschaftlichkeitsfaktoren und Ermittlung des optimalen Zeitpunkts.

Die Gesamtarchitektur des Analyse- und Berechnungsmodells besteht aus den drei Hauptkomponenten *Kostenmodell*, *Nutzenmodell* und *Wirtschaftlichkeitsmodell* (siehe Abbildung 4): Das **Kostenmodell** enthält alle Kosten des geplanten Investitionsvorhabens. Dazu gehören die einmaligen Investitionskosten für Soft- und

Hardware, Schulungskosten, externe Beratungskosten, aber auch die laufenden Betriebskosten für z. B. Personal sowie Wartung und Support der Lösung. Das **Nutzenmodell** umfasst die direkten und indirekten Nutzen, die sich aus der Investition ergeben. Dazu zählen: Kosteneinsparungen aufgrund weniger Stammdatenfehler; höherer Umsatz, da Produkte schneller zum Verkauf angeboten werden können; höhere Kundenzufriedenheit. Das **Wirtschaftlichkeitsmodell** führt mit den Einzelergebnissen aus Kosten- und Nutzenmodell die Investitionsberechnung anhand verschiedener Verfahren durch. Anschließend wird eine Risikobetrachtung durchgeführt und daraus eine abschließende Bewertung des Investitionsvorhabens abgeleitet. Das Wirtschaftlichkeitsmodell enthält verschiedene graphische Visualisierungen der Ergebnisse. Dabei wurden aus der Analysephase die in der Praxis relevantesten Investitionsrechnungsverfahren berücksichtigt (siehe Abbildung 2 Hervorhebung).

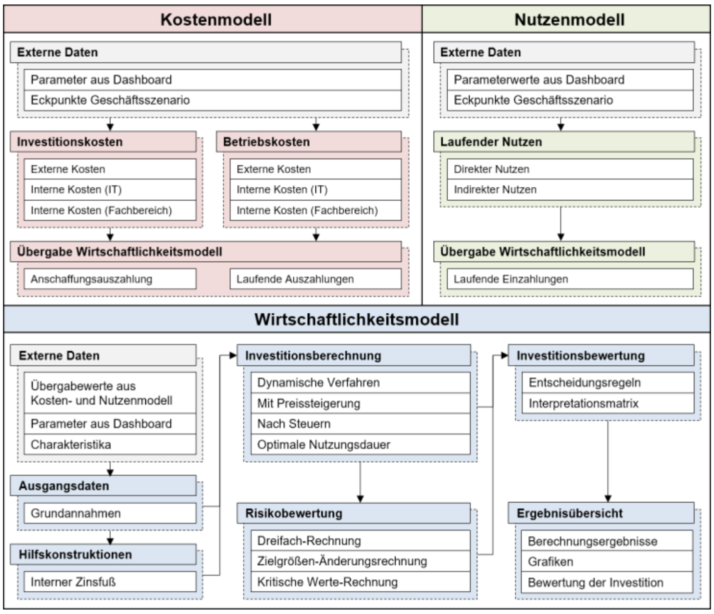


Abbildung 4. Gesamtarchitektur des Analyse- und Berechnungsmodells

Die graphische Visualisierung des Wertbeitrags eines Investitionsvorhabens erfolgt in ALADDIN in Anlehnung an ein wertorientiertes Portfoliomanagement [29] (siehe Abbildung 5). Dabei wird das Investitionsvorhaben der Dimension Profitabilität in Abhängigkeit vom Kapitalwert und der Dimension Wertbeitragsrisiko in Abhängigkeit von der Sensitivitätsanalyse (Dreifach-Rechnung, Zielgrößen-Änderungsrechnung und Kritische-Werte-Rechnung) zugeordnet. Ferner wird das Investitionsvorhaben dreifach und zwar mit Ausgangsdaten, mit Preissteigerung und unter Berücksichtigung der Steuern in das Investitionsportfolio eingezeichnet (siehe Abbildung 5). Somit erkennt das „investitionswillige“ Unternehmen auf einen Blick, ob das Investitionsvorhaben ein Wert-Erzeuger, Wert-Abschmelzer, Wert-Aufholer, Wert-Zerstörer [29] oder eben noch vorteilhaft (neutral) ist.

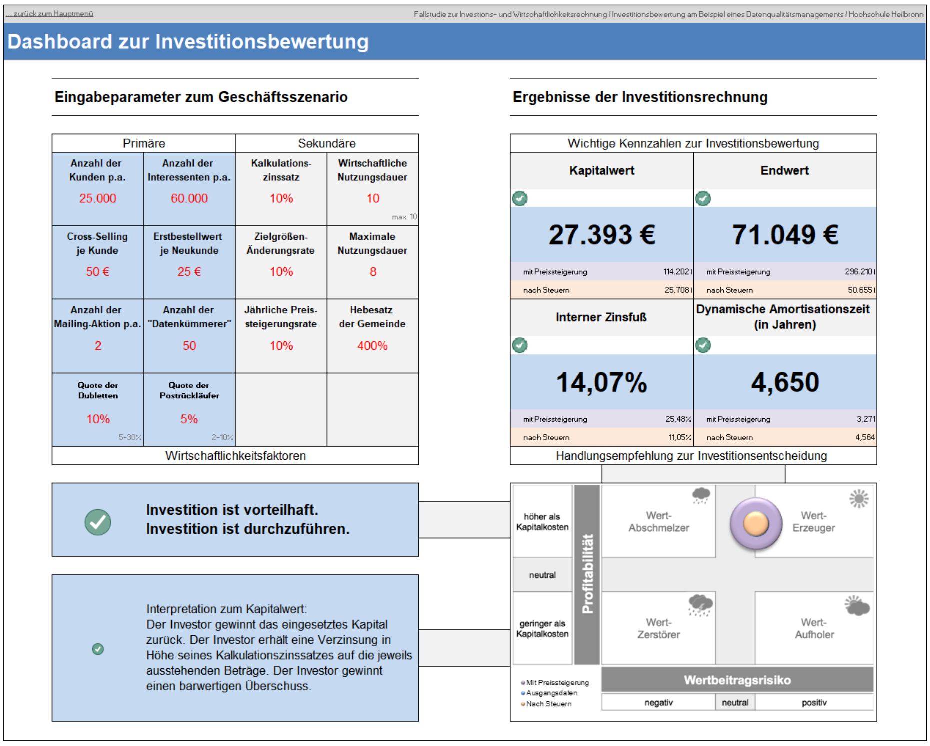


Abbildung 5. Dashboard von ALADDIN zur Investitionsbewertung eines Business Case

4.3 Evaluation der ersten Iteration anhand einer beispielhaften Anwendung

Das Ziel im ersten Evaluationsschritt bestand darin, das Analyse- und Berechnungsmodell (die aktuelle Version von ALADDIN) in Form einer Excel-Implementierung auf die Anwendbarkeit zu überprüfen. Dabei wurde ALADDIN genutzt, um anhand einer beispielhaften DQM-Investition diese zu bewerten. Die Leitfragen zur Evaluation des Modells waren:

- Erfüllt das Instrumentarium die Anforderungen und ist somit eine Investitionsbewertung durchführbar?
- Welcher Erkenntnisgewinn kann daraus abgeleitet werden?

Der beispielhafte Business Case umfasst ein Versandhandelsunternehmen (ein real existierender Fall), das halbjährig die neue Kollektion postalisch per Printkatalog an seine Interessenten und Kunden versendet. Ziel der Kampagne ist es, Neukunden zu einer Erstbestellung und Bestandskunden zum Cross-Selling zu gewinnen. Zur Vermeidung unnötiger Druck-, Versand- bzw. Folgekosten aufgrund schlechter Adressdaten soll der Adressbestand durch ein zu etablierendes unternehmensweites DQM geprüft und bereinigt werden. Die dafür notwenigen Investitionskosten für das DQM sind durch quantifizierbare Nutzenpotentiale auszugleichen.

Zur Simulation und Ermittlung des Wertbeitrags der DQM-Investition wurden zunächst die primären Wirtschaftlichkeitsfaktoren eines Kampagnenprozesses mittels Literaturanalyse und Fallstudie ermittelt. Dabei ergaben sich folgende Faktoren: Adressvolumen; Quoten zu fehlerhaften Adressdaten; Externe/interne Mailing-Kosten; Interne Aufwände aufgrund schlechter Datenqualität; Umsatzzuwächse (Neu und Cross-Selling); Validierungs- und Bereinigungskosten je Datensatz.

Diese Faktoren wurden als parametrisierte Grundannahmen in das Kosten- und Nutzenmodell übertragen. Ferner erfolgte die Festlegung der sekundären Wirtschaftlichkeitsfaktoren (wie bspw. mittelfristige Investition, abruptes Ende, lineare Abschreibung, etc.), die ebenfalls als parametrisierte Grundannahmen in das Wirtschaftlichkeitsmodell übertragen wurden. Abbildung 5 zeigt das Dashboard mit flexibler Parametereingabe zur Simulation und Bewertung dieser DQM-Investition. Ferner ist das Rechenergebnis, der Wertbeitrag im Investitionsportfolio und die automatisierte Investitionsempfehlung enthalten. Die genaue Berechnung mit allen Zahlenwerten kann bei den Autoren angefragt werden.

5 Diskussion und Fazit

Im Zuge der Digitalisierung bilden Stammdaten eine grundlegende Ressource, die nach Zeit,- Kosten- und vor allem Qualitätsgesichtspunkten entsprechend bewirtschaftet werden müssen [6]. Ein etabliertes und gut funktionierendes Datenqualitätsmanagement (DQM) leistet somit einen wichtigen Wertbeitrag zum Unternehmenserfolg. Vielen Unternehmen gelingt es jedoch nicht, dieses Potential vollständig auszuschöpfen, da Investitionen nicht im ausreichenden Maße getätigt werden [3, 15, 17].

Konkurrierende Investitionsvorhaben jeglicher Art fordern eine plausible Kosten-Nutzen-Argumentation, die gerade für eine DQM-Investition Schwierigkeiten bereitet. Als eine mögliche Lösung dieses Problems kann unser Analyse- und Berechnungsmodell ALADDIN eingesetzt werden, das die Unternehmen befähigt, eine gezielte Kosten-Nutzen-Analyse für DQM-Maßnahmen durchzuführen und somit die wirtschaftliche Vorteilhaftigkeit einer DQM-Investition zu ermitteln. Durch den modularen Aufbau und die Parametrisierung von ALADDIN ist eine schnelle Anpassung und flexible Investitionsrechnung eines spezifischen Business Case möglich. Einen wesentlichen Erfolgsfaktor und die zugleich größte Schwierigkeit stellt die Bestimmung der primären Wirtschaftlichkeitsfaktoren dar. Für die Anwendung im Unternehmen sind deshalb alle beteiligten Funktionsbereiche zu involvieren. Weitere Erkenntnisse der Entwicklung und ersten Iteration von ALADDIN sind:

- Verfahren zur Investitionsrechnung werden in der Praxis in unterschiedlichen Intensitäten genutzt. Eine Risikobewertung sowie die Definition eines Business Case sind dabei essentielle Bestandteile einer Investitionsbewertung.
- Datenqualität im Speziellen ist jedoch stark kontextabhängig und erfordert je Kombination aus Geschäftsprozess, Funktionsbereich und Datenobjekt eine spezifische Kosten-Nutzen-Analyse.

- Bislang gibt es wenig grafische Standard-Elemente zur Veranschaulichung der kritischen Kennzahlen einer Investitionsrechnung. Im Zuge der Modellentwicklung wurden zusätzlich grafische Elemente und ein Portfoliomanagement zum besseren Verständnis und der schnelleren Auffassung ausgearbeitet.
- Ein wesentlicher Erfolgsfaktor für eine Investitionsrechnung ist die umfangreiche Analyse von Wirtschaftlichkeitsfaktoren. Hierzu kann die Geschäftsleitung durch entsprechende Anweisungen massiv unterstützen. Ferner liefern Best-Practices gute Ansätze in der Abbildung und Berechnung eines Business Case.

Insgesamt bietet das Analyse- und Berechnungsmodell zur Investitionsbewertung einen weiteren wichtigen Erkenntnisgewinn für ein wertorientiertes DQM und gibt Impulse für zukünftige Forschungsaktivitäten unter Verwendung von interdisziplinären Ansätzen aus Betriebswirtschaftslehre und Informatik. Daraus ergeben sich Leitfragen, die durch weitere Forschungsaktivitäten zu beantworten sind:

- Welche präventiven Maßnahmen reduzieren die Korrektur-Maßnahmen von schlechten Daten und wie lassen sich diese quantifizieren?
- Welchen Wertbeitrag kann Datenqualität für ein Unternehmen leisten? Wie kann der Nutzen von DQM-Investitionen methodisch unterstützt bestimmt werden?
- Welches Maß an Datenqualität muss ein Unternehmen mindestens sicherstellen?

Als kritische Würdigung ist festzuhalten, dass zum jetzigem Zeitpunkt der Entwicklung von ALADDIN noch nicht abschließend bewertet werden kann, ob ALADDIN als Analyse- und Berechnungsmodell zur Investitionsbewertung für ein unternehmensweites Datenqualitätsmanagement vollständig und ausreichend ist. Daher sind in der weiteren Entwicklung von ALADDIN, insbesondere zur finalen Beantwortung der initialen Forschungsfragen, weitere Schritte notwendig. Als nächster Schritt ist die Handhabbarkeit von ALADDIN in der Praxis zu prüfen. Dabei sind weitere Funktionsanforderungen aufzunehmen und zu integrieren. Dies sollte auf Basis von konkreten Anwendungsfällen in Unternehmen verschiedener Branchen erfolgen. Für die Gesamtevaluationsphase ist ein multimethodischer Ansatz vorgesehen, der u. a. Interviews mit Experten auf dem Gebiet der Investitionsrechnungen und umfangreiche Befragungen mit Unternehmen im Sinne der Anwendung von ALADDIN vorsieht.

References

1. Bley, K., Leyh, C., Schäffer, T.: Digitization of German Enterprises in the Production Sector – Do they know how “digitized” they are? In: AMCIS 2016 Proceedings (2016)
2. Mathrani, S., Mathrani, A., Viehland, D.: Using enterprise systems to realize digital business strategies. JEIM 26, 363–386 (2013)
3. Schäffer, T., Leyh, C.: Master Data Quality in the Era of Digitization - Toward Inter-organizational Master Data Quality in Value Networks: A Problem Identification. In: Piazolo, F., Geist, V., Brehm, L., Schmidt, R. (eds.) Innovations in Enterprise Information Systems Management and Engineering (LNBIP, vol. 285), pp. 99–113 (2017)

4. Bloching, B., Leutiger, P., Oltmanns, T., Rossbach, C., Schlick, T., Remane, G., Quick, P., Shafranyuk, O.: Die digitale Transformation der Industrie. München (2015)
5. Leyh, C., Bley, K.: Digitalisierung. Chance oder Risiko für den deutschen Mittelstand? – Eine Studie ausgewählter Unternehmen. HMD 53, 29–41 (2016)
6. Otto, B., Österle, H.: Corporate Data Quality. Voraussetzung erfolgreicher Geschäftsmodelle. Springer, Heidelberg (2016)
7. Dalmolen, S., Moonen, H., van Hillegersberg, J.: Industry-wide Inter-organizational Systems and Data Quality. In: AMCIS 2015 Proceedings (2015)
8. de Corbière, F.: Interorganizational Information Systems and Data Quality Improvement. In: ICIQ 2007 Proceedings (2007)
9. Hüner, K.M., Schierning, A., Otto, B., Österle, H.: Product Data Quality in Supply Chains: The Case of Beiersdorf. EM 21, 141–154 (2011)
10. Legner, C., Schemm, J.: Toward the Inter-organizational Product Information Supply Chain. JAIS 9, 119–150 (2008)
11. Loshin, D.: Master Data Management. Morgan Kaufmann, Amsterdam (2008)
12. Schäffer, T. and Stelzer, D.: Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing. In: WI 2017 Proceedings (2017)
13. Ebner, V., Brauer, B.: Fallstudie zum Führungssystem für Stammdatenqualität bei der Bayer CropScience AG. HMD 48, 64–73 (2011)
14. Helmis, S., Hollmann, R.: Webbasierte Datenintegration. Vieweg+Teubner, Wiesbaden (2009)
15. Schäffer, T., Beckmann, H.: Stammdatenqualität 2016. Steinbeis-Edition, Stuttgart (2016)
16. Zillmann, M.: Revival der Stammdaten. Lünendock GmbH, Mindelheim (2016)
17. Schäffer, T., Beckmann, H.: Stammdatenqualität 2013. Steinbeis-Edition, Stuttgart (2014)
18. Wang, R.Y., Strong, D.M.: Beyond Accuracy: What Data Quality Means to Data Consumers. JMIS 12, 5–33 (1996)
19. Grabe, J., Däumler, K.-D.: Grundlagen der Investitions- und Wirtschaftlichkeitsrechnung. NWB, Herne (2014)
20. Brugger, R.: Der IT Business Case. Kosten erfassen und analysieren Nutzen erkennen und quantifizieren Wirtschaftlichkeit nachweisen und realisieren. Springer, Heidelberg (2005)
21. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75–105 (2004)
22. Hevner, A.R.: A Three Cycle View of Design Science Research. SJIS 19, 87–92 (2007)
23. Becker, J., Knackstedt, R., Pöppelbuß, J.: Entwicklung von Reifegradmodellen für das IT-Management. WI 51, 249–260 (2009)
24. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, xiii–xxiii (2002)
25. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: ECIS 2009 Proceedings (2009)
26. Eppler, M.J., Helfert, M.: A Classifikation and Analysis of Data Quality Costs. In: ICIQ 2004 Proceedings (2004)
27. Däumler, K.-D., Grabe, J.: Anwendung von Investitionsrechnungsverfahren in der Praxis. NWB, Herne (2010)
28. Zechmann, A.: Assessing the Economic Value of Data Assets. Working Report, Universität St. Gallen, St. Gallen (2016)
29. Dillerup, R., Stoi, R.: Unternehmensführung. Vahlen, München (2013)

Der Kunde in der Digitalen Transformation – Creating Customer Values

Teilkonferenzleitung

Christoph Lattemann
Susanne Robra-Bissantz

Design Principles for Co-Creating Digital Customer Experience in High Street Retail

Jan H. Betzing[1], Daniel Beverungen[2], and Jörg Becker[1]

[1] University of Muenster, ERCIS, Münster, Germany
{jan.betzing,joerg.becker}@ercis.uni-muenster.de
[2] University of Paderborn, Business Information Systems, Paderborn, Germany
daniel.beverungen@uni-paderborn.de

Abstract. While customers increasingly embrace online shopping, many retailers in high streets struggle to attract profitable customers to their stores and retain them. Establishing digital customer experiences may increase customers' value perceptions, improving the competitive position of high street retailers. Customer experience creation in retail is a multi-faceted construct that, amongst others, depends on the service interface, atmosphere, assortment, price, past experiences, and the social environment in which retail service is co-created while customers shop. However, extant customer experience theory insufficiently accounts for the transformative power of recent mobile technology that enables digital and contextual service. In a conceptual approach, we develop eight propositions to frame *digital* customer experience, enabled by mobile technologies. In line with these propositions, we propose eight design principles that enable and constrain IT artifacts for co-creating digital customer experience.

Keywords: Digital Customer Experience, High Street Retail, Context-aware Service, Location-based Service, Design Principles

The full version of this paper can be found on page 2083.

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Nachhaltiger IKT-Konsum durch Sharing Economy? Eine multimethodische Analyse

Rikka Wittstock[1], Danielle Warnecke[1] und Frank Teuteberg[1]

[1] Universität Osnabrück, Fachgebiet Unternehmensrechnung und Wirtschaftsinformatik, Osnabrück, Deutschland
{rikka.wittstock, danielle.warnecke, frank.teuteberg}@uni-osnabrueck.de

Abstract. In diesem Beitrag wird mittels einer multimethodischen Analyse untersucht, ob und inwieweit die Sharing Economy ein geeignetes Konzept zur Reduzierung der Umweltauswirkungen des Konsums mobiler Informations- und Kommunikationstechnologie (IKT) darstellen kann. Hierzu wird zunächst anhand einer systematischen Literaturanalyse eruiert, durch welche Mechanismen die gemeinschaftliche Nutzung von Gebrauchsgütern zur Reduzierung der Umweltbelastung beiträgt. Die Bereitschaft der Konsumenten zur gemeinschaftlichen Nutzung mobiler IKT wird durch eine quantitative Erhebung (n=329) ermittelt, während die Marktpotenziale Sharing Economy basierter Geschäftsmodelle durch Expertenbefragung erhoben werden. Der Beitrag kommt zu dem Schluss, dass sich Sharing Economy Geschäftsmodelle insbesondere für hochpreisige, kurzzeitig genutzte mobile IKT eignen.

Keywords: Sharing Economy, Nachhaltigkeit, Informations- und Kommunikationstechnologie, Multimethodenansatz

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

“Wat de Buur nich kennt, dat frett he nich!“ – Partizipation der Bevölkerung am Digitalisierungsprozess der Nahversorgung im ländlichen Raum

Pascal Meier[1], Jan Heinrich Beinke[1] und Frank Teuteberg[1]

[1] Universität Osnabrück, Unternehmensrechnung und Wirtschaftsinformatik, Osnabrück, Deutschland
{pascal.meier,jan.beinke,frank.teuteberg}@uni-osnabrueck.de

Abstract. Der demografische Wandel und dessen Auswirkungen stellen insbesondere ländlich strukturierte Regionen vor große Herausforderungen. Durch das oft attraktivere Lebensumfeld ziehen jüngere Menschen weiterhin verstärkt in die Städte. Das Aufrechterhalten der lokalen Versorgung und Infrastruktur gestaltet sich für die Kommunen zunehmend schwerer. Durch den Bevölkerungsrückgang in ländlichen, strukturschwachen Regionen geraten diese in eine Abwärtsspirale: Die Rentabilität privatwirtschaftlicher Angebote wie beispielsweise Dorfläden sinkt durch die sinkende Kaufkraft, die mit einer rückläufigen Bevölkerungszahl einhergeht. Dadurch folgt oft die (altersbedingte) Aufgabe der Dorfläden. Die Suche nach Nachfolgern gestaltet sich zumeist sehr schwierig. In diesem Beitrag wird ein Dorfladenkonzept vorgestellt, indem der Dorfladen als zentraler (Social-)Hub fungiert. Des Weiteren werden Leitfragen für einen partizipativen Digitalisierungsprozess bereitgestellt. Die partizipativ entwickelten Artefakte werden im Rahmen einer multi-methodischen Fallstudie in einer der ältesten Gemeinden in Westniedersachsen im Zuge des vom BMBF-geförderten Demografieprojekts „Dorfgemeinschaft 2.0“ veranschaulicht und im weiteren Verlauf des Projekts erprobt.

Keywords: Nahversorgung im ländlichen Raum, Digitale Transformation, Partizipation im Digitalisierungsprozess

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Towards an Omni-Channel Framework for SME Sales and Service in the B2B Telecommunications Industry

Robert Heidekrüger[1], Markus Heuchert[1], Nico Clever[1], and Jörg Becker[1]

[1] University of Münster – ERCIS, Münster, Germany
{firstname.lastname}@ercis.uni-muenster.de

Abstract. The use of mobile technology has become self-evident both for private and professional use. With business customers expecting a shopping and service experience similar to the usage in private, Business-to-Business (B2B) providers must come up with new solutions. Omni-Channel management is one trend that is increasingly gaining attraction. To this end, a profound understanding of influencing factors for sales and service to small and medium enterprises (SME) is missing. Telecommunications, an industry of major importance, serves as target domain for this research endeavor that presents an omni-channel framework for SME Sales and Service. Through collaboration with one of the three largest German providers, 22 interviews were conducted and serve as the data basis.

Keywords: Omni-Channel, SME, B2B, Telecommunications.

1 Introduction

The move towards a digital society in face of the ongoing digitalization of the world impacts both people's private and professional lives [1, 2]. Nowadays, the use of (mobile) technology has become self-evident, especially but not only for the younger generations, with 52.7 % of the world population accessing the Internet from their mobile phone in 2015 and a predicted penetration of 63.4 % for 2019 [3]. In light of this development, it can be observed that people adopt behaviors and expectations from their private lives and apply them to their work lives, equally.

Consequently, business customers increasingly expect shopping and service experiences similar to those they make as private consumers when procuring in the business context [4]. To respond to such changing expectations, Business-to-Business (B2B) retailers and service providers have to come up with new solutions and offers in order to excel through competitive advantage and service excellence. In this context, the concept of omni-channel (OC) management, in which information is integrated between different touchpoints and channels for a superior customer experience, gains attention and its use in B2B environments is increasingly becoming popular [5–7].

Business clients of telecommunications providers make use of numerous touchpoints throughout their customer journey. They do so to become customers in

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

the first place, procure additional products, resolve incidents, or eventually end or renew their contract. With mobile telecommunication being a commodity nowadays, providers need to extend their value proposition by additional products and services [8]. To this end, offering a seamless customer experience across channels to customers is a promising vision to facilitate superior purchasing and service encounters in this competitive market.

Yet, a common coherent understanding of the concept of B2B OC is missing in academic as well as in practical managerial literature. It remains unclear how firms can use it in order to improve their customers' journeys across different channels and touchpoints. Therefore, the objective of this work is twofold. On the one hand, the goal is to present the potential of OC environments in B2B environments in general. On the other hand, we aim to develop an OC framework for sales and service that enables telecommunications providers to offer a seamless customer experience to their Small and Medium Enterprise (SME) customers across all touchpoints and channels. In this, we will argue from the customer's point of view since OC, in our understanding, is a customer-centric approach.

The remainder of the paper is structured as follows: In section 2, the research background is laid out. In section 3, the research methodology to create the OC framework is described. Consequently, in section 4, the design of the aforementioned framework is explained. In section 5, the work at hand is discussed and a brief outlook to further research is given.

2 Research Background

This work draws on existing research in B2B environments, as well as the OC approach and customer experience management. The aim of this chapter is to introduce the relevant concepts, define the scope and clarify their interconnection later materialized in the framework design.

B2B Environments in Telecommunications. While private consumers perform transactions in Business-to-Consumer (B2C) environments, B2B exchanges exclusively occur between corporate entities. One peculiarity in this context are the so-called buying centers. The purchasing business typically sets up a situation-specific, interdisciplinary team of different hierarchical levels for reducing the risk of a purchasing decision of an organization [9]. Being more complex, of higher volume and even higher profitability in nature, B2B markets are featured by a higher degree of customization on the selling side. From this point on, we will use business customers interchangeably with customers. Typically, sales representatives (SR) enable direct sales through on-site meetings, along with other forms that enable the selling directly to the customer without an intermediary [10]. Indirect sales characterized by legally and economically independent firms that act as resellers play a subordinate role in B2B environments, as the personal relationship created between SRs and their customers facilitates tailored solutions to the customers' needs.

Furthermore, a highly customer-oriented service is enabled, normally realized through an internal sales support without customer contact that assists the SR in the field [11].

Omni-Channel Management. The proliferation of different means of contact between company and customer has led to challenges in managing channels. OC represents the latest development in this regard, where all communication between the two parties is linked, so that the customer has a superior experience regardless of the used channel [12]. The interface that the company provides to interact with the customer is termed touchpoint (TP) and here limited to company-owned TPs.
Straker et al. [13] provide a classification of channels that distinguishes *traditional* and *digital channels*. The former are comprised of primarily physical and non-digital ways of interaction, while e-mail will also be considered as a traditional channel due to its self-evident omnipresence. The defining characteristic of digital channels are the reliance on the internet as underlying technology.

Further, four key typologies of digital channels can be named [13]. *Functional* channels serve varying purposes (interaction, diversion, functional) in bi- or unidirectional communication for informative, supportive, promotional or revenue-generating content, while *social* channels are driven by the bidirectional interaction among the users and with the company and feature informative and promotional content. *Community* channels are run by users and used for diversion, i.e. participating in recreational activities. *Corporate* channels are unidirectional and have functional purpose seen in the provision of information.

Customer Experience Management. Customer experience has recently become a major buzzword in Marketing and strategic priority [14]. While being subjective in nature, companies seek to understand the customer's journey before, during and after a purchase to optimize the customer experience by utilizing a customer-centric view. Customers nowadays tend to use more than one channel in their journey, so that a clear connection to OC management becomes clear: The holistic view of customer experience management is mirrored in the postulated channel integration of OC management for a seamless and consistent experience. As the journey proceeds, different organizational units (e.g. sales or after-sales service personnel) are *touchpoints* to the customer, which additionally need to follow the one-face-to-the-customer paradigm. Concluding, the customer experience is influenced vertically (through different channels) and horizontally (over time through different touchpoints).

3 Research Methodology

To develop an OC framework for sales and service in the telecommunications industry, a design-oriented approach is taken. The framework—the artifact to create—can be classified as a model with respect to the terminology proposed by Hevner et al. [15]. For its development, the Design Science Research Methodology (DSRM) proposed by [16] is applied. The DSRM consists of six phases. The first phase—problem identification and motivation—and the second phase—define the objectives for the solution—have already been described in the introduction. In the next section, an analysis of customer communication channels as a basis for the development of the OC framework will be delineated. During this phase, 22 expert interviews were carried out in close collaboration with one of the three major telecommunications providers in Germany to elaborate the requirements for and the contents of the OC framework. Each interview took 30 to 60 minutes and the predefined questions were sorted into the corresponding phase of the defined selling process or customer journey at the provider, depending on the interviewee. This approach guaranteed coverage of all phases. The third phase—design and development—will be described in the fifth section. The demonstration—the fourth phase of the DSRM—will be described afterwards. Here, an approach of argumentative demonstration based on use cases is applied. The fifth phase—evaluation—could not yet be carried out and will be part of future research. The sixth phase—communication—is accomplished by this work and future publications, likewise.

4 Customer Communication Channel Analysis

In order to develop an OC framework for SME sales and service in the telecommunications industry, the relevant channels making up the OC customer experience have to be identified first. The interviews were mainly conducted to this end. Of the 22 interviewees, 17 can be identified as internal experts stemming from SME sales people, sales managers, customer service managers, product and software experts. The remaining five interviewees are customers which are at the center of attention when trying to improve the customer experience through a seamless OC offer.

During the analysis of the interviews, two major channel groups could be identified on the basis of the work on digital channel design by Straker et al. [13] as described in the research background: traditional channels and new, digital channels. In this, e-mail, although being digital by nature, will be regarded as a traditional channel due to its status as more and more legally accepted medium of communication. Furthermore, some of the digital channels were disregarded due to their low relevance in B2B environments, such as Pinterest or Flickr. In addition to the typology, some other channels were identified. These are mobile instant messaging (MIM), video conferencing, chat bots and professional corporate social platforms like Xing or LinkedIn. The full list of identified channels for customers, can be differentiated into traditional and digital channels. Traditional channels are (1) Face-to-Face and Point of Sale, (2) Phone, (3) Paper mail and fax, (4) E-Mail, which are described in the following along with digital channels shown in Table 1.

Face-to-face communication is essential, especially in the B2B context and when it comes to more complex issues and products. The interviews have shown that both internal experts as well as SME customers make a strong argument for the necessity of face-to-face appointments when dealing with complex topics or negotiations, e.g.

fixed net infrastructure. As for the *point of sale*, it is less common to consult branch shops for B2B issues. However, branch stores should be capable of providing all business customers with basic services like SIM card replacements or booking of standard tariff options in order to act as “one face to the customer”.

As it was found during the interviews, the *phone*—along with e-mail—is the most frequently used channel for customers and sales representatives to interact. In order to provide customers with an OC experience, each contact via phone has to be documented properly. Regarding hotlines, customers of telecommunications providers are currently faced with a great number of different phone numbers for different requests. Although it might make sense to have separate phone numbers for several reasons, the number of hotlines that are communicated to customers should be kept as low as possible in order to provide the most comfortable customer experience. Another finding of the interviews is that *paper mail and fax* are increasingly replaced by e-mail as a legally accepted way of business communication. However, an abolishment of paper mail and fax is not yet possible due to more traditional customers. Similar to the phone, *e-mail* is a very frequently used channel for both outbound and inbound communication of telecommunications providers. This trend was also supported by both internal and external interviewees.

Table 1. Identified digital customer touchpoints

Touchpoint group	*Customer touchpoint*
Functional	Self-Service Platform
	Mobile Apps
	Live Chats and Chat Bots
	Website, Web Enquiries, Newsletters
	Video Conferencing
Social	Social Networks (Facebook, Google+, Twitter)
	Corporate Social Platforms (LinkedIn, Xing)
	Mobile Instant Messaging
Community	Forums
	YouTube
Corporate	FAQ
	Digital Feedback Forms

Functional channels. In the interview analysis, a *self-service platform* was identified as the main application that is offered to customers. As mobile applications are considered separately, this channel will further be referred to as self-service platform within the framework that is being designed. While such platforms give customers the chance to handle several administrative and largely standardized tasks like contract extensions and SIM card exchanges on a client basis, the internal interviewees see a lot of potential for improvement and further development of these.

Mobile apps have become a self-evident part of people’s everyday life. In business environments—especially in large companies—the functionality of such apps is usually restricted to a certain degree. Here, the transfer of functionality from the existing self-service platforms to the mobile apps along with a corresponding role and rights management could enhance the customer experience in a B2B context.

Chats are increasingly present on all kinds of websites [7]. Chats can be differentiated in live chats with real service agents and virtual assistants or chat bots, which are powered by AI [17]. Either way, chats are suited for rather simple requests or quick questions about a product or service. Interview partners have stated that they prefer talking to someone on the telephone when it comes to more complex issues or use e-mails for documentation reasons. However, it should be part of an OC strategy to offer this additional, situation-sensitive channel to account for simple questions.

Websites are often the first touchpoint for potential customers. It is, therefore, of major importance that the information displayed on the website is consistent with the information available in all other channels, which is a basic principle of OC management. In B2B telecommunications, offers for customers are usually negotiated individually using master agreements. As a consequence, the websites of the major providers are held informative with less focus on specific prices and rather suggesting personal contact for individual quotes. Accordingly, websites should be individualized for different customer segments, e.g. self-employed people, SME customers, large corporations, and public authorities.

As an alternative to on-site face-to-face meetings, *video conferencing* is a communication channel that facilitates personal interaction. Instead of depending on third-party software, proprietary video conferencing tools or unified contact center solutions of the telecommunications provider can be used for example to bring together a customer, a sales representative and a technical expert on a complex fixed net project. However, video conferencing in general is not suited to be used with any touchpoint (e.g. hotline personnel), but rather with those where a personal relationship should be facilitated, e.g. with the dedicated SRs.

Social channels. The usage of *social networks* is not limited to the private lives of people anymore as they increasingly gain relevance in the B2B context [18]. Despite corporate restrictions that might hinder people from using social networks on behalf of their employer, people are used to these networks from their private experience. Similar to web chats, social media channels are more suited for simple standard requests, e.g. regarding tariff options or roaming. Nevertheless, social media channels can also be used in a simplex fashion (i.e. one-way communication) to announce news while focusing less on the interaction with the network. This can be easily followed by experts on the customer side, potentially generating up- and cross-selling potential

Corporate social platforms such as LinkedIn or Xing, are specifically made for the business context. Although the platforms are primarily used for recruiting purposes, they can also serve as a communication channel between telecommunications providers and their (potential) customers. During the interviews, one SR stated, for example, that he regularly employs Xing as a channel for social selling.

Seeing the proliferation of *mobile instant messaging* with providers like WhatsApp, such services increasingly gain importance in internal business communication, as it was also mentioned by some interviewees. Yet, using such third-party services as a means of communication caused mixed reactions among interviewees and are subject to privacy and security concerns. Here, to still be able to offer such channels to business customers if requested, proprietary MIM solutions could help ensuring privacy, security as well as proper documentation of the customer contact as part of an OC approach. However, a major downside of such solutions is the missing adoption among the customers.

Community channels. *Online forums* are a common platform that is offered to customers by all telecommunications providers where customers get to help each other and be assisted by service agents publicly. In the business context, such platforms are usually not offered. Similar to chats and social channels forums are suited for minor requests or simple questions, or to pool requests by multiple customers, which is often done in the B2C context. Forums offer a way for providers to communicate answers and solutions to common problems to a larger audience and make these solutions available in the long-term. However, according to the analysis, such platforms are less suited for B2B environments where requests are usually more specific and individual. In contrast to social channels that provide additional value when used as outbound channels for information and news, forums offer less added-value for providers in the business context so that they will probably not pose as an as important channel to offer to B2B customers as part of an OC approach.

As with the social networks before, it is beneficial for telecommunications providers to be present on common *online video platforms* with YouTube as the most prominent example to fulfil the expectations that consumers project into their professional lives.

Corporate channels. Offering *Frequently Asked Questions* (FAQ) for basic topics and promoting them to customers can help to prevent unnecessary requests by providing a way to solve questions themselves. The contents of the FAQ should be based on common customer enquiries and regularly be updated to improve them gradually. In the context of OC, the consistency of information is essential. It is, therefore, beneficial to limit the FAQ to questions applicable to all customer segments or clearly distinguish the corresponding customer segment in order to avoid confusion.

Providing customers with a means to give *feedback* is an essential part when it comes to improving the customer experience and is therefore crucial in an OC context. Making use of available customer contact data to request feedback after an interaction, e.g. by sending a SMS or an e-mail following a call to the hotline, is a convenient way to collect recent feedback. In the case of negative experiences, follow-up calls can be made to unsatisfied customers in order to mitigate damage and prevent customer churn. Apart from that, the possibility to provide feedback gives customers the feeling of being heard and taken seriously if it is accordingly answered and taken into account.

5 Design of the Omni-Channel Framework

From the analysis of traditional and digital customer touchpoints in the previous section, the structure for the OC design can be derived, i.e. the channels that should be offered by a telecommunications provider to facilitate an OC approach. Based on the channel design, implications for the organizational and IT infrastructures of B2B telecommunications providers can be derived. The resulting framework is depicted in Figure 1.

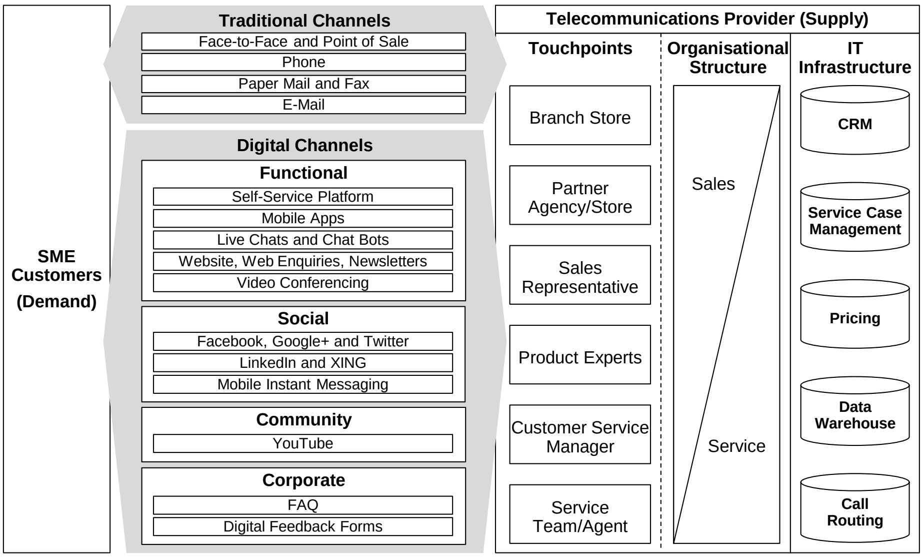


Figure 1. Omni-channel framework for B2B telecommunications providers

On the left-hand side of the framework, the customers are placed, since in the customer-centric OC approach the customers pose the starting point of analysis. Next to the customers, the previously elaborated channels are placed to illustrate their communication purpose. On the right-hand side, the telecommunications providers are located, divided into three sub-parts. The leftmost of these parts are the touchpoints which customers can reach via the different channels. Next to the touchpoints, the internal organizational structure of the provider is depicted. The dashed line is chosen to illustrate that the touchpoints are embedded into the organizational structure, but the structure is not directly visible and important to the customers. As the focus of this work lies on sales and service, other remaining departments are not part of the framework. The rightmost part shows the IT infrastructure. Similar to the organizational structure, only such systems and tools that were identified as relevant for OC and in the scope of this work are included.

SME customers. The customer side is not further specified in detail as there are no insights into the individual organizational structure and it should represent any SME customer a telecommunications provider might serve. Any part of the customer organization can turn to the provider's touchpoints.

Telecommunications provider. On the provider side, channels are used by several touchpoints, which is illustrated by both traditional and digital channels pointing at all touchpoints. Naturally, not all touchpoints use every channel so that a corresponding customer touchpoint management is required that regulates which channels are used by each individual touchpoint. As channels will be used equally by both sales and service, all touchpoints need to have access to the same data bases and software systems in order for OC to work. Correspondingly, there are implications for the organizational and IT infrastructures of telecommunications providers.

5.1 Implications for the Organizational Structure of Telecommunications Providers

With regard to the organizational structure, the traditional split of the departments sales and service into separate silos, which is common in many companies, needs to vanish and closer collaboration is required. This convergence is illustrated in the framework by the diagonal division into sales and service which stands for one integrated unit. It has to be noted that this does not mean that the creation of one enormous department under a new name is required, but rather calls for the necessity of holistic collaboration and a common understanding of customers. This involves the same segmentation of customers, the common access to all relevant customer data, no matter whether sales- or service-related, as well as seamless internal communication processes and tools in order to be able to offer an OC experience to customers. By taking a customer-centric perspective through a customer journey approach, the potential of such convergence can be laid out: As sales activities happen before after-sales service activities, the customer experience is fostered horizontally by making use of information collected in prior phases of the customer journey.

5.2 Implications for the IT Infrastructure of Telecommunications Providers

The importance of the IT infrastructure as a prerequisite for an OC strategy was highlighted before. As OC is a customer-centric approach that improves the customer's experience by facilitating seamless switching between channels and touchpoints, customer data is the central asset that has to be available across all touchpoints the customer can reach out to. Therefore, a Customer Relationship Management (CRM) system is essential for the management of customer data and represents the core system in the OC framework. However, no standard CRM solution is capable of offering all functionalities that are required by specific industries, e.g. telecommunications. Therefore, such systems either have to be customized or integrated with additional systems. Along with rather static master data like a customer's company name, address, or industry, there is transaction data, e.g. current customer requests. Especially this type of customer service-related data may not always be stored in the same software system as customer master data. However, in order to realize OC properly, the data should either be stored in the same system or, alternatively, in different fully integrated systems in order to ensure access to all data from each system. Especially for service departments, a system for service case management is relevant. In order to ensure that cases are handled quickly and by the right people, such a tool should work as a workflow management tool that implements clearly defined business processes based on the ideally integrated organizational structure of sales and service that was discussed before.

Moreover, a common tool for pricing is necessary to guarantee price consistency across touchpoints and enable both sales and service employees to give transparent information on individual conditions for each customer. Finally, although the common system environment should enable all employees to provide the same information, redirection of requests and cases on the phone will be necessary at times so that a state-of-the-art call routing system needs to be in place to connect customers seamlessly to other touchpoints without them having to take action, either triggered by service agents or by Interactive Voice Response systems. As mentioned before, the rules for redirection have to be clearly defined by business processes and should be supported by a workflow management tool that allows the assignment of tasks to other teams or employees, for instance, ideally in the CRM system itself.

6 Demonstration of the Omni-Channel Framework

In order to demonstrate the feasibility of the developed OC framework, a fictive use case is presented at this point. However, this use case was mentioned as a huge potential of such an approach in many of the interviews that were lead during the analysis phase. The use case was mentioned before and is presented as a *persistent shopping cart across touchpoints*.

Usually, a shopping cart in an online shop is limited to the website or at best also accessible using a mobile app. How storing the customer's shopping cart centrally can help improve both the customer experience as well as increase sales as part of an OC strategy, will be explained in the following.

In case the customer has doubts about tariffs, options or devices that are not easy to self-solve online, there should be an option to directly contact an agent using a chat, e-mail or phone. By automatically including the customer number in the chat or e-mail or automatically retrieving it when they call from their number recognizable by their provider, the customer can automatically be routed to a service agent who has all information on the shopping cart available. By giving the agents write-access to the customer's shopping cart on demand, the customer can be assisted in their choice of products and be helped quickly, e.g. by adjusting options or changing tariffs in the cart. These interactions with the customer can also be used as an opportunity for up- and cross-selling by the provider, especially when they contact their SR with questions, who should equally have full access to the customers' shopping carts.

Similarly, all quotes made by SRs or any other touchpoints should be available for customers to look at, including all individual conditions within the self-service platform. By storing this data centrally and accessible for all touchpoints, it is guaranteed that the customer can contact any touchpoint with questions and all touchpoints can take the information into account when advising customers or making their own quotes. It also helps to avoid that offers are made by other touchpoints, when SRs have already made one and thereby prevent customers from making confusing experiences that may deteriorate satisfaction.

Most importantly, it has to be ensured that the self-service solution reflects the exact pricing conditions that the customer would get at any other touchpoint so that the customer does not refrain from using it in the future. For example, sometimes price reductions are given for online orders as part of promotion in order to shift more orders to online channels.

Apart from the case when customers actively contact a provider with a request concerning their shopping cart, they can also leverage the information from their customers' online behaviors proactively, e.g. by sending notifications upon open, not ordered shopping carts, individualized product information or newsletters to corresponding customers.

7 Discussion

In this article, the development of an OC framework for sales and service in the B2B telecommunications industry on the basis of the Design Science Research Methodology was described. While the need for such a framework was shown in the motivation, the lack of it was emphasized in the research background. During the development, 22 expert interviews with both internal sales experts and external customers were carried out to verify, extend and abolish the typology of communication channels—both traditional and digital—by Straker et al. [13]. On the basis of four traditional channels and twelve digital channels—grouped into functional, social, community, and corporate channels—the OC framework was derived afterwards showing the way of communication along with the respective internal touchpoints, the internal organizational structure as well as the IT infrastructure needed for such an OC approach. The feasibility of the framework was demonstrated on the basis of a fictive, yet in the interviews often-mentioned use case.

Up to now, the evaluation of the OC framework for telecommunications providers in practice is still missing. This evaluation will be carried out in close collaboration with one of the three major German telecommunications providers, which has already provided access to the internal experts and the customers for the expert interviews in this work. Moreover, the framework and the implications for both the organizational and IT infrastructures should be verified and/or adapted by testing them out in a bigger context, probably not limited to the German telecommunications market.

References

1. Evangelista, R., Guerrieri, P., Meliciani, V.: The economic impact of digital technologies in Europe. Econ. Innov. New Technol. 23, 802–824 (2014)
2. Hagberg, J., Sundstrom, M., Egels-Zandén, N.: The digitalization of retailing: an exploratory framework. Int. J. Retail Distrib. Manag. 44, 694–712 (2016)
3. EMarketer: Mobile phone internet user penetration worldwide from 2014 to 2019, https://www.statista.com/statistics/284202/mobile-phone-internet-user-penetration-worldwide/ (Accessed: 29.09.2017)
4. Forrester Consulting: Building The B2B Omni- Channel Commerce Platform Of The Future B2B Buyer Expectations Are Driving Sellers To Deliver Fully Functional Omni-Channel Experiences. Technical report (2014)
5. Forbes: For B2B, Omnichannel is not just a buzzword, it means business, https://www.forbes.com/sites/brianwalker/2014/12/11/for-b2b-omnichannel-is-not-just-a-buzzword-it-means-business/ (Accessed: 29.09.2017)
6. Forrester Consulting: Mastering Omnichannel B2B Customer Engagement. Technical report (2015)
7. Intershop: Taking the Fast Track into the Digital Future of B2B Commerce. Technical report (2016)
8. Klein, A., Jakopin, N.: Consumers' willingness-to-pay for mobile telecommunication service bundles. Telemat. Informatics. 31, 410–421 (2014)
9. Kleinaltenkamp, M., Saab, S.: Technischer Vertrieb: Eine praxisorientierte Einführung in das Business-to-Business-Marketing. Springer, Heidelberg (2009)
10. Detroy, E.N., Behle, C., Vom Hofe, R.: Handbuch Vertriebsmanagement. mi-Fachverlag, München (2009)
11. Bürli, R., Friebe, P.: Distribution: Grundlagen mit zahlreichen Beispielen, Repetitionsfragen mit Lösungen und Glossar. Compendio Bildungsmedien, Zürich (2010)
12. Verhoef, P.C., Kannan, P.K., Inman, J.J.: From Multi-Channel Retailing to Omni-Channel Retailing. Introduction to the Special Issue on Multi-Channel Retailing. J. Retail. 91, 174–181 (2015)
13. Straker, K., Wrigley, C., Rosemann, M.: Typologies and touchpoints: designing multi-channel digital strategies. J. Res. Interact. Mark. 9, 110–128 (2015)
14. Lemon, K.N., Verhoef, P.C.: Understanding Customer Experience and the Customer Journey. J. Mark. 80, 1–62 (2016)
15. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q. 28, 75–105 (2004)
16. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. J. Manag. Inf. Syst. 24, 45–77 (2007)
17. Kreutzer, R.T., Land, K.-H.: Digitale Markenführung. Springer Fachmedien Wiesbaden, Wiesbaden (2017)
18. Forbes: The Real Truth About B2B Marketing And Social Media, https://www.forbes.com/sites/ajagrawal/2016/10/03/the-real-truth-about-b2b-marketing-and-social-media/ (Accessed: 29.09.2017)

Value Co-creation Ontology—A Service-dominant Logic Perspective

Michael Blaschke[1], M. Kazem Haki[1], Stephan Aier[1], and Robert Winter[1]

[1] University of St.Gallen, Institute of Information Management, St.Gallen, Switzerland
{michael.blaschke,kazem.haki,stephan.aier,robert.winter}@unisg.ch

Abstract. Marketing research apprises scholars in different disciplines of a paradigmatic reorientation from a traditional goods-dominant (G-D) to a service-dominant (S-D) logic. S-D logic re-conceptualizes the notion of economic exchange. The cornerstone of this reorientation is the concept of *value co-creation*—a collaborative process of reciprocal value creation among various actors. Owing to S-D logic's significance, information systems (IS) research discusses its prospective implications on core elements of the IS knowledge base. However, an equivocal understanding of value co-creation's foundations, semantics, and use emphasizes its underlying theoretical ambiguity in IS and marketing research. Through employing *Methontology*, a well-structured methodology to build ontologies, we develop a value co-creation ontology for IS from an S-D logic perspective. The developed ontology not only offers a multidisciplinary glossary of value co-creation's constituent concepts, but also thoroughly depicts their relationships. The resultant ontology represents a first step toward reflecting S-D logic in IS analysis and design.

Keywords: Value Co-creation, Service-dominant Logic, Ontology, Information Systems Analysis and Design.

1 Introduction

The notion of economic exchange has been notably affected by the paradigm of service-dominant (S-D) logic. Seminal studies [1–3] shed light on the emergence of a paradigm shift from a goods-dominant (G-D) logic to S-D logic. Central to S-D logic is the proposition that all social and economic actors are co-creators of value—that is, *resources* do not have value per se; rather value is created *jointly* with other actors when resources are used [1, 2]. This moves the locus of value creation from *exchange* to *use* [2]. This *usage* determination—as opposed to an *exchange* determination of G-D logic—emphasizes that value unfolds only as soon as actors initiate a *value co-creation* process in that they *start* using goods (products or services) and reciprocally integrate resources with goods providers during usage time [4–7]. Therefore, in S-D logic, value is determined by the quality of a *value-in-use* experience and not just by the quality of goods' *value-in-exchange* [8, 9]. For instance, in the airline industry, jet turbine manufacturers used to follow G-D logic by selling turbines to airlines. However, since

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

airlines do not create value by owning turbines, but rather by the realization of airtime, manufacturers nowadays sell airtime to airlines instead of jet turbines [e.g., 10–12].

The information systems (IS) community has started investigating IS phenomena from an S-D logic and value co-creation perspective [e.g., 12–14]. Consequently, the implications of S-D logic and its core concept of value co-creation for IS are discussed in prevalent IS literature [15]. Against this backdrop, we argue that, since S-D logic fundamentally re-conceptualizes economic exchange [5, 16], this paradigmatic reorientation requires to rethink management of business and economic exchange. This, in turn, requires to rethink the way business is analyzed, designed, and managed—and, consequently, the way IS for business support are analyzed and designed.

To this end, we offer a *value co-creation ontology*—comprising constituent concepts of value co-creation and their inherent relationships—as one of the first steps toward reflecting S-D logic and value co-creation in IS analysis and design (ISAD). In this study, ontology refers to an artifact, constituted by a specific set of concepts and their inherent relations used to describe a certain reality [17]. However, the surge in academic and practical interest in value co-creation highlights an equivocal understanding of its conceptual boundaries and its real-world implications. McColl-Kennedy et al. [18], for instance, developed a catalogue of 27 different definitions of value co-creation, emphasizing underlying theoretical, semantical, and terminological ambiguity. The same holds true for constituent concepts of value co-creation from the lens of S-D logic such as the term service that has been defined in many different ways [19]. To wit, a multidisciplinary reference vocabulary in the form of an unambiguous value co-creation lexicon is not available. This lack—hindering the grounding of IS on value co-creation—has been raised not only in IS literature [20], but also in marketing research as the reference discipline of S-D logic [21, 22]. As such, the purpose of this study is (i) to develop an unambiguous and multidisciplinary *value co-creation glossary* of its constituent concepts; and (ii) to eventually derive an *ontology of value co-creation* comprising its constituent concepts and their relationships.

The remainder of this paper is structured as follows. Section 2 presents the theoretical background of S-D logic and value co-creation in marketing research and reviews the presence of S-D logic and value co-creation in IS and ontology research. Section 3 explicates the employed research methodology and its stages. Section 4 presents the resulting ontology. Section 5 discusses the ontology and concludes.

2 Theoretical Background and Research Motivation

Since this study aims to develop an ontology of value co-creation from an S-D logic perspective, we first provide an overview of S-D logic and value co-creation from their root discipline of marketing research. We then review the presence of S-D logic and value co-creation in both IS and ontology literature to motivate and position our study.

S-D Logic and Value Co-creation in Marketing Research. The cumulative effort of bringing S-D logic to the forefront of marketing research has resulted in its core concept of value co-creation [4, 7, 23]—that is, "the processes and activities that underlie resource integration and incorporate different actor roles in the service ecosystem" [15, p. 162]. The basic conceptualization of value co-creation is pivotal for S-D logic in that it integrates all the related concepts and their relations. The process of value co-creation thus integrates S-D logic's core concepts namely, actor, resource, service, institutional arrangement, and service ecosystem [3]. In also explicating their relationships, value co-creation holds that actors integrate resources through service exchange which is configured by institutional arrangements through which service ecosystems endogenously emerge [3, p. 3]. S-D logic finds its root in marketing research, where it gained momentum since its inception by the landmark study of Vargo and Lusch [1]. To develop a comprehensive conceptual foundation, Vargo and Lusch proposed—and further made amendments on [e.g., 2, 3]—a set of foundational premises for S-D logic to distinguish it from G-D logic.

S-D Logic and Value Co-creation in IS Research. S-D logic and its core concept of value co-creation have been influential in theorizing IS phenomena [14]. Extant IS research promotes and employs the notion of S-D logic and value co-creation through four streams. While the first stream offers an *early introduction* to S-D logic and value co-creation [e.g., 16, 24], the second stream introduces and examines *service innovation* [e.g., 12, 15]. In the third stream of research, *realization and practical implications of S-D logic and value co-creation* are illustrated [e.g., 25, 26]. Finally, since S-D logic offers new, distinctive philosophical assumptions about economic exchange, the fourth stream of research promotes *S-D logic and value co-creation as a theoretical lens* to study various IS phenomena. This lens has been dominantly applied for studying co-creation of IT value [e.g., 27].

S-D Logic and Value Co-creation in Ontology Research. Extant research uses the notion of *ontology* as a pragmatic approach to structure and codify knowledge about the concepts, relationships, and axioms/constraints pertaining to a domain (i.e., value co-creation) [28]. As such, extant IS research refers to the use of ontologies in ISAD as *ontology-driven IS* in which an explicit ontology plays a central role in analyzing and designing all of the system aspects and components [17, 29]. There are few studies aiming at one or a few fragmented aspects of value co-creation in ontology development endeavors. Nevertheless, no such work on a value co-creation ontology from an S-D logic perspective is available. We thus argue that the shift to S-D logic requires to eventually account for the central role of value co-creation in IS analysis and design by means of a value co-creation ontology. The latter would support the development of methods and techniques for business analysis from a value co-creation perspective and, thereby, enable the analysis and design of IS that support such business context. Therefore, in line with current applications in IS research, we posit that S-D logic's conceptual foundations with its unifying core concept of value co-creation provide a sound theoretical basis to offer a complementary view on ontology for IS. A sound ontology of value co-creation is fundamental in facilitating ontology-driven IS [17, 30], which eventually fosters the realization of value co-creation through its reflection in the design of new IS. The latter thereby lays emphasis on the role of IS in materializing S-D logic.

3 Methodology

To systematically develop a value co-creation ontology, we opt for *Methontology*, a step-by-step and well-structured methodology to build ontologies [31]. *Methontology* has been frequently used in different disciplines such as artificial intelligence, computer science, IS, and law [e.g., 32, 33] to guide ontology development in various domains. As shown in Figure 1, while stages I to IV of *Methontology* are within the scope of the paper at hand, stages V and VI are out of scope owing to our primary focus on synthesizing concepts and relations entailing the concept of value co-creation. Table 1 summarizes *Methontology*'s stages.

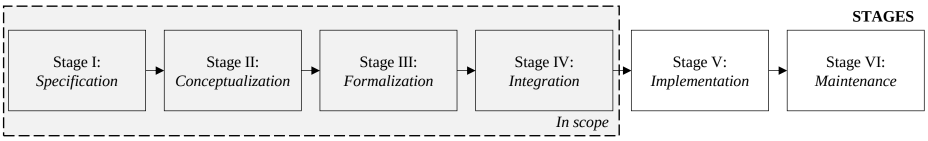


Figure 1. Stages of *Methontology* for Ontology Development [adapted from 31]

Table 1. The Applied Stages of *Methontology* in our Ontology Development [28, 31]

Stage	*Purpose*	*Activities*
I. Specification	Identification of area and scope	*Scoping:* Defining the ontology's purpose, application scenarios, and morphological analysis; *Cue-N-Anchoring:* Grounding the ontology on foundations of S-D logic
II. Conceptua-lization	Structuration of domain knowledge	*Baselining:* Conducting a systematic literature review to extract main concepts and relations; *Listing:* Developing a glossary of concepts' definitions, explanations, and examples
III. Formalization	Representation of specifiable ontology	*Crisscrossing:* Identifying and representing relationships; *Evaluating:* Verifying the ontology's soundness against theoretical foundations in seminal S-D logic studies
IV. Integration	Reuse of existing ontologies' constituents and documentation	*Reusing:* Integrating relevant components of existing ontologies; *Documenting:* Presenting the developed ontology through UML class diagram

Table 2. The Scope of the Value Co-creation Ontology Based on [28]

Dimension	*Scope*	*Justification*	*Compared to*
Positioning	Computa-tional ontology	Focus on future computational implementation and use in the pursuit of other pragmatic objectives in a specific application *instead of* the purely academic pursuit to know about the nature of reality [17].	Philosophical ontology; Conceptual schema
Domain	Ontology *for* IS	Focus on specifically and expressly designing the ontology as foundation *for* IS analysis and design *instead of* an attempt to capture a comprehensive ontology *of* all the concepts and relationships that are pertinent to understanding and reasoning about the IS universe [28].	Ontology *of* IS
Type	Domain-level	Focus on the specific domain of value co-creation from an S-D logic perspective *instead of* (i) very general concepts that are not specific to any domain (top-level) or (ii) processes/tasks to be accomplished in a specific application (application-level) [28].	Top-level; Application-level
Formality	Semi-formal	Focus on expressing content in a restricted and structured form, representable in an artificial formally defined language (e.g., UML) *instead of* no (informal) or very rigorously specified (mathematical, programmable) definitions, rules, and structures (formal) [34].	Informal; Formal

Specification (Stage I) is concerned with analyzing, explicating, and documenting—in natural language—the ontology's purpose (i.e., problem, intended uses, and users), level of formality (i.e., informal, semi-formal, or formal), area, and scope. Drawing on the ontology's *ex ante* intended purpose and application scenarios, this stage comprises *scoping* and *anchoring* activities. *Scoping* defines the area and scope of the ontology for which we use morphological distinctions offered by Kishore and Sharman [28] (see Table 2). *Anchoring* specifies a set of domain-specific, context-specific, or literature-specific ideas as anchors. The proposed ontology should be adequately grounded in these anchors so that the development effort is both guided and protected from loss of direction (*cue-n-anchor* notion) [28]. To this, we base our ontology of value co-creation on the theoretical foundations of S-D logic. The ontology's specified level of formality, area, and scope are the results of this stage.

Table 3. Glossary of Value Co-creation's Core Concepts from the S-D Logic Perspective

Concept	*Definition/Explanation*	*Example*
Value	In G-D logic, is fundamentally derived and determined in *exchange* (i.e., embedded in a firm's output and captured by price). In S-D logic, value is fundamentally derived and determined in *use* (i.e., the integration and application of resources in a specific context) [35, p. 145]. As such, in S-D logic, there is no value created until goods are used—that is, experience and perception are essential to value determination [1–3].	Dependable jet pro-pulsion at optimal efficiency
Actor	In S-D logic, actors are any social and economic agent with varying sizes (e.g., individual organism, family, firm), who provide input to the value creation processes for the benefit of the other actor [36, 37]. Compared to G-D logic, S-D logic thus posits a "generic actor" [3, p. 3] abstraction that disassociates actors from predesignated roles (e.g., "producers" and "consumers"). As such, the role of actors is to co-create value through resource integration and service provision in a network of other actors [15, 36].	Passenger, jet turbine producer, airline, airport operator
Resource	While in G-D logic resources have historically been viewed as tangible things that humans use for support, often natural resources that are fixed or limited in supply, S-D logic, however, refers to resources as anything an actor can draw on for support [1, 15]. As such, S-D logic underscores that resources comprise not only tangible, natural, and static resources, but also intangible and dynamic functions of human ingenuity and appraisal. Furthermore, they can be internal to actors and under their control or external to actors but capable of being drawn on for support [38].	Turbine, fuel, sensor, flying skills, turbine maintenance skills
Service	In S-D logic, service is the fundamental basis of economic exchange, which refers to "applying specialized competences (knowledge and skills) through deeds, processes, and performances for the benefit of another actor or the actor itself" [15, p. 158]. While in G-D logic service refers to a unit of output in exchange for another good (e.g., money), in S-D logic service refers to the *processes and activities* of applying specialized competencies for the benefit of and in conjunction with another actor [5, 39].	Airtime
Institutional Arrange-ment	Being central to understanding human systems and social activity in general [3, p. 7], *institutional arrangements* are fundamental in understanding the structure and dynamics of value co-creation. They refer to sets of interrelated humanly devised rules, norms, and beliefs that enable and constrain actors' actions and that make the exchange of service in their respective ecosystem predictable and meaningful (sometimes referred to as *institutional logics*) [3, p. 7].	Federal Aviation rules, regulation, guidelines
Service Ecosystem	G-D logic underscores a dyadic process of *value exchange* in traditional supply chains (i.e., neoclassical industrial perspective). S-D logic posits a collaborative process of *value co-creation* in service ecosystems (i.e., network-centric perspective). A service ecosystem is "a relatively self-contained, self-adjusting system of mostly loosely coupled social and economic (resource integrating) actors connected by shared institutional logics and mutual value creation through service exchange" [15, p. 162].	Aviation ecosystem

Conceptualization (Stage II) is concerned with structuring the domain knowledge in a conceptual model that describes the problem and its solution in terms of the domain vocabulary. Drawing on the ontology's level of formality, area, and scope (from *specification* stage), this stage comprises *baselining* and *listing* activities. While in the *baselining* activity we review relevant literature to uncover all relevant concepts and to eliminate redundancies and ambiguities, we develop a glossary as well as a tentative structure of the ontology in the *listing* activity [28]. We conduct *baselining* by means of a literature review[1], and carry out *listing* through developing a glossary of value co-creation concepts comprising definitions, explanations, and examples of the constituent concepts. The ontology's glossary of concepts and tentative structure are the results of this stage (see Table 3 and Table 4). We provide an example for each of the derived concepts by employing the jet turbine case of shifting from G-D to S-D logic [10–12].

Table 4. Glossary of Value Co-creation's Extended Concepts in S-D Logic

	Concept	*Definition/Explanation*	*Example*
Value	*Value Proposition*	"Value propositions establish connections and relationships among service systems" [35, p. 148]. Value propositions thus exist only to facilitate the value co-creation process and do not have value per se. Once a service beneficiary (e.g., customer) accepts a value proposition, he is an integrator of the value proposition (offered by service offeror) with his other resources in order to create value. This indicates and focuses on the notion of *value-in-use* [4, 24, 40].	Availability of airtime
	Value-in-exchange	"Value-in-exchange is the negotiated measurement offered and received (e.g., money and value proposition) among exchange partners" [35, p. 150]. As such, value-in-exchange occurs when the offeror offers a value proposition to the beneficiary or a service system offers a value proposition to another service system [1, 2, 35].	Money for airtime availability
	Value-in-use (-context)	"The customer's experiential evaluation of the product or service proposition beyond its functional attributes and in accordance with his/her individual motivation, specialized competences, actions, processes, and performances" [23, p. 293]. Value-in-use (context) indicates the use and co-creation of value proposition in the specific context of a service beneficiary or the specific context of a service system [1, 2, 4, 35].	Aircraft mobility during an actual flight
Actor	*Service Offeror*	Actor that (1) makes offers to other actors so that they cannot deliver value, but only offer value propositions; and (2) "collaboratively creates value following acceptance of value propositions, but cannot create and/or deliver value independently" [2, p. 7]. As such, the service offeror is both value facilitator (creator of value-in-exchange that does not create any value per se) as well as value co-creator (in fulfilment step). The offeror co-creates value during direct engagement and interaction with beneficiary [4, 5, 8, 39].	Jet turbine producer
	Service Beneficiary	Actor that (1) benefits from other actors that supply him with service or resources; and (2) always uniquely and phenomenologically determines value [2, p. 7]. As such, the service beneficiary is the realizer of value, i.e., co-creator of value-in-use to generate value based on the proposed resources from the offeror.	Airline, passenger
	Context	Uncontrollable, exogenous environmental conditions of surrounding social, ecological, and governmental contingencies [35]. For instance, environmental resources such as time, weather, and laws are always integrated in and relied on the value co-creation process. As such, "all actors are connected with other actors and other resources, and these connections provide the *context* for the actors to experience value" [15, p. 159].	Weather, air traffic control, operational issues
Resourc	*Operan**d** Resource*	Tangible, static, and passive components of goods or products that an actor acts on to obtain support (i.e., they enable or facilitate) (Vargo and Lusch, 2004). They are "seen as vehicles for service provision, rather than primary to exchange and value creation" [41, p. 374].	Sensor, aircraft, fuel

[1] We include marketing journals that are ranked *(world) leading* (tagged with *) by at least one of the ratings included in the 57th Harzing Journal Quality List (2016). We search in the Business Source Premier database employing the EBSCOhost search engine since S-D logic's inaugural year 2004 [1]. 30 selected papers have the phrases *"service-dominant"*, *"service logic"*, or *"dominant logic"*, on the one hand, and *"co-creat*"* or *"cocreat*"*, on the other hand, in abstract and in the title. In addition, we include studies on and/or using S-D logic that are published in the AIS basket-of-eight journals. This adds another 15 papers, most of which are part of the MISQ special issues on "Service Innovation in the Digital Age" [12] and on "Co-creating IT Value" [27]. The selected 45 papers are coded to identify constituent concepts of value co-creation guided by two recent literature reviews [7, 23] on value co-creation.

Table 4. (*continued*) Glossary of Value Co-creation's Extended Concepts in S-D Logic

	*Operan**t** Resource*	Intangible, dynamic, and active resources (e.g., human knowledge and skill) that act on other resources to produce effects rather than being operated on (i.e., they initiate or trigger) [1]. They are seen as "the fundamental source of competitive advantage" [2, p. 7].	Flying skills, maintenance skills
Service	*Service Platform*	"A modular structure that consists of tangible and intangible components (resources) and facilitates the interaction of actors and resources (or resource bundles)" [15, p. 166]. As such, actors employ service platforms as a vehicle to facilitate their day-to-day service exchanges and to mutually co-create value.	Jet turbine
Service Ecosystem	*Service System*	"A dynamic value-cocreation configuration of resources, including people, organizations, shared information (language, laws, measures, methods), and technology, all connected internally and externally to other service systems by value propositions" [16, p. 399]. As such, service system is the basic abstraction of service science and S-D logic [16, 42]. In effect, a service ecosystem is a network of different service systems.	Airport operator system, airline system
	Service Exchange	Service systems' simultaneous, reciprocal activities and processes of accessing, adapting, and integrating operan***t*** resources to create value for themselves and for other service systems [35]. Service exchange between different service systems is thus the basis for economic exchange and is a learning process towards value co-creation [1, 2, 35]. As such, the crux of the contrast between S-D and G-D logics lies in the basis of exchange. S-D logic focuses on the action of operan***t*** resources, whereas G-D logic focuses on the exchange of operan***d*** resources [1].	Exchange of airtime service between airline and turbine producer

Table 5. Relationships among Value Co-creation Ontology's Constituent Concepts

Relation	*Association to S-D logic's Foundational Premises (FPs) and Axioms*
R1: Actors *determine* value.	FP10 (AXIOM): Value is always uniquely and phenomenologically determined by the beneficiary. FP7: Actors cannot deliver value but can participate in the creation and offering of value propositions.
R2: Actors *integrate* resources.	FP6 (AXIOM): Value is co-created by multiple actors, always including the beneficiary. FP9 (AXIOM): All social and economic actors are resource integrators.
R3: A service instance *is composed of* at least one operant resource.	FP3: Goods are distribution mechanisms for service provision. FP4: Operant resources are the fundamental source of strategic benefit.
R4: Actors *exchange* service for service.	FP2: Indirect exchange masks the fundamental basis of exchange. FP5: All economies are service economies.
R5: Service *creates* value.	FP1 (AXIOM): Service is the fundamental basis of exchange.
R6: Institutional arrangements *configure* (i.e., enable/constrain) service and exchange.	FP11 (AXIOM): Value co-creation is coordinated through actor-generated institutions and institutional arrangements.
R7: Actors *create* institutional arrangements.	
R8: Institutional arrangements *determine* value.	
R9: Institutional arrangements *govern*/evaluate the emergence of service ecosystems.	
R10: A service ecosystem *is composed of* actors.	FP5: All economies are service economies. FP7 Actors cannot deliver value but can participate in the creation and offering of value propositions.

Formalization (Stage III) is concerned with compiling the specifiable ontology according to the chosen degree of formality from the *specification* stage (i.e., informal, semi-formal, formal). Drawing on the ontology's glossary of concepts and tentative structure (from *conceptualization* stage), this stage encompasses *crisscrossing* and *evaluating* activities. *Crisscrossing* structures the baseline glossary into a specifiable ontology comprising all structural constituents required in ontology construction. We execute *crisscrossing* by (i) extracting relationships from S-D logic literature and (ii) representing the extracted relationships. For representation, we use existing labels and notations of prevalent modelling languages as well as further labels that we extract from S-D logic literature (see Figure 2 and Figure 3). *Evaluating* is concerned with testing the emerging structure for soundness after the addition of each single new relationship in a coherent fashion [28]. We *evaluate* ontology's soundness against the outlined eleven foundational premises as well as the five constraints/axioms in [3] (see Table 5).

Integration (Stage IV) is concerned with reusing existing ontology constituents (i.e., concepts, relationships, definitions, constraints/axioms, and attributes) that are already built into other ontologies. Drawing on the specifiable ontology, this stage is made of *reusing* and *documenting* activities. *Reusing* selectively and sequentially integrates existing ontologies into the given ontology and documents the ontology after each integration step [28]. In developing an ontology of value co-creation, we reuse components from existing ontologies that focus on S-D logic and/or value co-creation. *Documenting* represents the developed ontology in a document for communication purposes [31] for which we use a UML class diagram (see Figure 2 and Figure 3). Following these four stages, *implementation (Stage V)* is concerned with codifying the ontology in a formal programming language (e.g., C++, Prolog, Ontolingua), and *maintenance (Stage VI)* is concerned with modifying the ontology's constituents (e.g., definitions) during usage time. Stages V and VI are out of scope. In the next section, we explicate how the realization of the four stages and their activities have yielded the *value co-creation ontology form the perspective of S-D logic*.

4 Value Co-creation Ontology

Here we report on the value co-creation ontology (see Figure 2 and Figure 3). We developed it by iteratively (i) extracting relationships between value co-creation's concepts and (ii) representing the relationships in UML class diagrams (see Stage III).

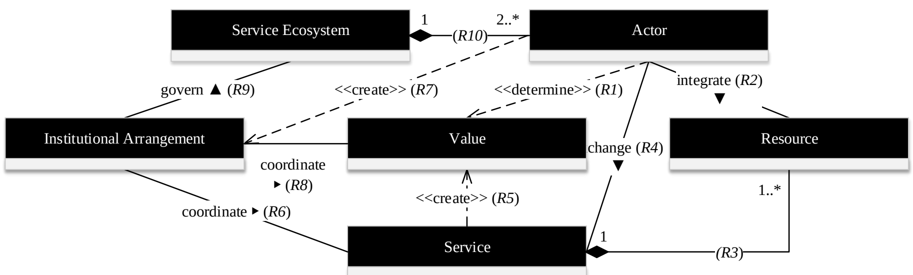


Figure 2. Value Co-creation Ontology for Core Concepts

Drawing on value co-creation ontology's core concepts (see Table 3 and Figure 2), value is always uniquely and phenomenologically determined by the actor being supplied with a service (*R1*). As such, S-D logic underscores that value occurs through the processes and activities underpinning usage and consumption by an actor who integrates its resources in realizing the value (*R2*). Therefore, a service requires at least one resource (*R3*) and it is exchanged between actors to access, adapt, and integrate resources among themselves for the benefit of another actor or the actor itself (*R4*). An instance of a designed, offered, and exchanged service creates an instance of value once the service is received and used by an actor (*R5*). To enable and constrain exchange of service, each service instance is configured guided by institutional arrangements (*R6*). Notably, rules, norms, beliefs, and their interrelation (e.g., civil law, criminal law) are not nature-given entities; they are rather created by actors (*R7*). Since these institutional arrangements are humanly devised, they also determine value in that they impact an actor's interpretation and determination of what actually is valuable and what is not (*R8*). Actors are then connected by their shared institutional logics and mutual value co-creation, both of which govern and evaluate the emergence of the nested and overlapping service ecosystems (*R9*). As such, institutional arrangements are key in fostering cooperative and coordinated behavior among actors. Eventually, service ecosystems are composed of at least two loosely coupled actors (*R10*). Building on the ontology's extended concepts (see Table 4), Figure 3 further represents all concepts (both core and extended) and their relationships. Table 5 specifies the relationships among the ontology's constituent concepts.

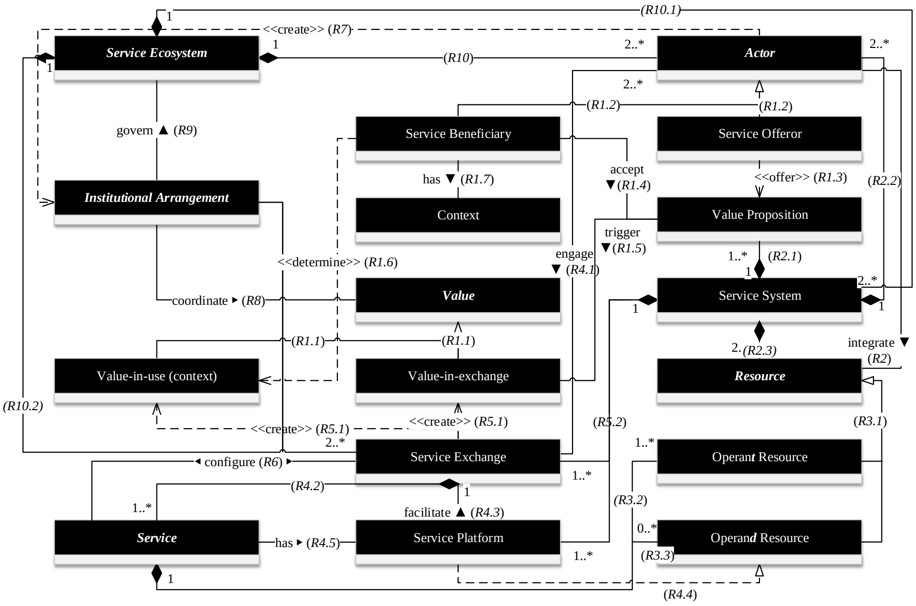


Figure 3. Value Co-creation Ontology for Extended Concepts (including **Core Concepts**)[2]

[2] Relation 4 (R4) (see Figure 2) is not present in Figure 3 due to the inclusion of the *service exchange* concept.

5 Discussion and Conclusion

This study starts with the premise that information systems (IS) analysis and design (ISAD) should account for reorientation from goods-dominant (G-D) to service-dominant (S-D) logic. One of the primary steps toward reflecting S-D logic, and its core concept of value co-creation, in ISAD is to develop an ontology of value co-creation based on S-D logic's theoretical foundations. To this end, this study develops an ontology of value co-creation for IS from an S-D logic perspective. Through employing *Methontology*, we synthesize the evolving S-D logic literature into a glossary of value co-creation's constituent concepts and represent their inherent relationships. We do so in the structure of a computational, semi-formal, and domain-level ontology for IS.

Practitioners are provided with an organizing language to understand, develop, and apply value co-creation. By applying the ontology, they can more clearly define the specific aspects required in facilitating value co-creation. Practitioners might anticipate areas of concerns and take appropriate measures by means of the ontology. Reflecting the ontology can be valuable for organizations that may be motivated to account for and realize S-D logic, but may not be aware of inherent intricacies and managerial actions to cope with. We are hopeful that such organizations would benefit from (1) reflecting on value co-creation's impact on organizations; and from (2) consciously utilizing relevant conceptual knowledge embedded in this work in their ISAD activities.

The study's results draw on, integrate in, and extend current studies in both IS and marketing research. Concerning IS research, our study integrates into both (i) IS studies that emphasize and employ S-D logic and value co-creation in advancing the discipline [e.g., 14, 15, 43]; and (ii) IS research that call for the use of ontologies in ISAD as *ontology-driven IS* [17, 29]. As such, we extend IS research in offering a *value co-creation ontology* aimed at grounding ontology-driven IS on value co-creation. We also demonstrate the consolidation of so far detached *activities* of ontology building for IS [28] in traditional *stages* of ontology development [31]. Concerning marketing research, this study draws on pertinent, seminal studies on value co-creation and S-D logic [e.g., 3] to synthesize and integrate value co-creation's constituents. Following marketing scholar's call for a multidisciplinary reference vocabulary lexicon [2, 24, 44], we extend existing marketing research through synthesizing the diverse body of value co-creation knowledge and providing a systematic representation of its underlying concepts and their relationships.

However, since the world of ontologies is vast and capturing conceptual boundaries and empirical constituents of value co-creation is complex as well, we anchor our approach in S-D logic with its rigorous, yet specific, interpretation and conceptualization of value co-creation. For instance, scholars deem *co-production* a dimension of value co-creation, which is potentially abstracted by, but not explicitly integrated in S-D logic [23]. Moreover, this study does not realize *Methontology*'s stages V and VI—concerned with the ontology's implementation and maintenance—due to our primary focus on specifying and formalizing the ever-growing domain of interest. This calls for future research in using the developed ontology as a basis to its further implementation and refinement. The developed ontology can also inform prospective endeavors in developing a value co-creation modeling language.

References

1. Vargo, S.L., Lusch, R.F.: Evolving to a New Dominant Logic for Marketing. J. Mark. 68, 1–17 (2004).
2. Vargo, S.L., Lusch, R.F.: Service-dominant logic: continuing the evolution. J. Acad. Mark. Sci. 36, 1–10 (2008).
3. Vargo, S.L., Lusch, R.F.: Institutions and axioms: an extension and update of service-dominant logic. J. Acad. Mark. Sci. 44, 5–23 (2016).
4. Payne, A.F., Storbacka, K., Frow, P.: Managing the co-creation of value. J. Acad. Mark. Sci. 36, 83–96 (2008).
5. Grönroos, C.: Value co-creation in service logic: A critical analysis. Mark. Theory. 11, 279–301 (2011).
6. Grönroos, C., Ravald, A.: Service as business logic: implications for value creation and marketing. J. Serv. Manag. 22, 5–22 (2011).
7. Galvagno, M., Dalli, D.: Theory of Value Co-creation: A Systematic Literature Review. Manag. Serv. Qual. Int. J. 24, 643–683 (2014).
8. Prahalad, C.K., Ramaswamy, V.: Co-creating unique value with customers. Strategy Leadersh. 32, 4–9 (2004).
9. Macdonald, E.K., Kleinaltenkamp, M., Wilson, H.N.: How business customers judge solutions: Solution quality and value in use. J. Mark. 80, 96–120 (2016).
10. Neely, A.: Exploring the financial consequences of the servitization of manufacturing. Oper. Manag. Res. 1, 103–118 (2008).
11. Fichman, R.G., Dos Santos, B.L., Zheng, Z.: Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. MIS Q. 38, 329-A15 (2014).
12. Barrett, M., Davidson, E., Prabhu, J., Vargo, S.L.: Service Innovation in the Digital Age: Key Contributions and Future Directions. MIS Q. 39, 135–154 (2015).
13. Eaton, B., Elaluf-Calderwood, S., Sørensen, C., Youngjin Yoo: Distributed Tuning of Boundary Resources: The Case of Apple's Ios Service System. MIS Q. 39, 217-A12 (2015).
14. Srivastava, S.C., Shainesh, G.: Bridging the Service Divide Through Digitally Enabled Service Innovations: Evidence from Indian Healthcare Service Providers. MIS Q. 39, 245-A19 (2015).
15. Lusch, R.F., Nambisan, S.: Service Innovation: A Service-dominant Logic Perspective. MIS Q. 39, 155–175 (2015).
16. Maglio, P.P., Vargo, S.L., Caswell, N., Spohrer, J.: The service system is the basic abstraction of service science. Inf. Syst. E-Bus. Manag. 7, 395–406 (2009).
17. Fonseca, F., Martin, J.: Learning The Differences Between Ontologies and Conceptual Schemas Through Ontology-Driven Information Systems. J. Assoc. Inf. Syst. 8, 129–142 (2007).
18. McColl-Kennedy, J.R., Vargo, S.L., Dagger, T.S., Sweeney, J.C., van Kasteren, Y.: Health care customer value cocreation practice styles. J. Serv. Res. 1094670512442806 (2012).
19. Goldstein, S.M., Johnston, R., Duffy, J., Rao, J.: The service concept: the missing link in service design research? J. Oper. Manag. 20, 121–134 (2002).
20. Alter, S.: Challenges for Service Science. J. Inf. Technol. Theory Appl. 13, 22 (2012).
21. Godsiff, P.: Service systems and requisite variety. Serv. Sci. 2, 92–101 (2010).
22. O'Shaughnessy, J., O'Shaughnessy, N.J.: The service-dominant perspective: a backward step? Eur. J. Mark. 43, 784–793 (2009).
23. Ranjan, K.R., Read, S.: Value co-creation: concept and measurement. J. Acad. Mark. Sci. 44, 290–315 (2016).

24. Lusch, R.F., Vargo, S.L., Wessels, G.: Toward a conceptual foundation for service science: Contributions from service-dominant logic. IBM Syst. J. 47, 5–14 (2008).
25. Ordanini, A., Pasini, P.: Service co-production and value co-creation: The case for a service-oriented architecture (SOA). Eur. Manag. J. 26, 289–297 (2008).
26. Yan, J., Ye, K., Wang, H., Hua, Z.: Ontology of collaborative manufacturing: Alignment of service-oriented framework with service-dominant logic. Expert Syst. Appl. 37, 2222–2231 (2010).
27. Grover, V., Kohli, R.: Cocreating It Value: New Capabilities and Metrics for Multifirm Environments. MIS Q. 36, 225–232 (2012).
28. Kishore, R., Sharman, R., Ramesh, R.: Computational Ontologies and Information Systems I: Foundations. Commun. Assoc. Inf. Syst. 14, 158–183 (2004).
29. Guarino, N.: Formal Ontology and Information Systems. In: Proceedings of International Conference on Formal Ontology in Information Systems 1998. pp. 81–97. , Trento (1998).
30. Chen-Huei Chou, Mariam Zahedi, F., Huimin Zhao: Ontology-based Evaluation of Natural Disaster Management Websites: A Multistakeholder Perspective. MIS Q. 38, 997–1016 (2014).
31. Fernández-López, M., Gómez-Pérez, A., Juristo, N.: Methontology: from ontological art towards ontological engineering. (1997).
32. Corcho, O., Fernández-López, M., Gómez-Pérez, A., López-Cima, A.: Building legal ontologies with METHONTOLOGY and WebODE. In: Law and the semantic web. pp. 142–157. Springer (2005).
33. Fernández López, M., Gómez-Pérez, A., Pazos Sierra, A., Pazos Sierra, J.: Building a chemical ontology using methontology and the ontology design environment. (1999).
34. Sharman, R., Kishore, R., Ramesh, R.: Computational Ontologies and Information Systems II: Formal Specification. Commun. Assoc. Inf. Syst. 14, 184–205 (2004).
35. Vargo, S.L., Maglio, P.P., Akaka, M.A.: On value and value co-creation: A service systems and service logic perspective. Eur. Manag. J. 26, 145–152 (2008).
36. Spohrer, J.C.: On looking into Vargo and Lusch's concept of generic actors in markets, or "It's all B2B …and beyond!" Ind. Mark. Manag. 40, 199–201 (2011).
37. Vargo, S.L., Lusch, R.F.: It's all B2B…and beyond: Toward a systems perspective of the market. Ind. Mark. Manag. 40, 181–187 (2011).
38. Akaka, M.A., Vargo, S.L.: Technology as an Operant Resource in Service (Eco)systems. Inf. Syst. E-Bus. Manag. 12, 367–384 (2014).
39. Grönroos, C.: Service logic revisited: who creates value? And who co‐ creates? Eur. Bus. Rev. 20, 298–314 (2008).
40. Gummesson, E.: Extending the service-dominant logic: from customer centricity to balanced centricity. J. Acad. Mark. Sci. 36, 15–17 (2008).
41. Pels, J., Vargo, S.L.: Toward a transcending conceptualization of relationship: a service-dominant logic perspective. J. Bus. Ind. Mark. 24, 373–379 (2009).
42. Maglio, P.P., Spohrer, J.: Fundamentals of service science. J. Acad. Mark. Sci. 36, 18–20 (2008).
43. Nambisan, S.: Information Technology and Product/Service Innovation: A Brief Assessment and Some Suggestions for Future Research. J. Assoc. Inf. Syst. 14, 215–226 (2013).
44. Lusch, R.F., Vargo, S.L.: Service-dominant logic: reactions, reflections and refinements. Mark. Theory. 6, 281–288 (2006).

Critical Success Factors for Introducing Smart Services: A Supplier's Perspective

Sonja Dreyer[1], Jan Zeren[1], Benedikt Lebek[2], and Michael H. Breitner[1]

[1] Leibniz Universität Hannover, Information Systems Institute, Hannover, Germany
{dreyer,breitner}@iwi.uni-hannover.de
janzeren@hotmail.de
[2] BHN Dienstleistungs GmbH & Co. KG, Aerzen, Germany
lebek.benedikt@bhn-services.com

Abstract. Smart services enable providers to win new customers and to strengthen the relations with existing ones. Established companies emphasize the importance of providing smart services in addition to their current portfolio. To be able to offer smart services successfully, various preconditions must be fulfilled. Critical success factors are identified that can be used for companies which want to offer smart services in the future, focusing potential customers and their needs. Implementing smart services in existing business environments to offer added value to customers is a challenge. Due to a systematic literature review, critical success factors for introducing smart services are extracted from academic literature with the objective of an introduction in the interests of customers. Experts from practice evaluated and complemented the success factors. Both a theoretical and a practical perspective is taken to support introducing smart services.

Keywords: Smart service, smart service introduction, smart service implementation, critical success factors, supplier's perspective

1 Introduction

Smart services enable to satisfy individual and continuously changing customer needs [1]. Such services become increasingly important for providers to differentiate from competitors. Although there are companies exclusively focusing on providing smart services, most of the providers are existing companies, e.g. in the field of machine component supply. They aim at using smart services to create new business opportunities. The development of a solution wanted by a specific customer leads to increasing loyalty and relationships [2]. Requirements of customers are highly individual and continuously changing. From a customer's perspective, working with a company that is able to satisfy these needs is predestined to be a long-term partner. Thus, both the provider and the customer benefit from successful smart services. As smart services are complex, integrating their provision in existing sales structures is a practical challenge [3].

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

In academic literature, smart services are of growing interest in the last years. The smart service transition from theoretical design to practical implementation is addressed in several publications. Most of the publications emphasize the importance of considering customer's requirements as a critical success factor when implementing smart services. But several preconditions have to be fulfilled to be able to satisfy the needs of the customers and consequently, to implement smart services successfully. Although several publications present success factors for introducing smart services in an existing business environment, there is no systematic overview from a provider's perspective focusing the customer. To address this gap we proposed the following two research questions:

> *RQ1: What are critical success factors from a provider's perspective for introducing smart services that focus individual customer needs?*
> *RQ2: How do the critical success factors named in academic literature agree with success factors relevant in practice?*

To answer these research questions a systematic literature review was conducted based on Webster and Watson (2002) [4]. Using predefined search terms, we searched through different academic databases to identify relevant literature for smart service introduction. With basis on the identified literature, critical success factors for introducing smart services were worked out. Experts that were asked in interviews extended and evaluated the identified success factors.

The remainder of this article is structured as follows: In the second chapter the selected research proceeding is explained. Smart services are defined and existing literature in the field of introducing smart services is presented in chapter three. Experts to verify the results are consulted afterwards. The results are discussed and implications are presented systematically. The article concludes with limitations, approaches for further research and conclusions in the chapters four and five.

2 Research Design

Based on existing literature in the field of smart service implementation and expert interviews, the objective of this study was to identify critical success factors. The article is motivated by a growing interest in smart services, reflected in academic literature. Introducing smart services in an existing business environment is a real challenge in the field. However, there is no systematic overview of critical success factors in literature taking the provider's perspective. Such success factors support organizations to focus the customers and their needs in the smart service implementation process. Additionally, researchers gain knowledge regarding the smart service transition.

With the objective to get an overview of existing literature in the field of smart service introduction, we conducted a systematic literature review. According to Webster and Watson (2002) we searched in six different databases, including AISeL, IEEEXplore, JSTOR, ScienceDirect, SpringerLink, and Taylor&Francis, using predefined search terms [4]. To reduce the number of hits, in the first step the results were limited to English articles. We also excluded whitepapers, book chapters and other non-academic publications. A number of twelve articles were identified to be relevant by reading the titles and the abstracts. If we were not able to decide whether a publication is relevant for our topic, we took a look at the whole text. Both a forward and a backward search were conducted with basis on the identified articles. This led to two further articles that were included in our overview. Finally, we identified 14 articles concerning the introduction of smart services from a provider's perspective with focus on the customer.

The relevant literature was screened afterwards. Critical success factors for introducing smart services were worked out under the precondition that they are explained from a supplier's perspective. Critical success factors are understood as factors that are of particular importance for a successful smart service introduction and consequently, to remain competitive [5]. According to that definition, in this article an in literature named aspect is listed as a critical success factor if it is estimated to be necessary for a successful introduction. In total, nine different success factors were identified. Each publication was screened regarding the considered success factors. It was differentiated between partially considered factors, what means that they were named but not described in detail, and considered success factors, what means that they were discussed, analyzed or critically reviewed. They formed the basis for semi-structured interviews with twelve experts from eight different companies in the fields of IT, electronics, manufacturing and medical technology. All companies have in common that they plan to provide smart services and partially already offer digital services. From literature extracted critical success factors were both critically evaluated regarding their importance in practice and extended. It resulted in critical success factors that can be used as guidelines for the introduction of smart services, ensuring a focus on the customer. Figure 1 shows the research design that was used.

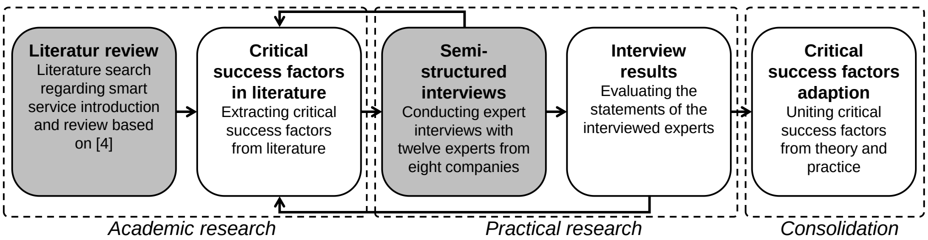


Figure 1. Research design

3 Critical Success Factors for Smart Service Introduction

Smart services are enabled through information and communication technology. The objective of smart services is to address customer-individual requirements [6]. This is realized through both interacting with the customer and data-collection in real-time [7]. In literature, this is called value co-creation. With the help of connected systems and machine intelligence [8], context information are used to offer services adjusted to the requirements of customers [9]. The information that is used concerns the technology, the environment as well as the social context [10], [11]. Information from several sources is combined and forms the basis for data analyses. Results can be presented in dashboards or used to make suggestions proactively [12]. High dynamics [13] and the use of collaboration technology enables continuous communication and feedback which is also essential for smart services [14]. Allmendinger and Lombreglia (2005) conclude the most important characteristics of smart services [15]. Smart services are enabled through information. Context information, frequently collected in real-time, is used for analyses, e.g. using machine intelligence. Smart services enable to avoid unexpected events and thus create a new kind of value. In the present article, a provider is a company or organization that offers smart services to customers. The customer can be a business or private user [1]. The named aspects are taken as the basis for the following investigations.

3.1 Extracting Critical Success Factors from Literature

According to Priller et al. (2014), cyber security and the accompanying privacy *(1)* have a great share when introducing smart services [16]. Smart services require a large amount of data and information, partially in real-time. As this data may be sensitive, an effective security concept is inevitable. Demirkan et al. (2015) agree and state that both a security and privacy solution has to be designed before offering smart services [6]. Data and information form the basis for competitive advantages which is why security and privacy concerns are an important aspect for potential customers. In addition, legal bases *(2)* have to be considered in an early stage of introducing smart services [17], [18]. Referring to the large amount of data and information, transparency is necessary about the ownership and usage permissions. The influence of technology developments on legal issues has to be clarified [19].

Employee qualification *(3)* is another success factor for introducing smart services that is named in literature. Although the qualification of employees is always important when offering products or services to customers, the way how to train employees changed [20]. The age of digitalization is the reason for changing requirements on the employees of a company that plans to offer smart services and to support the introduction in the customer's company. Competencies are named in this context and summarize knowledge and skills that are necessary for successfully introducing smart services [21]. Smart services are very customer-centric which is why sales training is highlighted as important when talking about employee qualification [18]. This also helps to understand customer needs. Apart from employee's individual competencies the interplay with other employees from different departments *(4)* is named a success factor. The interplay of different capacities and experiences plays a role [22]. Nevertheless, Pétercsák et al. (2016) point out that the stakeholders have to be selected carefully [23]. It is necessary that they are willing to share their knowledge and to cooperate with the other team members. From a customer's perspective, it is helpful to have a clear contact that is responsible. According to Allmendinger and Lombreglia (2005) it is also important that the management participates *(5)* on the new developments [15]. Otherwise, the project of offering smart services fails because the provider is not assessed to be innovative by potential customers. The management level is responsible for creating an innovation culture in companies because an innovative strategy has to be integrated in the overall corporate strategy [19]. Considering the customer needs *(6)* is the central aspect of successful smart services. This has also to be included in the smart service introduction phase [24]. The motivation why a customer wants to use a smart service has to be reflected in the offer. Klötzer and Pflaum (2017) share the same opinion and emphasize that additional customer value has to be created [21]. The customer is estimated to be more important than the technology [18]. Cooperation with the customers is suggested to successfully introduce smart service offerings. An advantage is that the lifecycle of a smart service *(7)* can better be understood in the specific application [24]. An integrated lifecycle model considering hybrid products and their services contributes to the usage of the value creation potential [25]. Investigating what kinds of activities a customer carries out enables to adapt a smart service. Requirements can be detected and the way the customer works can be supported [15]. It enables to identify what type of smart service is relevant for potential customers.

Table 1. Literature categorized by named critical success factors

Author	*1) Security and privacy*	*2) Legal basis*	*3) Employee Qualification*	*4) Interdisciplinary teams*	*5) Management involvement*	*6) Consideration of customer needs*	*7) Consideration of product lifecycle*	*8) IT infrastructure*	*9) Reference model use*
Allmendinger, Lombreglia 2005 [15]					●	●	●		
Baars, Ereth 2016 [26]	●		●					●	
Barile, Polese 2010 [19]		●			●	●			
Bullinger et al. 2015 [22]			●	○		●		●	●
Demirkan et al. 2015 [6]	●	●							○
Geum et al. 2016 [24]						●	●	●	
Klötzer, Pflaum 2017 [21]			●	○		●		●	
Lê Tuán et al. 2012 [27]	●							●	●
Lerch, Gotsch 2015 [20]			●					○	○
Lesjak et al. 2014 [28]	○			●					
Pétercsák 2016 [23]			●	●	●			●	
Priller et al. 2014 [16]	●	●							
Theorin et al. 2016 [17]		●						●	
West, Gaiardelli 2016 [18]		●	●		●	●	●		

● Considered ○ Partially considered

The compatibility of a new service with the IT infrastructure *(8)* has a substantial impact on successful implementation [22]. Baars and Ereth (2016) identify several challenges that arise out of introducing smart services [26]. A major point is the data integration. It has to be solved due to an appropriate IT infrastructure. Theorin et al. (2016) describe how an IT infrastructure can look like that is suitable for smart services [17]. Modern hardware in form of large and high-speed storages are fundamental [26]. This does not only apply for the smart service provider but also for the customer. Looking at the customer's location, an appropriate IT infrastructure is necessary to use the full potential of a smart service. In academic publications, reference models *(9)* are developed that help to orientate when starting to offer smart services. Bullinger et al. (2015) recommend using a suitable reference model when introducing smart services [22]. This is underpinned by the fact that the introduction of smart services requires comprehensive planning [6]. Table 1 summarizes the identified literature in conjunction with the considered critical success factors.

3.2 Comparing Theoretically Developed Success Factors with Practice

To compare the theoretical results with practical impact twelve experts were asked regarding critical success factors for smart services (Table 2). The experts agreed that smart services enable a better communication with the customer (B, G2). In contrast to selling products, the offer of services means both continuous contact and revenues. Products often have already been exploited which is why services in connection with products are a logical expansion for providers (A2). Additionally, it is a consistent move to stay competitive (D). Although smart services are increasingly in focus, already existing internal processes should be used and adapted (D). Processes carried out by the customer also should be considered when introducing a new smart service. This is especially important for large companies because they often lack agility (F). As a major challenge an expert named the fact that smart services are often location-independent (G1). This means that the services can be performed all over the world.

Table 2. Interviewed experts and their settings

Company/participant	*Company characteristics/department*
Company A	IT sector, IT service provider
Participant A1	Demand management
Participant A2	Innovation
Company B	IT sector, process optimization
Participant B	Programming
Company C	Electronics industry, communication systems supplier
Participant C1	Consulting management
Participant C2	Technical service
Company D	Manufacturing industry, component producer
Participant D	Technical service
Company E	Manufacturing industry, automotive supplier
Participant E1	Planning
Participant E2	Planning and programming
Company F	Manufacturing industry, automotive and industrial supplier
Participant F	Strategy management
Company G	Manufacturing industry, automation systems producer
Participant G1	Innovation
Participant G2	Product management, technical service
Company H	Medical technology industry, medical instruments producer
Participant H	Programming and development

Security and privacy concerns *(1)* are considered as important factor when starting to offer smart services because new requirements result. Machines are more and more connected with external networks such as the internet (A1). But not only machines and further technical devices have to be considered when talking about security and privacy but also the people that work in the new environment. Regularly executed trainings are named by an expert as possible solution to increase sensitivity for the topic (E2). The possible future customer has to have confidence in the service provider (A2). This applies especially for smart services as sensible data are exchanged and processed. But experts emphasized that complying security and privacy standards do not add value for a customer (F, G1). Nevertheless, these are prerequisites to be able to offer smart services successfully. Similar considerations apply to legal bases *(2)*. As the whole field of using large amounts of data from the customer is a fast and dynamic development, there is often a lack of clarity regarding the legal framework (E1). An expert highlighted that projects in the field of smart services have to be checked constantly regarding legal compatibility (E2). In contrast, some experts held the opinion that the legal basis is a purely matter of definition which is why it does not take up much time (A1, H). However, legal bases have to be unambiguous, internally in the provider's and customer's company as well as between the involved partners.

The impact of employee qualification *(3)* regarding the success of introducing smart services is estimated as comparatively low by the experts (A1, C1, C2, G1). Some of them explained that it is important to sensitize for the topic of smart services (F, G2) and to ensure that all employees have the same understanding of it (D). The handling and maintenance of the systems that are necessary for the offer of smart services is carried out by people (E2). Therefore, by some interviewees it is considered important and decisive for the success that employees are integrated in the introduction process of smart services (E2, G1). As processes are partially changing, trainings may help to reduce uncertainties and to increase efficiency (A1). Another tendency is that the systems will be much simpler. Several experts highlighted that smart services have to be structured in such a way that comprehensive trainings are not necessary (F, G1). Only selected employees should go through trainings, such as employees from the sales department. An expert from the electronics industry was of the opinion that capacity often is much more important than qualification (C2). That is why trainings are not always the path to success. However, it has to be ensured that the customer of a smart service has a reliable partner to contact who supports smart service concerns.

By all experts, smart services are estimated as highly complex, which is why decisions of individual employees are assessed to be inefficient (E1). Interdisciplinary teams *(4)* help to make decisions that satisfy various departments. The process of introducing smart services is estimated to take much longer if there would not be this type of teams (G1). An interviewee of the manufacturing industry pointed out that interdisciplinary teams are not limited to the company but the potential customers should also be involved when making decisions (G2). As smart services always have to be individualized, a cooperation simplifies the introduction in the customer's company. The importance of involving the management *(5)* is perceived very differently by the experts. Smart services are a new development direction that has to be exemplified by the management (D). But the project management *(10)* participates on the operating business and is generally seen as more important (E2). The importance of management involvement depends on the company structure (A2). The consideration of customer needs *(6)* is regarded the top-ranked critical success factor. Smart services always have to be developed for the customer which is why they play the decisive role in a successful introduction. All already named success factors are connected with the customer. Individually added value has to be created which is why a development away from the customer has necessarily to be avoided (C2). An expert stated that customer requirements are more important than a fast introduction of a smart service (F). In the ideal case the expectations are more than met (C1, C2).

Service lifecycles *(7)* are usually shorter than lifecycles of classic hardware. Nevertheless, it is recommended to consider the lifecycles of hardware and services together because profitability often results from combining products and services (G2). An expert notes that this only applies for smart services related to products (A1). Completely new smart services have a diverging lifecycle which is why existing lifecycles cannot be applied. In addition, the lifecycle of a smart service is highly externally controlled (C2). Individual circumstances in the company of the customer should be considered. Additionally, already implemented services should be taken into account when introducing a new smart service (D). This facilitates the introduction and the use of the smart service because duplication of work is avoided.

Availability plays a decisive role in the context of smart services. In this, a suitable IT infrastructure *(8)* forms the basis for a smart service (E2). Although this is a prerequisite for offering a smart service, it is not a top-ranked success factor. Regarding the IT infrastructure there were two diverging opinions. On the one hand it was stated that smart services should follow the existing infrastructure (E2). On the other hand it was explained the other way round, the IT should always follow the solution (A1, F). Which way is selected by the companies often depends on the level of investment. This applies both for the service provider and the customer.

When introducing a smart service for the first time, reference models *(9)* might be useful. An expert agreed and complemented that it highly depends on the model. The better a reference model goes with the company's settings the better it can be used (F, G1). Therefore, very high-level models offer low added value (E1, E2, H). Smart services are dynamic which is why it is difficult to use static reference models (A1). Requirements are changing continuously and therefore, feedback of the customers is estimated to be much more valuable (F).

Clearly defined requirements not only from the customers but also from internal stakeholders *(11)* were named by two experts as an additional success factor because they help to avoid department specific problems (E1, E2). A flexible project management is required and named as essential for a successful introduction of smart services (H). As main reason, the dynamic environment of that kind of service was named. Due to the dynamic environment, both in the provider's and the customer's company, a quick adaption to changes is vitally important. It is decisive how well it succeeds to implement new smart services in existing business models *(12)* and the corporate strategy (A2, D). In this, it is important to develop a suitable pricing strategy (A1). The use of smart services usually is linked to additional costs for the customer. This has to be considered in the pricing. The price of such a service has to be justified, otherwise the smart service would not be requested by customers. It is stated that it can only be avoided through a consistent orientation on the requirements of the customers. An exemplary solution for pricing is to reduce the costs of the product related to the smart service.

Table 3. Critical success factors for introducing smart services

Critical success factors	*Description*
1) Security and privacy	Data that is used for smart services is sensitive. An effective security and privacy concept is inevitable.
2) Legal basis	Transparency is necessary regarding the ownership and usage permissions of data and information.
3) Employee qualification	It is important to sensitize for the topic of smart services. Trainings help to understand customer needs and to improve smart service introduction.
4) Interdisciplinary teams	Smart services are highly complex what requires interdisciplinary teams, both internally and between provider and customer.
5) Management involvement	An innovation culture can only be created when the management exemplifies it. The importance of the management depends on the company structure.
6) Consideration of customer needs	Smart services always have to be developed for the customer. Only services that satisfy customer needs can be successful.
7) Consideration of product lifecycle	The lifecycle of smart services is partially different to those of products. This should be taken into account.
8) IT infrastructure	An appropriate IT infrastructure is a precondition for smart services because data has to be collected, transmitted, processed and analyzed.
9) Reference model use	Reference models that are suitable to the setting of a company are useful as orientation.
10) Project management	A flexible project management enables a quick adaption to the changing environment and contributes to implement smart services in an existing strategy.
11) Consideration of internal requirements	Requirements of internal stakeholders are useful as milestones and should be considered to avoid internal inconsistencies.
12) Business model	New smart services have to be embedded in business models. A suitable pricing strategy has to be worked out as well as suitable cooperation partners.
13) Consideration of the market	Cultural aspects have to be considered when introducing a worldwide applicable smart service. The market positioning is important.
14) Standardization	Smart services should be standardized to be able to adapt them easily and to avoid complications with partners and customers.

The consideration of the market *(13)* is emphasized to be highly important by several experts from practice (A2, C1, C2, D, G2). On the one hand, cultural aspects do not have to be neglected when introducing a worldwide applicable smart service (G2). The openness towards new kinds of products and services differs across countries and continents. This influences the best way to access a new market. Additionally, aspects such as the suitable marketing strategy derive from this. On the other hand, the positioning on the market is important (G1). A high market presence can enable to win new customers and to establish standards. Standardization *(14)* is named important to reduce efforts regarding the introduction of smart services (A2, C1, D, G2). The introduction process is shorter and less complicated, especially regarding technical aspects. The consideration of the market and related aspects is proved the most important success factor that is not in focus in literature. Table 3 summarizes all critical success factors from a provider's perspective for introducing smart services that were worked out both from literature and from expert interviews.

4 Limitations and Further Research

Academic literature concerning a smart service implementation from a provider's perspective formed the basis of the investigations. Relevant literature was found through using English search terms. Search terms in other languages were not taken into account. All search terms included the term "smart service". Other terms that might be used as synonyms were not considered. Including search terms in other languages and considering synonyms for smart services might have led to further results. With twelve interviews in eight different companies, the number of experts consulted is not necessarily representative. Asking further experts regarding critical success factors for introducing smart services might lead to additional aspects. This applies especially to experts from not yet considered industries. As smart services are still under development, in most of the companies there are not yet best practice approaches for introducing a new smart service that focuses the customer and how to include them in an existing business environment. Furthermore, it is not evaluated yet what kind of customer is predestined for smart service usage. Finally, it cannot be said whether the critical success factors will change in the future.

5 Conclusions

Critical success factors from a provider's perspective for introducing smart services were worked out. The objective was to provide guidelines that help to provide smart services that satisfy the customer. A smart service can only be introduced successfully when it is implemented in the interests of the customers. A number of 14 publications formed the basis for identifying critical success factors from literature. Nine critical success factors were found and analyzed. Theoretically extracted factors were compared to requirements from the field. Twelve interviews were conducted with experts form eight different companies. The experts came from the IT sector, from the electronics and manufacturing industry and from the medical technology industry. On the one hand, already extracted success factors were evaluated. It resulted in the finding that all critical success factors extracted from literature in average are estimated to be relevant, apart from the reference model use. On the other hand, the success factors were extended to be practically relevant. A success factor that was named several times to be very important and that was not focused in literature was the consideration of the market. As smart services are often provided worldwide, cultural aspects and market dynamics have to be taken into account. With focus on the potential customers, providers who want to offer smart services have to choose carefully how to access a new market and how to obtain a high market presence. As both a theoretical and practical perspective were taken the elaborated critical success factors are useful as guidelines for companies that want to start offering smart services to customers.

References

1. Massink, M., Harrison, M., Latella, D.: Scalable Analysis of Collective Behaviour in Smart Service Systems. In: 25th ACM Symposium on Applied Computing, pp. 1173–1180. ACM, New York (2010)
2. Rabetino, R., Kohtamäki, M., Gebauer, H.: Strategy Map of Servitization. International Journal of Production Economics 192, 144–156 (2016)
3. Knote, R., Blohm, I.: It’s not about Having Ideas — It’s about Making Ideas Happen! Fostering Exploratory Innovation with the Intrapreneur Accelerator. In: 24th European Conference on Information Systems, Paper 3 (2016)
4. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. Management Information Systems Quarterly 26(2), xiii–xxiii (2002)
5. Rockart, J.F.: Chief executives define their own data needs. Harvard Business Review 57(2), 81–93 (1979)
6. Demirkan, H., Bess, C., Spohrer, J., Rayes, A., Allen, D.: Innovations with Smart Service Systems: Analytics, Big Data, Cognitive Assistance, and the Internet of Everything. Communications of the Association for Information Systems 37(1), 733–752 (2015)
7. Wünderlich, N.V., von Wangenheim, F., Bitner, M.J.: High Tech and High Touch: A Framework for Understanding User Attitudes and Behaviors Related to Smart Interactive Services. Journal of Service Research 16(1), 3–20 (2012)
8. Gavrilova, T., Kokoulina, L.: Smart Services Classification Framework. In: 6th Federated Conference on Computer Science and Information Systems, pp. 203–207 (2015)
9. Oh, J.S., Park, J.S., Kwon, J.R.: Ubiquitous Infrastructure and Smart Service on City Gas Environments in Korea. In: 5th International Conference on Future Information Technology. IEEE Press, New York (2010)
10. Alahmadi, M., Qureshi, M.R.J.: Improved Model to Test Applications Using Smart Services. Science International (Lahore) 27(3), 2275–2280 (2015)
11. Lee, J.Y., Kim, M.K., La, H.J., Kim, S.D.: A Software Framework for Enabling Smart Services. In: 5th IEEE International Conference on Service-Oriented Computing and Applications. IEEE Press, New York (2012)
12. Kynsilehto, M., Olsson, T.: Checkpoints, Hotspots and Standalones — Placing Smart Services over Time and Place. In: 7th Nordic Conference on Human-Computer Interaction, pp. 209–218. ACM, New York (2012)

13. Byun, J., Park, S.: Development of a Self-Adapting Intelligent System for Building Energy Saving and Context-Aware Smart Services. IEEE Transactions on Consumer Electronics 57(1), 90–98 (2011)
14. Wünderlich, N.V., Heinonen, K., Ostrom, A.L., Patricio, L., Sousa, R., Voss, C., Lemmink, J.G.A.M: 'Futurizing' Smart Service: Implications for Service Researchers and Managers. Journal of Services Marketing 29(6/7), 442–447 (2015)
15. Allmendiger, G., Lombreglia, R.: Four Strategies for the Age of Smart Services. Harvard Business Review 83(10), 131–145 (2005)
16. Priller, P., Aldrian, A., Ebner, T.: Case Study: From Legacy to Connectivity — Migrating Industrial Devices into the World of Smart Services. In: 19th IEEE Emerging Technologies and Factory Automation. IEEE Press, New York (2014)
17. Theorin, A., Bengtsson, K., Provost, J., Lieder, M., Johnsson, C., Lundholm, T., Lennartson, B.: An Event-Driven Manufacturing Information System Architecture for Industry 4.0. International Journal of Production Research 55(5), 1297–1311 (2016)
18. West, S., Gaiardelli, P.: Driving the Servitization Transformation through Change Management: Lessons Learnt from Industrial Cases. In: Spring Servitization Conference (2016)
19. Barile, S., Polese, F.: Smart Service Systems and Viable Service Systems: Applying Systems Theory to Service Science. Service Science 2(1–2), 21–40 (2010)
20. Lerch, C., Gotsch, M.: Digitalized Product-Service Systems in Manufacturing Firms: A Case Study Analysis. Research-Technology Management 58(5), 45–52 (2015)
21. Klötzer, C., Pflaum, A.: Toward the Development of a Maturity Model for Digitalization within the Manufacturing Industry's Supply Chain. In: 50th Hawaii International Conference on System Science, pp. 4210–4219 (2017)
22. Bullinger, H.J., Meiren, T., Nägele, R.: Smart Services in Manufacturing Companies. In: 23rd International Conference on Production Research (2015)
23. Pétercsák, R., Maccani, G., Donnellan, B., Helfert, M., Connolly, N.: Enabling Factors for Smart Cities: A Case Study. In: 37th International Conference on Information Systems (2016)
24. Geum, Y., Jeon, H., Lee, H.: Developing New Smart Services Using Integrated Morphological Analysis: Integration of the Market-Pull and Technology-Push Approach. Service Business 10(3), 531–555 (2016)
25. Blinn, N., Nüttgens, M., Schlicker, M., Thomas, O., Walter, P.: Lebenszyklusmodelle hybrider Wertschöpfung: Modellimplikationen und Fallstudie an einem Beispiel des Maschinen- und Anlagenbaus. In: Multikonferenz Wirtschaftsinformatik, pp. 711–722 (2008)
26. Baars, H., Ereth, J.: From Data Warehouse to Analytical Atoms — The Internet of Things as a Centrifugal Force in Business Intelligence and Analytics. In: 24th European Conference on Information Systems, Paper 3 (2016)
27. Lê Tuán, A.L., Quoc, H.N.M., Serrano, M., Hauswirth, M., Soldatos, J., Papaioannou, T., Aberer, K.: Global Sensor Modeling and Constrained Application Methods Enabling Cloud-Based Open Space Smart Services. In: 9th International Conference on Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing. IEEE Press, New York (2012)
28. Lesjak, C., Ruprechter, T., Bock, H., Haid, J., Brenner, E.: ESTADO — Enabling Smart Services for Industrial Equipment through a Secured, Transparent and Ad-hoc Data Transmission Online. In: 9th International Conference for Internet Technology and Secured Transactions, pp. 171–177. IEEE Press, New York (2014)

Evaluation der organisationalen UX-Gestaltungskompetenz

Dominique Winter[1] und Gunnar Stevens[2]

[1] Hochschule Emden/Leer, Emden, Deutschland
dominique.winter@designik.de
[2] Hochschule Bonn-Rhein-Sieg, Sankt Augustin, Deutschland
gunnar.stevens@h-brs.de

Abstract. Organisationen wollen Produkte mit guter User Experience herstellen. Durch die Evaluation der organisationalen UX-Gestaltungskompetenz können Organisationen erkennen, wie stark ihre momentane UX-Gestaltungskompetenz ausgeprägt ist und wie die Kompetenz gezielt gesteigert werden kann. Für die Abbildung der aktuellen Kompetenz werden ein Fragebogen zur theoretischen Kompetenz und ein Fragebogen für die Produktevaluation kombiniert. Durch diese Kombination wird die Kompetenz der Organisation aus der Handlungs- und der Ergebnisperspektive betrachtet.
Für die Erarbeitung von Handlungsfeldern zur Verbesserung der Kompetenz werden qualitative Interviews durchgeführt und mit den Ergebnissen der quantitativen Erhebungen verknüpft. Durch einen anschließenden Ergebnisworkshop erarbeiten sich die Mitglieder der Organisation einen effizienten Weg zur Steigerung der organisationalen UX-Kompetenz.

Keywords: User Experience, Kompetenz, Evaluation, Organisation

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Product Updates: Attracting New Consumers Versus Alienating Existing Ones

Jens Foerderer and Armin Heinzl[1]

[1] Business School, University of Mannheim, Germany.
{foerderer,heinzl}@uni-mannheim.de

Abstract. Are product updates—in terms of producers' decision to add new features "over-the-air"—an effective means in stimulating greater product demand and appeal? Our difference-in-differences analyses of a matched sample of 17,247 mobile apps in Google Play over a period of 24 weeks documents mixed consumer reactions. Whereas updates attracted new consumers, existing consumers rated an app 1.1% worse than before the update and compared to a control group of not updated apps. Why did existing consumers react negatively to updates? Our data provides little support that economic reasoning—i.e., direct costs or learning costs imposed by the update—underlies these reactions. Instead, negative reactions appear—at least to a certain degree— as a behavioral phenomenon, as consumer reviews show an increased density of affective vocabulary after an update. We conclude that updates may stimulate new demand but may alienate existing consumers.

Keywords: Software update, Mobile apps, behavioral economics, difference-in-differences, propensity score matching.

1 Introduction

Product updates have become an essential instrument in firms' product management repertoire. Once representing solely a means of providing bug fixes and minor improvements, updates are increasingly being used to provide consumers with new features over-the-air [1–3]. Such "feature updates" allow producers to add new functionality after a product's market release and while it is in use by consumers [4, 5]. Since 2008, for example, Apple has added several hundreds of features to its iPhone, including the "Siri" personal assistant, "facetime" video calls or "iCloud", a feature to synchronize files across devices. The increasing use of software makes feature updates also relevant for many products that traditionally relied little on information technology. For example, cars are increasingly software-based, enabling automakers to push substantial new features to their fleet without requiring consumers to bring their vehicle into the dealer's garage [1, 6]. Tesla, for example, recently made a product update to improve driving performance [7]. Even beyond cars and phones, any other products have become updateable, including televisions, washing machines, or books [3]. In sum, updates enable producers to alter a product in use by consumers, instead of being limited to introducing features only over subsequent new product releases. In this

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

sense, updates offer producers new possibilities in creating long-lasting demand and appeal for their products [1, 8], eventually representing “a new set of strategic choices related to how value is created and captured” [2].

In this paper, we study the effects of updates on product demand and ratings. Updates are not trivial to study because they affect two distinct audiences: *existing* consumers of a product as well as potentially *new* consumers attracted by an update. Updates may be an effective means to increase demand for a product by attracting new consumers. As updates introduce new features, they may increase the likelihood of a product to enter the consideration set of consumers. This presumption stems foremost from traditional utility theory, which suggest that each additional product attribute that consumers perceive positively increases consumers’ utility [9, 10]. First empirical evidence seems to confirm this suggestion. In the context of browser add-ons, Tiwana [5] finds frequent updates linked to higher downloads. Similarly, consumers seem to choose digital products that offer more features over comparable ones with fewer features [11–14].

Reactions of existing consumers are less clear. Tiwana et al. [5] find a link between updates and higher product ratings. Fleischmann et al. [4] observe in an experimental setting that updates positively disconfirm existing consumers, leading to a greater intention to continue using a product. These contributions notwithstanding, there is anecdotal and empirical evidence that existing consumers react negatively. Anecdotes of product failures suggest for instance a “feature creep”, meaning that the ongoing addition of features may result in bloated and over-complicated products, ultimately making consumers abandon a product [15]. In addition, simple Google searches yield hundreds of tutorials for how to undo recent updates of Instagram, Facebook, or Snapchat, to only name a few. Research so far suggests both rational and behavioral drivers for negative reactions of existing consumers. A rational reason might be that updates confront consumers with costs for switching from the product they know to the updated one [16–19]. Switching costs include transaction costs for an update (e.g., update fees, data plans) and learning costs (e.g., handling new features, changes in the user interface). If switching costs imposed by an update exceed potential utility, consumers may react negatively to the update. Apart from rational considerations, negative reactions may also reflect a behavioral phenomenon driven by psychological ownership [20, 21], reluctance toward novelty [22], endowment [23, 24], or routine-seeking behavior [25]. For example, consumers may be reluctant toward novel things such as new features regardless of potential utility. In sum, however, conclusive empirical evidence on the effect of updates as well as the mechanisms driving these effects remains scarce.

Studying the effects of updates is difficult because it requires detailed product-level data that allows identifying changes made by updates and isolating the reactions of existing consumers and new consumers. An investigation should also account for unobservable and observable differences among products and producers that may confound consumer reactions. Finally, the above considerations underscore that such an investigation requires identifying causality. The decision to update is endogenous to producers, and likely to suffer from reverse causality. Because of this complexity, the accurate measurement of the consequences of updates has proven elusive.

Our study uses data on mobile apps, which allows us to track product-level information on consumer reactions (i.e., ratings, downloads, and reviews) as well as textual information on updates released by app developers (i.e., changelogs). Our dataset comprises a weekly balanced panel of 17,247 distinct apps listed in the Google Play Store U.S., the largest market for apps worldwide, over 24 weeks in 2016. To address endogeneity bias, we use propensity score matching [26, 27]. We match apps that are updated (treatment group) to apps that are not updated (control group) but equivalent given our observational data. In our context, we have access to several thousands of potential control observations that we can use as inputs for the matching procedure, which bolsters our control group design [27, 28].

2 Theoretical Background

Updates are "self-contained modules of software that are provided to the user for free in order to modify or extend a software after it has been rolled out and is already in use" [4]. Updates are no stand-alone products but rather are integrated into the base product [29, 30]. In this paper we are interested in product updates that add new functionalities to a product. We refer to such updates as "feature updates" and use the term "update" for the sake of brevity. An example for feature updates are Apple's regular updates for its iPhone. The 2016 update for instance brought more than 50 new features.

Updates are evaluated by two distinct audiences, which also require differentiated theoretical considerations. One evaluating audience consists of *new* consumers attracted by an update. The other evaluating audience encompasses *existing* consumers of a product. The main argument of our paper is that new consumers and existing consumers react differently to updates. In the following, we discuss this argument in more detail.

Feature updates may attract new consumers by making it more likely that the product enters the consideration sets of consumers. This presumption comes from traditional utility theory, which has modeled consumer preferences using an additive utility function [10]. Utility models assume each additional product attribute that consumers perceive positively to increase consumers' utility. This idea finds itself implemented in many market research techniques, such as the conjoint analysis or discrete choice models [e.g., 9]. Because these models predict consumer outcomes based on expected utilities or part-worths for each product feature, the conclusion is that each positively valued feature adds to the success of a product, compared to not having the feature.

Empirical evidence supports the prediction that consumers choose products that offer more features over comparable products with fewer features [11, 12, 14]. One consistent observation is that adding attributes to a product increases consumers' perceptions of its capability, resulting in improved product evaluations before ownership [12]. Consequently, feature updates are likely to attract new consumers, since they make the product more likely to be considered by a larger number of consumers. Thus, we argue:

Hypothesis 1: A feature update increases the number of new consumers of the product.

Reactions of existing consumers are less clear. Various evidence and arguments suggest negative reactions. To structure our discussion, we classify these arguments as rational (economic) and behavioral. “Rational” refers to the argument that actors, in terms of consumers, maximize their utility and are capable of gathering and evaluating all information for this purpose themselves [31]. “Behavioral”, by contrast, denotes systematic biases in an actor’s decision-making that cannot, or only in a very limited way, explained by rational logics [23]. From an economic perspective, existing consumers face costs of switching from the product they know to the updated product [17–19, 32]. If switching costs imposed by an update exceed the obtained utility, consumers may react negatively to the update. Based on Nilssen [32] and Klemperer [18], there are at least two types of switching costs relevant for our consideration of updates, namely transaction costs and learning costs. Transaction costs are immediate costs incurred by the update, in terms of the time, effort, and money in changing from the original product to its updated version [16, 18, 33]. Transaction costs include fees producers charge for the update. For example, some providers of mobile games charge for new game levels and gimmicks. Other transaction costs incur for data plans or opportunity costs associated with installing an update and fixing potential errors encountered in this process. Learning costs represent the effort required by a consumer “to reach the same level of comfort or facility with a new product as they had for an old product” [16]. For example, a feature update for a banking app might introduce a new verification procedure for making money transfers, which requires consumers to understand and learn the new procedures before being again able to transfer money. Another example might be that feature updates imply a reorganization of the user interface, eventually confronting consumers with costs for learning and understanding the handling the product.

Various studies yielded evidence for consumers’ economic considerations. Several studies document learning costs in the context of product features [11, 12, 34, 35]. Thompson et al. [34] asked study participants to choose between three variants of a digital device. More than sixty per cent of participants chose the variant with the most features. Similarly, when the researchers gave subjects the chance to customize their product, freely choosing from twenty-five features, subjects also maximized features. However, when the researchers asked the subjects to use the device, subjects evaluated products with many features more negatively than the ones with less features. Mukherjee and Hoyer [12] observe in the case of high-complexity products that additional features reduced product ratings because consumers made learning-cost inferences about these features. Finally, the findings of Meyer et al. [11] suggest that while consumers are more likely to adopt products with added features, they subsequently avoid using these features due to inferred learning costs. In sum, switching costs imposed by an update might explain negative reactions of existing consumers.

Negative reactions of existing consumers may also represent a behavioral phenomenon that exists aside from rationality. If the negative reactions are behavioral, consumers react negatively even if updates provided new capabilities and implying no costs at all. Among others, behavioral drivers include psychological ownership [20, 21], reluctance toward novelty [22], loss aversion or endowment effects [24], and routine-seeking behavior [25]. These behaviors have been documented across disciplines, and they provide theoretical arguments that existing consumers evaluate feature updates negatively. Transferred to our context, existing consumers may react negatively to updates because they represent novelty in the products they use. In a similar way, endowment or loss aversion tendencies of consumers might let them prefer keeping the "status quo" of their product rather than being provided with potentially useful new features [23, 24]. In sum, we argue:

Hypothesis 2: A feature update decreases existing consumers' evaluations of the product.

3 Method

We test our hypotheses in the context of mobile apps. Apps are a type of software for a specific and particular purpose, optimized for mobile devices [36, 37]. Typical examples of apps are email, calendar, stock market, and weather. Apps optimize the appearance of displayed data, taking into consideration the screen size and resolution [36]. The functionality of mobile apps is usually limited by the unique characteristics of mobile devices: they have comparably little processing power, are controlled by touch gestures, and used "on the go" [36]. We particularly study mobile apps that run on Google's Android platform. At the time of our study, more than 80% of all smartphones worldwide run Android (Gartner, 2016). This setting has the advantage that we can collect data directly from the Google Play Store, the largest store for Android apps. In Google Play, consumers can compare, rate, review, and obtain apps [38]. Producers can update their apps at any time. Updates are rolled out immediately and automatically to existing consumers "over-the-air". Consumers are not charged any direct costs for an update, yet producers may adjust upfront prices or in-app prices along with an update.

To allow for causal inference, we employ a matching strategy [26, 28, 39]. Matching strategies pair each observation that experiences the treatment of interest at a given point in time (in our case, apps that experience an update) with one or several similar observations that do not experience the treatment at that time (the control group). We observe each app at four subsequent points in time: two weeks before the update (t-2), one week before the update (t-1), one week after the update (t+1), and two weeks after the update (t+2). We estimate the effects of updates by calculating the difference-in-differences (DID) between updated and not-updated apps, before and after the update [28, 40].

We obtained a list of all apps in the Google Play Store as of June 2016 from a mobile analytics firm. We selected a random sample of 100,000 apps from this list, for which we collected app-specific information, including ratings, updates, prices, and text reviews, in an automated way in a weekly panel format. We filtered the obtained dataset as follows. Besides apps, Google Play lists content, including television shows, music, and books, and hedonic applications, including games. In order to ensure comparability, we excluded apps labeled as "books & references", "comics", "education", "libraries & demos", "news & magazines", "wallpaper", "widgets", and “games”. To ensure comparability, we dropped apps with less than ten downloads. In the following, we discuss the variables included in our study in more detail.

DOWNLOADS. To assess whether updates attract new consumers we use the number of downloads for an app. Google Play, as other app markets, does not provide precise measures of app downloads. Instead, Google gives a categorical indicator of the number of downloads (e.g., 5-10, 100-500, 500,000-1,000,000). To obtain a more detailed measure of downloads, we combined information on download intervals with the number of ratings for an app. In order to submit a rating, users must have downloaded an app, so the number of ratings can be considered a conservative lower bound to consumer demand [41]. The resulting variable DOWNLOADS is then the mean between the midpoint of the download interval for an app and its number of ratings, which we logged.

RATING. We assess consumer reactions to updates by the rating consumers give to apps. Consumers may evaluate apps by rating it from one to five “stars”, where one star represents a low rating and five stars represent a high rating of the app. Apps with higher ratings are perceived to fulfill user expectations, have an agreeable and engaging interface, and are well-suited to audiences’ needs [38]. Consumers can renew their ratings after an app update, which allows us to distinguish between ratings of existing consumers versus new consumers [38]. The Google Play Store provides the mean rating of all consumers, rounded to one decimal, which we report as RATING.

Focal predictors (UPDATE and AFTER). To identify feature updates, we hired two independent assistants who manually inspected the changelogs (or, release notes) app producers publish along an update. In changelogs, producers describe key aspects of an app update [cf. 42]. Prior work used version numbers (e.g., 2.0, 2.1) to identify updates, which may serve as a proxy of feature updates [e.g., 5]. Although it is an informal convention that integer increases in version numbers indicate major changes [42], this standard is not enforced in many contexts and subject to certain ambiguity. Moreover, version numbers do not allow inferring the extent of features added to an app. By contrast, changelogs provide detailed insights into the changes made [42]. In the Google Play Store, changelogs are displayed below the product description in a section entitled “What’s new”, which makes them an important aspect of producers’ communication. Changelogs are limited to 500 characters, which requires producers to precisely describe the update [38].

The central predictor in our model is the dichotomous variable UPDATE, which is one if the focal app was updated with a new feature and zero otherwise. DID analyses require a second indicator for distinguishing the periods before and after the event that is studied. Thus, we include the dichotomous indicator AFTER in our models, which is one for the weeks after the update. The DID estimator is then given by interacting AFTER with UPDATE.

We construct further variables to gain insights into economic and behavioral reasoning behind negative ratings. First, the variable FEATURES ADDED is the count of the features added in an update. Second, we count the number of words in the changelog, as captured in WORDS IN CHANGELOG. Third, we obtained the time since the last feature addition. The variable WEEKS SINCE LAST UPDATE counts the number of days since the last update. Fourth, we measure direct price changes associated with an update (PRICE, continuous, in US-Dollar). The variable FREE is an indicator of apps that charge no up-front price.

Finally, consumers can include text reviews with their ratings, and these reviews may offer further insights. If reactions to updates are behaviorally driven, consumers' textual responses may give an indication. To analyze consumer reviews we implemented natural language processing techniques. We use the standard semantic text analysis software LIWC [43] to capture major text semantics. We are particularly interested in capturing consumer reasoning in a review. LIWC offers for this purpose the category "affective processes" [43]. Technically, each category consists of a list of identifying words. LIWC scores the wordlists against a text, and subsequently assigns a numerical score depending on how many of the category words were observed in a text. The category we employ in our analyses is "affective mechanism". It includes 1393 word stems, among them "happy", "worried", "hate" or "ugly". In a five-word text, "I hate the new update", the output by LIWC is 20 (per cent) for the affective mechanism dictionary (i.e., one affective word "hate" divided by a total of five words in the text, and multiplying by 100%). Before conducting the LIWC analyses, we cleansed the reviews. We removed fill words from the text, lemmatized each word, and removed reviews that were not written in English language. We then score the affective mechanism dictionary against consumers' review texts for each app-week. This procedure resulted in a numerical score assigned to each app-week, which indicates the mean score for the affective category. We include this score as the continuous variable AFFECTIVE in our analyses.

Controls. We estimate our models with app-level and time (i.e., week) fixed effects. App fixed effects adjust for static differences among apps (e.g., functionality, usability, producer etc.). Time effects control for external events (e.g., announcements by Google) or trends (e.g., an increasing number of apps are published), in terms of that app producers vary their decisions to update in response to temporal events or short-term trends.

We followed Shadish et al. [27] to build our matched set of control apps. We relied on propensity score matching [26]. We match on observational characteristics and the time of the update. The critical task in both PSM is to choose matching criteria. Matching criteria are inherently context-specific [28]. Although our context has received some attention in prior literature [e.g., 36, 37, 41], evidence is scarce when it comes to indicators of updating. We therefore followed the procedure employed by Pahnke et al. [44] and used informal interviews with producers, analysts, and industry experts to derive suitable matching criteria. The interviews converged on a number of app-specific factors. The interviews revealed that updates are costly for producers. Producers may thus tend to invest only in "promising" apps, depending on downloads and ratings received. Thus, we used DOWNLOADS and RATING as criteria for the matching procedure. Producers also seemed to update apps more often for which they charged an upfront price. Thus, we added PRICE as a matching criterion. While not evident from our interviews, we also added WEEKS SINCE LAST UPDATE as a matching criterion to account for temporal differences.

We used nearest neighbor matching without replacement in the first week of the pretreatment period. The final sample consists of 68,988 app-weeks, spanning a period of 24 weeks from 1 July 2016 to 15 December 2016. We used Ordinary Least Squares (OLS) with heteroscedasticity-robust standard errors clustered on app to estimate the following baseline equation:

$$y_{i,t} = \beta_0 + \beta_1 \text{AFTER}_t \text{ x UPDATE}_i + V_i + \tau_t + \epsilon_{i,t} \quad (1)$$

The subscripts i and t index for app and week, respectively. The dependent variable is $y_{i,t}$. UPDATE_i is an indicator variable for whether app i is in the treatment group, AFTER_t equals 1 if the current week is after the treatment, V_i are app fixed effects and T_t are time fixed effect. The DID coefficient of interest is β_1, which can be interpreted as the relative change of the treatment group compared to the control group, caused by the treatment. The main effects, AFTER and UPDATE, are absorbed.

4 Results

We first turn toward analyzing the consequences of feature updates for app downloads. We estimate equation (1) with DOWNLOADS as dependent variable, which gives us the effect of feature updates on app downloads. Table 1 shows the results. In Model 1, we observe that the interaction of interest, AFTER x UPDATE is positive and significant. All other things being equal, a feature update increases downloads by approximately 1.8% on average, plus minus .3%. This finding indicates that feature updates attract new consumers, supporting Hypothesis 1.

Table 1: Main Results: The Effect of Updates on App Downloads and Ratings.

	(1) Log(Downloads)	*(2) Rating*
After x Update	$.018^{***}$	$-.040^{***}$
	(.003)	(.002)
Log(Downloads)		$.062^{***}$
		(.012)
Constant	9.202^{***}	3.560^{***}
	(.009)	(.110)
Specification	OLS	OLS
Adjusted R2	.02	.05

Note: Accounts for app and time fixed effects. *, **, *** indicate significance at the 5%, 1%, and .1% levels, respectively.

How do existing consumers react to feature updates in terms of ratings? We estimate equation (1) with RATING as dependent variable, in terms of the effect of feature updates on ratings by existing and new consumers. To isolate the effect of product updates on existing consumers' ratings, we control for the increase in an app's new consumers with DOWNLOADS. In Model 2, the coefficient of AFTER x UPDATE now gives the effect of feature updates on existing consumers' ratings. The coefficient of AFTER x UPDATE is negative and strongly significant. All other things being equal, existing consumers rate an app 1.1% more negative after an update, on average. Thus, existing consumers react negatively to feature updates, supporting Hypothesis 2. So far, our findings indicate that consumers rate updated apps worse than before the update. What explains this discount? Our literature background presented economic and behavioral explanations, which we seek to explore in our data in the following. As it is infeasible to design a formal test that allows fully rejecting either economic or behavioral reasons, we conduct counterfactual analyses. We warrant, however, that our analyses do not allow definite conclusions. Rather we conduct these analyses to provide a more detailed picture of the mechanisms in place. If the negative reactions of existing consumers are a behavioral phenomenon, then the semantical analyses of consumer reviews should provide us with an indicator. To explore behavioral drivers, we assessed the semantic text analyses of consumer reviews. If updates cause an increase in the usage of affective words in reviews of existing consumers, then we should have an indicator of a behavioral bias. Figure 1 plots AFFECT for updated and not updated apps, before and after the update. The plot shows an increasing number of affective words (e.g., hate, annoyed, bad) in consumer reviews for updated apps, whereas control apps remain almost at pre-update levels. The figure indicates an increase in affective vocabulary in existing consumers' reviews for updated apps. Econometric analyses confirm this observation.

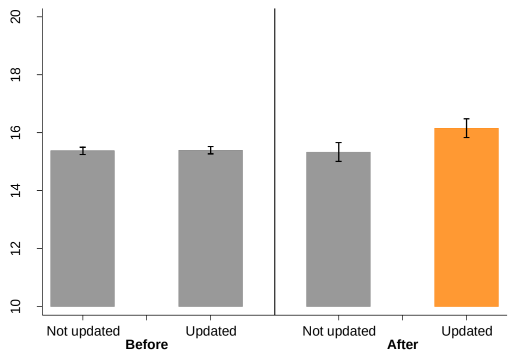


Figure 1: Mechanisms: Content of Consumer Reviews Before and After a Feature Update.

5 Discussion and Conclusion

Do updates cause greater product demand and appeal? In our data, updates appeared as an instrument to create new demand. All other things being equal, we find a feature update to increase downloads by approximately 1.8% on average. When considering the reactions of existing consumers, we found evidence of negative effects. All other things being equal, an update causes ratings of existing consumers to decline by approximately 1.1% on average, when compared with their ratings prior the update and to similar but not updated apps. These findings remained robust to various matching strategies and parameters, and account for app and time heterogeneity. In subsequent analyses, we found support that negative reactions of consumers are a behavioral phenomenon.

This paper directly responds to calls for understanding the management of digital products [1, 3, 8] and the "strategic choices related to how value is created and captured" [2]. The strategic choice we investigated is whether firms should invest in feature updates over the lifetime of their products—a phenomenon that has received increasing interest of scholars and managers. Our findings align with the observation in prior work that updates attract new consumers, yet object the observation that updates are perceived positively by existing consumers [4, 5].

Our paper sought to disentangle economic and behavioral drivers behind consumer reactions. If future research can confirm our finding of behavioral triggers behind existing consumers' negative reactions, producers' possibilities in reducing negative reactions might be limited. Here, we are left to speculate whether these reactions reflect an "update resistance", which might be a phenomenon of interest for product management [1, 3], consumer choice [e.g., 12–14], and software management literatures [e.g., 45–48]. The behavioral mechanisms actually underlying such an update resistance—such as reluctance toward novelty [22] or endowment [23]—require further decoding in future studies. Interviews with consumers and subsequent, experimental testing may help teasing out these mechanisms.

Finally, our context offered insights in the mobile app industry [37, 41, 49]. Our

Our findings suggest actionable patterns for managers. First, our findings support the effectiveness of feature updates in attracting new consumers to a product. The purposeful design of feature updates to attract particular consumer segments may allow producers to continuously adjust their products to various consumer groups as well as to be a "moving target" for competition. Second, our data do only little support economic reasons behind existing consumers' negative reactions. At least in our very context, the success of mitigation strategies focused on reducing the switching costs invoked by a feature update—such as reducing the extent and frequency of an update or providing guidelines and tutorials—might be limited.

References

1. Iyer, B., Venkatraman, N.V.: What Comes After Smart Products, https://hbr.org/2015/07/what-comes-after-smart-products.
2. Porter, M.E., Heppelmann, J.E.: How Smart, Connected Products Are Transforming Competition. Harvard Business Review. 92, 64–88 (2014).
3. Yoo, Y., Henfridsson, O., Lyytinen, K.: Research Commentary - The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. Information Systems Research. 21, 724–735 (2010).
4. Fleischmann, M., Amirpur, M., Grupp, T., Benlian, A., Hess, T.: The Role of Software Updates in Information Systems Continuance: An Experimental Study from a User Perspective. Decision Support Systems. 83, 83–96 (2016).
5. Tiwana, A.: Evolutionary Competition in Platform Ecosystems. Information Systems Research. 26, 266–281 (2015).
6. Svahn, F., Mathiassen, L., Lindgren, R.: Embracing Digital Innovation in Incumbent Firms: How Volvo Cars Managed Competing Concerns. MIS Quarterly. (forthcoming).

7. Taub, E.A.: Your Car's New Software Is Ready. Update Now?, http://www.nytimes.com/2016/09/09/automobiles/your-cars-new-software-is-ready-update-now.html.
8. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.: Digital Business Strategy: Toward a Next Generation of Insights. MIS Quarterly. 37, 471–482 (2013).
9. Green, P., Srinivasan, V.: Conjoint Analysis in Consumer Research: Issues and Outlook. Journal of Consumer Research. 5, 103–123 (1978).
10. Lancaster, K.: Consumer Demand: A New Approach. Columbia Universiry Press, New York, NY (1971).
11. Meyer, R.J., Zhao, S., Han, J.K.: Biases in Valuation Vs. Usage of Innovative Product Features. Marketing Science. 27, 1083–1096 (2008).
12. Mukherjee, A., Hoyer, W.D.: The Effect of Novel Attributes on Product Evaluation. Journal of Consumer Research. 28, 462–472 (2001).
13. Nowlis, S.M., Simonson, I.: The Effect of New Product Features on Brand Choice. Journal of Marketing Research. 33, 36–46 (1996).
14. Sela, A., Berger, J.: How Attribute Quantity Influences Option Choice. Journal of Marketing Research. 49, 942–953 (2012).
15. Surowiecki, J.: Feature Presentation, http://www.newyorker.com/talk/financial/2007/05/28/070528ta_ talk_surowiecki.
16. Chen, P.-Y., Hitt, L.M.: Measuring Switching Costs and the Determinants of Customer Retention in Internet-Enabled Businesses: A Study of the Online Brokerage Industry. Information Systems Research. 13, 255–274 (2002).
17. Farrell, J., Klemperer, P.: Coordination and Lock-in: Competition with Switching Costs and Network Effects. In: Handbook of Industrial Organization. pp. 1967–2072 (2007).
18. Klemperer, P.: Markets with Consumer Switching Costs. The Quarterly Journal of Economics. 102, 375–394 (1987).
19. Shapiro, C., Varian, H.R.: Information Rules: A Strategic Guide to the Network Economy. Harvard Business Press, Cambridge, MA (1998).
20. Pierce, J.L., Kostova, T., Dirks, K.T.: Toward a Theory of Psychological Ownership in Organizations. Academy of Management Review. 26, 298–310 (2001).
21. Pierce, J.L., Delbecq, A.L.: Organization Structure, Individual Attitudes and Innovation. Academy of Management Review. 2, 27–37 (1977).
22. Boudreau, K.J., Guinan, E.C., Lakhani, K.R., Riedl, C.: Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science. Management Science. (forthcoming).
23. Kahneman, D.: Maps of Bounded Rationality: Psychology for Behavioral Economics. The American Economic Review. 93, 1449–1475 (2003).
24. Kahneman, D., Knetsch, J.L., Thaler, R.H.: Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. The Journal of Economic Perspectives. 5, 193–206 (1991).
25. Oreg, S.: Resistance to Change: Developing an Individual Differences Measure. Journal of Applied Psychology. 88, 680–693 (2003).
26. Rosenbaum, P.R., Rubin, D.B.: The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika. 70, 41–55 (1983).
27. Shadish, W.R., Cook, T.D., Campbell, D.T.: Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Mifflin Company, Boston, USA (2002).
28. Angrist, J.D., Pischke, J.-S.: Mostly Harmless Econometrics: An Empiricist's Companion. Princeton Press, Princeton, NJ (2009).
29. Clements, P., Northrop, L.: Software Product Lines: Practices and Patterns. Addison-Wesley, Boston, MA (2002).
30. Ulrich, K.: The Role of Product Architecture in the Manufacturing Firm. Research Policy. 24, 419–440 (1995).

31. Kahneman, D., Tversky, A.: Prospect Theory: An Analysis of Decision Under Risk. Econometrica. 47, 263–291 (1979).
32. Nilssen, T.: Two Kinds of Consumer Switching Costs. The RAND Journal of Economics. 23, 579–589 (1992).
33. Whitten, D., Wakefield, R.L.: Measuring Switching Costs in IT Outsourcing Services. The Journal of Strategic Information Systems. 15, 219–248 (2006).
34. Thompson, D.V., Hamilton, R.W., Rust, R.T.: Feature Fatigue: When Product Capabilities Become Too Much of a Good Thing. Journal of Marketing Research. 42, 431–442 (2005).
35. Thompson, D.V., Norton, M.I.: The Social Utility of Feature Creep. Journal of Marketing Research. 48, 555–565 (2011).
36. Ghose, A., Goldfarb, A., Han, S.P.: How Is the Mobile Internet Different? Search Costs and Local Activities. Information Systems Research. 24, 613–631 (2013).
37. Ghose, A., Han, S.P.: An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet. Management Science. 57, 16711691 (2011).
38. Salz, P.-A.: Apponomics: An Insider's Guide to a Billion Dollar App Industry. Tyson Press, Chester, UK (2014).
39. Dehejia, R.H., Wahba, S.: Propensity Score-Matching Methods for Nonexperimental Causal Studies. Review of Economics and Statistics. 84, 151–161 (2002).
40. Bertrand, M., Duflo, E., Mullainathan, S.: How Much Should We Trust Differences-in-Differences Estimates? The Quarterly Journal of Economics. 119, 249–275 (2004).
41. Yin, P.-L., Davis, J.P., Muzyrya, Y.: Entrepreneurial Innovation: Killer Apps in the iPhone Ecosystem. American Economic Review. 104, 255–59 (2014).
42. Kemerer, C.F., Slaughter, S.: An Empirical Approach to Studying Software Evolution. IEEE Transactions on Software Engineering. 25, 493–509 (1999).
43. Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K.: The Development and Psychometric Properties of LIWC2015. UT Faculty/Researcher Works. (2015).
44. Pahnke, E.C., Katila, R., Eisenhardt, K.M.: Who Takes You to the Dance? How Partners' Institutional Logics Influence Innovation in Young Firms. Administrative Science Quarterly. 1–38 (2015).
45. Banker, R.D., Davis, G.B., Slaughter, S.A.: Software Development Practices, Software Complexity, and Software Maintenance Performance: A Field Study. Management Science. 44, 433–450 (1998).
46. Barry, E.J., Kemerer, C.F., Slaughter, S.A.: Environmental Volatility, Development Decision, and Software Volatility: A Longitudinal Analysis. Management Science. 52, 448–464 (2006).
47. Kannan, K., Rahman, M.S., Tawarmalani, M.: Economic and Policy Implications of Restricted Patch Distribution. Management Science. 62, 3161–3182 (2016).
48. Telang, R., Wattal, S.: An Empirical Analysis of the Impact of Software Vulnerability Announcements on Firm Stock Price. IEEE Transactions on Software Engineering. 33, 544–557 (2007).
49. Bresnahan, T., Davis, J.P., Yin, P.-L.: Economic Value Creation in Mobile Applications. Stanford University Graduate School of Business Research Paper No. 14-08. (2014).

An IS Perspective on Omni-Channel Management along the Customer Journey: Development of an Entity-Relationship-Model and a Linkage Concept

Markus Heuchert[1], Benjamin Barann[1], Ann-Kristin Cordes[1], and Jörg Becker[1]

[1]Univeristy of Münster - ERCIS, Department, Münster, Germany
{heuchert; barann; cordes; becker}@ercis.uni-muenster.de

Abstract. The digital transformation forces organisations to increasingly embed technology to catch up with customer demands. The omni-channel approach is one recent trend that requires taking the customer's perspective and offering a consistent experience across channels and touchpoints. While this development clearly necessitates IT for implementation, past research primarily stems from the marketing domain. In this article, we present an entity-relationship-model and a linkage model that takes an IS perspective and thereby enables communication between marketing and IT.

Keywords: Omni-Channel Management, Entity-Relationship-Model, Customer Experience Management

1 Introduction

Every organisation faces the digital transformation and is challenged by utilizing new technologies in its business [1]. The rapid adoption of these new technologies in broad levels of the population made customers outpace companies. Especially, retailers need to catch up in order to satisfy customer demands and therefore stay competitive [2]. However, historically grown, complex and heterogeneous system landscapes impede substantial changes and make an integrated system very expensive, as changing requirements are mirrored continuously from the customer side.

One recent trend evoked by the digital transformation is "omni-channel retailing" [3]. It is defined as linking all retail channels to provide a superior customer experience (CE) along the customer journey. But while the term is used extensively, in most of the retail companies is no prospect of a realisation since the technological requirements are too ambitious [4]. Marketing experts are required to understand the customer's needs and IT experts have to implement the needed technology, respectively. This cannot be done independently of each other nor subsequently but needs to be a joint effort between the two parties. However, a unified understanding from both perspectives and a basis for communication are missing.

Stemming primarily from the marketing literature, the customer experience is seen as a prime factor influencing the buying decision. As it is seen as a composition of rather soft factors [5], the incorporation of this research stream in technology is

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

challenging. Even the smart retail domain, which presents industry-specific technical opportunities, mainly focusses on customer acceptance, adoption, and experience [33]. While there are first attempts to facilitate decision support [e.g. 19, 42], missing cross-functional communication is another roadblock for the implementation of these technologies – especially if it comes to the linkage of different channels and touchpoints [9, 10]. It is the raison d'être of the IS discipline to foster the missing exchange of business and IT. Therefore, the goal of this article is to create IT artefacts that foster the communication between the different areas. This article aims to answer the research questions: (1) How can the terms of the customer experience management be structured in order to provide a common understanding about the terms and their relations and (2) how can the interrelation of the terms of customer touchpoint integration be defined to support the understanding and the strategy development.

To come up with the research results, a literature review has been performed according to Webster and Watson [11]. The identification of relevant articles is based on keyword search in the field of marketing, retail and service science. In total, 48 publications have been identified as relevant for the purpose of this article. As a post-processing step related to the keyword search, the main terms regarding the customer experience management (customer experience, customer journey and customer touchpoints) and touchpoint integration (multi-, cross-, and omni-channel integration as well as integration, linkage, and connectivity) have been identified. These build the foundation for answering the research questions by discussing the synonyms and defining the terms and their interrelations resulting in IT artefacts which provide a common understanding for the IS and marketing perspective.

The remainder of this paper is structured as follows. First, in section two the research background is introduced by explaining the underlying concepts of customer experience, the customer journey, as well as channels and touchpoints. Afterwards, in section three the first IT artefact – an entity-relationship-model – is deduced as the first step to structure terms of the domain and their connection to each other. In section four the second IT artefact – an omni-channel (OC) linkage concept – is created. It disambiguates the terms of (cross-/omni) channel and touchpoint integration by making use of the systems thinking theory. Finally, section six concludes the article by summarising the results and giving an outlook.

2 Research Background

2.1 Customer Experience

The importance of experiences in the social and economic sphere have initially been stressed by Pine and Gilmore [12]. More recently, the notion of customer experience has become a buzzword in marketing and its perfection a principal management objective [5]. To enable companies to empathise their customers, it is necessary to adopt a customer-centric view. Furthermore, the service-dominant logic [13] serves as theoretical underpinning, so that goods are merely seen as the vehicle to transport value-creating services to customers. As an experience is always subjective, its unfolding must depend on both the customer's perception, as well as the delivered service by the company. Hence, CE is co-created between the customer and the company, but the company's services can be designed with respect to CE [14].

The constituents of the CE seen in the academic literature are discussed by Lemon and Verhoef [5], who define CE as "a multidimensional construct focussing on a customer's cognitive, emotional, behavioural, sensorial and social responses to a firm's offering during the customer's entire purchase journey" [5]. The listing emphasises the influence of "soft" factors that are hard to measure and thus hard to manage.

With CE becoming a hegemon in service design, multidisciplinary teams consisting of marketing (business) and IT personnel become necessary for including the right ideas and implementing them, respectively [15]. It has always been the role of the IS discipline to facilitate the cooperation of both parties. While marketing is the prime driver, the feasibility of implementation is mostly ignored. As of now, the literature does not offer a unified view capable of including both parties.

2.2 Customer Journey

CE as a whole is a complex construct. One approach to reduce this complexity is to model CE separately at direct or indirect points of contact between customer and company. This idea is manifested in the notion of the "customer journey" (CJ) (or customer corridor [16], customer decision process, purchase journey [5]). A CJ consists of a series of touchpoints to the customer [17]. While it is admitted that the whole CE is more than the sum of the CE at the touchpoints [18], this simplification realises a process view and attribution of its touchpoint elements. The time-logical sequence around a purchase can be structured into successive stages relating to theories of consumer decision-making [19] or behaviour [20]. Attribution increases the detail level at touchpoints and diminishes abstraction, e.g., by displaying the used channel that shows media disruption while keeping the customer's perspective.

By using personas (i.e., a documented set of archetypal people, who are involved with a product [21]) or empirical data, so-called CJ maps can be created to describe the as-is path of a customer and subsequently improve it. Being named a good communicator and visualiser of business activities, the question arises how such modelling can be conducted. While there are various approaches, especially in grey literature, sound modelling techniques are sparse. The "Multi-Level Service Design" approach [14] is found promising, yet CE itself is not explicitly part of the model. Practitioner's approaches lack a defined syntax and can be summarised as superficial.

2.3 Customer Touchpoints and Channels

A touchpoint (TP) is similar to a service encounter [22] when considering interactions with the company. However, factors outside the company's control also shape CE [23] and hence must be included. Brand-, partner-, customer-owned and external TPs can be distinguished [5]. Consensus exists that the importance of TPs in relation to the overall CE varies, which has been conceptualised in "moments of truth" [24].

TPs belong to a specific channel. Neslin et al. [25] define channels as the medium, through which the customer and the firm interact. Broadening this understanding by including customer-to-customer interactions builds the working definition for the article. In former times customers were staying in a single channel, but the digital age has transformed their decision-making: Customers may search information through one channel, purchase through another and retrieve the product through a third channel [10, 24]. Thus, the importance shifts from channels to TPs [28], as a TP provides the asked service and channels are only utilised selectively at the respective TPs instead of broadly. TPs need to be connected and integrated for a superior CE.

3 Customer Experience Management Entity-Relationship-Model

Due to several streams of the research being conducted in separate fields, different notions have emerged. While publications in marketing, retail and service science make use of several central notions, the ties between the constructs remain blurry. At most, specific relationships are explained. By analysing these indications, an extended entity-relationship-model (ERM) based on [29], shown in Figure 1, is created. The extension allows specialisation using the triangle element. The model highlights the degree of influence of the marketing and technical view, as well as their intersection. The IS perspective is seen as their union. Observable objects are colour-coded to contrast to theoretical constructs. Entity types are *highlighted* in the following paragraphs.

Fostering a superior CE must be the overall goal of a company to assert itself against its competition. By definition, an experience is subjective so that it cannot be referenced multiple times and every *CJ* has one unique experience. *Customer Experiences* can influence each other and thus are related.

A *journey* entails a *reference object*, which is typically a product or service of the company that satisfies a customer's need. With "customer decision process" [5] being used synonymously to CJ, a time-logical order can be applied, which is seen in different consecutive phases around a customer's purchase. Both academic and practice literature provides a plethora of models [18, 30, 31] that can be aggregated to at least three *customer journey stages*: pre-sales, sales, after-sales.

A TP links theoretical considerations of CE with an actual event so that they are located between the marketing and technical view. A *CJ stage* encompasses *TP classes*, and a *TP class* is always related to a *channel*. By employing the logic of abstract classes and instantiations thereof, the *TP class* is defined to be an abstract interaction interface to the customer. The *TP instance* describes an actual episode of contact and is thus being situated in the *journey* and including an *experience*. Consequently, *TP instances* as chronological events have a sequence. An illustrative example is the Facebook wall of the company as the *TP class* with a customer post as the *TP instance*. This *TP instance* is the successor of a prior *TP instance* in the journey.

A TP class is specialised in four dimensions. First, TP classes can be critical (moments of truth) and non-critical [24]. Second, TP classes can be divided into four distinct types [5]. Additionally, Straker et al. [9] provide a typology, which is adapted into the remaining two specialisations of the TP class entity type: Simplex and duplex TPs signify the direction of communication, while the purpose of each TP is split into being a functional (comprising a clear objective by a company or customer) or diversion TP (enabling recreational and social activities by customers).

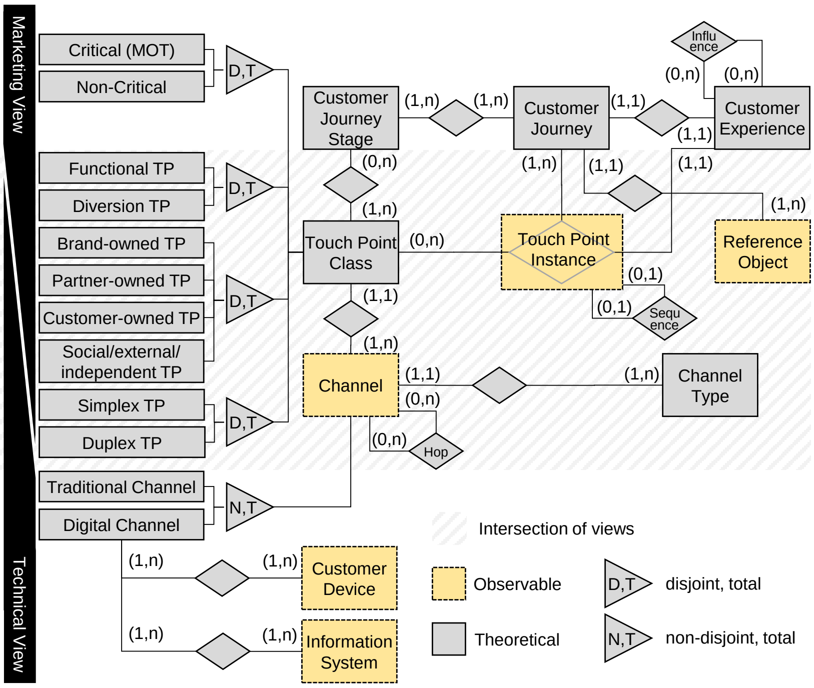


Figure 1. Customer Experience Management Entity-Relationship-Model

The *channel*, which can metaphorically be seen as the shell around TPs, is important from a technical perspective, as all TPs within a channel share technological characteristics that determine service design and issues of integration. Also, a company first needs that shell to establish TPs within. Customers on the other hand principally decide on using a channel before considering specific TPs. While the latter can serve a particular purpose, channels themselves do not. Consider an electronics retailer: First, the decision is made whether to check out the retailer's Facebook presence or a local store (channel) for obtaining product information. Then, after choosing the latter, one walks into the outlet and approaches the service counter or respective department (TP). While it is common in the retail domain to interpret a channel as a sales channel, communication channels are also subsumed here.

The advent of the internet introduced new channels and enabled dynamics so that abstract *channel types* are suitable means to capture fundamental similarities regardless of manifestations: Facebook can be the channel and social media its channel type.

A more technical view is realised through the non-disjoint specialisation of *channels* by putting emphasis on digital technologies. Customers need a *device* (e.g. smartphone) to use a *digital channel*, and the company requires an information system to serve the channel. As a homogenous system landscape fosters an easier channel connectivity, it is desirable to serve multiple channels with the same system.

First and foremost, the utility of this artefact is seen in a structured view on the different notions that arise in the domain. An instantiation of the entities for a specific company can be conducted and then used to define a coherent OC management approach within a digital transformation strategy. While the ERM shows the possible relationships between channels and touchpoints, in reality, these are rarely exploited by systems. The following investigates channel management approaches to get a holistic view of OC management innovation projects.

4 Omni-Channel Linkage Concept

4.1 An Introduction to Multi-, Cross- and Omni-Channel Management

Companies are increasingly introducing new channels and TPs that can be accessed from various devices independent of time and space. In this context, the channel management concepts multi-, cross- and omni-channel have emerged. However, these often have a blurred meaning in academic literature [2, 28]. For this reason, existing literature is first investigated in order to clarify these terms.

Retailers introduce several channels to allow the customer to choose a preferred channel in each stage of the CJ. However, the channels are organised autonomously and managed separately without an overlap or integrated objectives [28, 32]. Multi-channel retailers have silo structures and a lack of strategy [9, 10, 33]. According to Beck & Rygl [2], a multi-channel retailer sells "merchandise or services through more than one channel or all widespread channels, whereby the customer cannot trigger channel interaction and/or the retailer does not control channel integration" [3].

Beck & Rygl [2] consider channel interactions and integration as the distinctive characteristic of a cross-channel environment. They explain that a cross-channel retailer is a multi-channel retailer, who controls the integration of at least two channels or whose customers can trigger the interaction between at least two of these. Channel integration can be understood as technical back-stage connectivity, which enables information exchange between the different channels (e.g. customer, pricing and inventory data). Channel interaction enables new pathways through the CJ across channels such as the possibility to buy online and pick up offline [2]. Besides, a cross-channel retailer establishes managerial channel synergies through well-coordinated channel objectives, design, and deployment that are planned across multiple business functions, which result in benefits for the customer [5, 34]. To reach this, companies need to integrate business functions and collaborate with external partners [5].

This state is sometimes already called an OC environment [28]. Indeed, cross-channel is an inherent part of this environment [32]. However, in an OC setting, the borders between the channels disappear and a seamless channel switch is possible [28]. Besides, the focus is no longer on the channels, but instead on the distinct customer TPs within these channels [28]. Content-wise OC retailers try to establish thematic cohesion and consistency between all TPs. Furthermore, from a technical perspective connectivity and context-sensitivity between TPs are established [35]. Thereby, many variable customer value-adding journeys are enabled [32]. This increases the importance of CE management across multiple TPs [5]. Following Beck & Rygl [2] and Verhoef et al. [28], OC retailing is defined as: The selling of merchandise or services through all widespread channels, which have been seamlessly linked together from a customer perspective by enabling full interaction between all channels and/or from a retailer perspective by establishing full control of the channel integration. Through the systematic management of TPs and coordination of processes as well as technologies across these, the CE and performance across channels are optimised.

4.2 Understanding Channel Integration

In the context of cross- and omni-channel management, several ways exist how channels can be seamlessly linked. Most of the existing literature discusses these approaches under the terms multi-, cross- or omni- channel integration. While channel integration can be seen from a marketing perspective (i.e. thematic cohesion, consistency and context sensitivity of TPs), there is also a technique perspective (i.e. connectivity), which supports or facilitates the marketing perspective [35]. Therefore, channel integration is defined as: A company's efforts to synchronise its marketing efforts across all channels to optimise the customers seamless shopping experience [34, 36] and to functionally integrate these channels for seamless transition and interchangeability across the different stages of the CJ [34, 37, 38].

Saghiri et al. [32] list seven integration routines, which summarise how channels can be integrated: Integrated promotion, integrated transaction, integrated product information, integrated pricing information, integrated order fulfilment, integrated reverse logistics, and integrated customer service.

By considering these integration routines and by increasing channel information visibility (i.e. exchange of data) across channels and journey stages, a company can enhance towards a fully integrated OC system [32]. Channel integration can result in a stronger sales growth [5, 34], an increase in the "perceived quality of the online channel" [5] and the reduction of service inconsistencies [37]. Additionally, channel integration can achieve synergies such as "improved customer trust, improved customer awareness, consumer risk reduction, and coverage of diverse shopping preferences" [39]. Furthermore, from an operational perspective, it allows companies to actively maintain customer contacts and develop a proactive CE management strategy through increased customer insights [34, 40]. Besides, adaptive digital TPs enable new forms of digital marketing [41]. For example, by introducing firm-initiated mobile TPs, retailers can "provide tailored, time-sensitive, and location-sensitive advertising and promotions in store" [5]. This is further supported by the increased data integration and analysis abilities [34]. Finally, the interconnection of channels makes it harder for competitors to imitate the company. It could increase the customer's value proposition and thus reduces the competitive pressure [33, 37].

4.3 Deduction of an Omni-Channel Linkage Concept

As mentioned in the previous subchapter, there are several possibilities how channels can be linked together. However, in an OC environment, the notion of channel integration can be extended to the broader idea of TP integration [5]. The digitalisation of TPs is an important aspect of reaching full channel integration. However, the topic of TP integration is quite novel. So, what is the difference between the channel and TP integration? When investigating these terms, it becomes apparent, that the term *integration* is used without a clear concept in multi-, cross- and omni-channel literature.

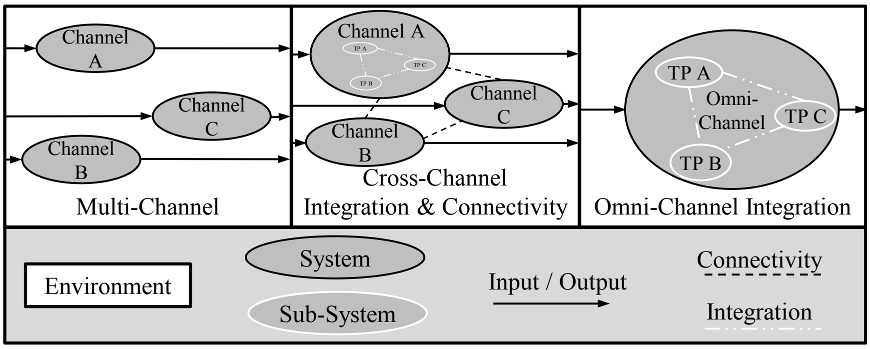


Figure 2. Multi-, Cross- and Omni-Channel from a System Thinking Perspective

According to the systems thinking theory, integration is the connection of components within a system. On the other hand, connectivity is the connection of components across different systems [42]. Figure 2 explains this idea in the context of channel management. In multi- and cross-channel environments, the channels itself can be seen as systems. While neither integration nor connectivity exists in a multi-channel environment, cross-channel integrations discussed in literature should be considered as channel connectivity instead. Only if this idea is extended to TPs, the connection of those within a certain channel should be considered as TP integration. This is in line with Homburg et al. [35] who explains that connectivity of TPs is the functional integration of "multiple touchpoints across online and offline environments for seamless transitions between one and another". However, moving forward to an OC environment, one could argue that the concept of channels does not exist anymore as the channel borders disappear [28]. In this case, only one OC system exists that consists of several TPs. If TPs are seamlessly linked in this environment, it could be considered as TP or OC integration.

To get around the ambiguity of the term *integration* and get a holistic idea of how TPs can be integrated and connected within an OC environment, the concept of OC linkage is deducted. Based on the idea of horizontal (i.e. "over time" [43]) and vertical integration (i.e. "coherence within each stage" [43]) and through the analysis of the touchpoints of a top 25 retailer by revenue worldwide [44], four different forms of OC linkage are identified along two dimensions. While the first deals with cross-stage integration, the second deals with the previously mentioned cross-channel connectivity (see Figure 3).

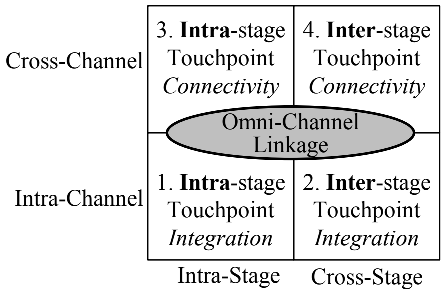


Figure 3. Omni-Channel Linkage Concept

1. Channel- and stage-internal *integration* is the first form of linkage in an OC environment. This type of integration deals with TP from a single channel that provide complementary information, are integrated from a process perspective and are aware of the interaction with others or itself historically within a specific stage. One example would be the possibility to add products to an online shopping cart from a digital cooking recipe. With smart retailing, this type of integration is also enabled in an offline channel and improves the in-store organisation [6, 45]. One example would be the usage of beacons and smartphones to enable context- or location-sensitivity [5, 8, 35]. Thus, this type of integration puts more emphasis on each single TP.

2. Channel-internal and stage-crossing *integration* is the second form of OC linkage. Saghiri et al. [32] explain that integration among channel stages ensures a smoother CE through context awareness that is passed on between stages. This type of integration is already quite common in online environments, where data is stored in user profiles and is analysed in order to provide i.e. product recommendations based on previous purchases. An easier example would be a service hotline that gives advice to the customer, places an order and, once the product arrives, offers after-sales support. Using smart retail technology, this could be transferred to an offline environment as well. E.g. smart shelf technology could be used to enable special promotions based on the customer's previous offline purchases [8].

3. Channel-crossing and stage-internal *connectivity* happen when TPs of different channels are connected to each other within a stage of the CJ in a cross-channel environment. Here, TPs of other channels provide complementary information for the same stage and are aware of each other. Furthermore, channels may be interchangeable and therefore allow a seamless switch [32]. One example would be the purchase of goods via an app and the pick-up in a local store (i.e. integrated order fulfilment).

4. Channel- and stage-crossing *connectivity* is the last form of OC linkage and combines the last two types of linkage. One example would be that a call centre agent first consults a customer and then adds products to his online shopping cart. The customer can then complete the purchase in the online store. Another example in a smart retail environment would be the provisioning of product recommendations at an arbitrary online TP after interacting with a specific offline TP.

The presented linkage concept does not only enable a consistent understanding of the terms involved in the digitalisation and integration of customer touchpoints but also delivers a holistic view of the technical possibilities for strategy development.

5 Conclusion and Outlook

Research on OC management is just starting to emerge [28]. The goal has been to contribute to the body of knowledge by shifting the view away from marketing concepts and viewing OC management from an IS perspective that combines marketing and technical views. The developed ERM provides a solid ground for future research by embedding constructs of the domain that are currently used frivolously. Furthermore, through the clear differentiation of TP integration and connectivity and by having a closer look at the directions in which TPs can be linked, IT and marketing gain another instrument to better facilitate their communication.

The ongoing research should evaluate and refine the presented artefacts through expert interviews and establishes a starting point for further research. Following the BPM life-cycle [46], it is planned to develop further IT artefacts. First, for the *design* of OC services, a modelling language will be created that covers a marketing and a technical perspective supporting cross-functional communication. Second, by having a closer look at the socio-technical characteristics, the business requirements have to be analysed that are necessary to implement (i.e. *model*) the proposed services (i.e. strategy, structure, culture, processes). Furthermore, the impacts of digitalised TPs on the operations of a retailer will be investigated. Based on the results, it is expected to deliver tangible structures and measures to facilitate decision-making. Third, implementing an IT artefact for the analysis, and presentation of process related information gathered at the different TPs would support the *execution* of the OC services. This requires the embedding of logging and monitoring mechanisms into digitalised TPs. Finally, these can drive the *monitoring* and *optimisation* of the CJ.

6 Acknowledgements

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant agreement No 645751.

References

1. Hess, T., Benlian, A., Matt, C., Wiesböck, F.: Options for Formulating a Digital Transformation Strategy. MIS Q. Exec. 15, 123–139 (2016).
2. Beck, N., Rygl, D.: Categorization of multiple channel retailing in Multi-, Cross-, and Omni- Channel Retailing for retailers and retailing. J. Retail. Consum. Serv. 27, 170–178 (2015).
3. Rigby, D.: The Future of Shopping, http://cormolenaar.nl/wp-content/uploads/2012/02/The-Future-of-Shopping-Harvard-Business-Review1.pdf, (Accessed: 11.11.2017).
4. Hillebrand, R., Finger, L.: Einkaufen in der Zukunft: Wie die Digitalisierung den Handel verändert. In: Digitales Neuland. pp. 89–101. Springer Fachmedien Wiesbaden, Wiesbaden (2015).

5. Lemon, K.N., Verhoef, P.C.: Understanding Customer Experience Throughout the Customer Journey. J. Mark. 80, 69–96 (2016).
6. Roy, S.K., Balaji, M.S., Sadeque, S., Nguyen, B., Melewar, T.C.: Constituents and consequences of smart customer experience in retailing. Technol. Forecast. Soc. Change. (2016).
7. Willems, K., Smolders, A., Brengman, M., Luyten, K., Schöning, J.: The path-to-purchase is paved with digital opportunities: An inventory of shopper-oriented retail technologies. Technol. Forecast. Soc. Change. (2016).
8. Inman, J.J., Nikolova, H.: Shopper-Facing Retail Technology: A Retailer Adoption Decision Framework Incorporating Shopper Attitudes and Privacy Concerns. J. Retail. 93, 7–28 (2017).
9. Straker, K., Wrigley, C., Rosemann, M.: Typologies and touchpoints: designing multi-channel digital strategies. J. Res. Interact. Mark. 9, 110–128 (2015).
10. Oh, L., Teo, H., Sambamurthy, V.: The effects of retail channel integration through the use of information technologies on firm performance. J. Oper. Manag. 30, 368–381 (2012).
11. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii–xxiii (2002).
12. Pine, B.J., Gilmore, J.H.: The Experience Economy: Work Is Theatre & Every Business a Stage. Harvard Business School Press, Boston (1999).
13. Vargo, S.L., Lusch, R.F.: Evolving to a New Dominant Logic for Marketing. J. Mark. 68, 1–17 (2004).
14. Patrício, L., Fisk, R.P., Falcão e Cunha, J., Constantine, L.: Multilevel Service Design: From Customer Value Constellation to Service Experience Blueprinting. J. Serv. Res. 14, 180–200 (2011).
15. Teixeira, J., Patrício, L., Nunes, N.J., Nóbrega, L., Fisk, R.P., Constantine, L.: Customer experience modeling: from customer experience to service design. J. Serv. Manag. 23, 362–376 (2012).
16. Meyer, C., Schwager, A.: Understanding Customer Experience. Harv. Bus. Rev. 1–11 (2007).
17. Zomerdijk, L.G., Voss, C.A.: Service Design for Experience-Centric Services. J. Serv. Res. 13, 67–82 (2010).
18. Dhebar, A.: Toward a compelling customer touchpoint architecture. Bus. Horiz. 56, 199–205 (2013).
19. Engel, J.E., Kollat, D., Blackwell, R.D.: Consumer Behavior. Dryden Press, New York (1978).
20. Howard, J.A., Sheth, J.N.: The theory of buyer behavior. Wiley, New York (1969).
21. Saffer, D.: Designing for interaction : Creating Innovative Applications and Devices. New Riders (2010).
22. Bitner, M.J., Ostrom, A.L., Morgan, F.N.: Service blueprinting: A practical technique for service innovation. Calif. Manage. Rev. 50, 66–94 (2008).
23. Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M., Schlesinger, L.A.: Customer Experience Creation: Determinants, Dynamics and Management Strategies. J. Retail. 85, 31–41 (2009).
24. Carbone, L.P., Haeckel, S.H.: Engineering Customer Experiences. J. Mark. Manag. 3, 18–19 (1994).
25. Neslin, S.A., Grewal, D., Leghorn, R., Shankar, V., Teerling, M.L., Thomas, J.S., Verhoef, P.C.: Challenges and Opportunities in Multichannel Customer Management. J. Serv. Res. 9, 95–112 (2006).

26. Dholakia, U.M., Kahn, B.E., Reeves, R., Rindfleisch, A., Stewart, D., Taylor, E.: Consumer Behavior in a Multichannel, Multimedia Retailing Environment. J. Interact. Mark. 24, 86–95 (2010).
27. Magnini, V.P., Karande, K.: Understanding consumer services buyers based upon their purchase channel. J. Bus. Res. 64, 543–550 (2011).
28. Verhoef, P.C., Kannan, P.K., Inman, J.J.: From Multi-Channel Retailing to Omni-Channel Retailing. J. Retail. 91, 174–181 (2015).
29. Vossen, G.: Datenmodelle, Datenbanksprachen Und Datenbankmanagementsysteme (German Edition). Oldenbourg Wissensch.Vlg (2008).
30. Puccinelli, N.M., Goodstein, R.C., Grewal, D., Price, R., Raghubir, P., Stewart, D.: Customer Experience Management in Retailing: Understanding the Buying Process. J. Retail. 85, 15–30 (2009).
31. Edelmann, D.C.: Branding in the Digital Age: You're Spending Your Money in All the Wrong Places, https://hbr.org/2010/12/branding-in-the-digital-age-youre-spending-your-money-in-all-the-wrong-places (Accessed on: 01.11.2017)
32. Saghiri, S., Wilding, R., Mena, C., Bourlakis, M.: Toward a three-dimensional framework for omni-channel. J. Bus. Res. 77, 53–67 (2017).
33. Herhausen, D., Binder, J., Schoegel, M., Herrmann, A.: Integrating Bricks with Clicks: Retailer-Level and Channel-Level Outcomes of Online-Offline Channel Integration. J. Retail. 91, 309–325 (2015).
34. Cao, L., Li, L.: The Impact of Cross-Channel Integration on Retailers' Sales Growth. J. Retail. 91, 198–216 (2015).
35. Homburg, C., Jozić, D., Kuehnl, C.: Customer experience management: toward implementing an evolving marketing concept. J. Acad. Mark. Sci. 45, 377–401 (2017).
36. Berger, P.D., Bolton, R.N., Bowman, D., Briggs, E., Kumar, V., Parasuraman, A., Terry, C.: Marketing Actions and the Value of Customer Assets. J. Serv. Res. 5, 39–54 (2002).
37. Jiang, K., Xu, L., Bao, X.: The impact of channel integration on channel reciprocity in the multi-channel retailing context. In: International Conference on Industrial Engineering and Engineering Management (IEEM). pp. 1840–1844. IEEE (2015).
38. Yan, R., Wang, J., Zhou, B.: Channel integration and profit sharing in the dynamics of multi-channel firms. J. Retail. Consum. Serv. 17, 430–440 (2010).
39. Wu, I.-L., Wu, S.-M.: A strategy-based model for implementing channel integration in e-commerce. Internet Res. 25, 239–261 (2015).
40. Grewal, D., Levy, M., Kumar, V.: Customer Experience Management in Retailing: An Organizing Framework. J. Retail. 85, 1–14 (2009).
41. Kannan, P.K., Li, H. "Alice": Digital marketing: A framework, review and research agenda. Int. J. Res. Mark. 34, 22–45 (2017).
42. Matook, S., Brown, S.A.: Characteristics of IT artifacts: a systems thinking-based framework for delineating and theorizing IT artifacts. Inf. Syst. J. 27, 309–346 (2016).
43. Adam Richardson: Touchpoints Bring the Customer Experience to Life, https://hbr.org/2010/12/touchpoints-bring-the-customer (Accessed: 01.11.2017)
44. Dhiraj, A.B.: Giants Of Retail: The World's 25 Largest Retailers By Revenue, 2017, http://ceoworld.biz/2017/01/24/giants-retail-worlds-25-largest-retailers-revenue-2017/ (Accessed: 01.11.2017)
45. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: A survey. Comput. Networks. 54, 2787–2805 (2010).
46. Houy, C., Fettke, P., Loos, P.: Empirical research in business process management – analysis of an emerging field of research. Bus. Process Manag. J. 16, 619–661 (2010).

Kundennutzen von VR-basierten 360° Panoramen für den Erwerb beratungsintensiver Güter und Dienstleistungen: Eine Case Study im Garten- und Landschaftsbau

Florian Remark[1], Lisa Berkemeier[1], Oliver Thomas[1]

[1] Informationsmanagement und Wirtschaftsinformatik, Universität Osnabrück, Germany
{florian.remark, lisa.berkemeier,oliver.thomas}@uni-osnabrueck.de

Abstract. Virtual Reality (VR) steigert die Teilhabe und Autonomie des Kunden im Kaufprozess von ortsgebundenen und beratungsintensiven Gütern und Dienstleistungen. Im Rahmen einer Case Study wird der Nutzen von Virtual Reality Technologien im Garten- und Landschaftsbau untersucht, um zu zeigen, dass mit diesen immersiven Szenarien die Planung eines Projektes unterstützt wird und eine Verbesserung gegenüber konventionellen Präsentationsmedien in der Entscheidungsfindung besteht. Die Anforderungsanalyse, prototypische Implementierung und Evaluation (N=61) eines VR-Systems, welches die Entscheidungsbasis des Kunden fundamental erweitert, tragen zur Wissensbasis bei. Ein praktischer Beitrag entsteht durch einen unternehmensspezifischen Prototyp, der sowohl branchenintern, als auch –übergreifend einfach auf andere Unternehmen adaptiert werden kann.

Keywords: Virtual Reality, VR, beratungsintensive Dienstleistungen, Case Study Research

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

Digitale Wertschöpfung durch Crowd Services: Neue Formen des Kundensupports am Beispiel Mila und Swisscom

Volkmar Mrass[1], Christoph Peters [1,2]

[1] Universität Kassel, Fachgebiet Wirtschaftsinformatik, Kassel, Deutschland
{volkmar.mrass,christoph.peters}@uni-kassel.de
[2] Universität St.Gallen, Institut für Wirtschaftsinformatik, St.Gallen, Schweiz
christoph.peters@unish.ch

Abstract. Die digitale Transformation verändert gegenwärtig die Art und Weise, wie Arbeit erbracht und organisiert wird. Bestehende Geschäftsmodelle verlieren an Attraktivität, neue Geschäftsmodelle kommen hinzu. Zwei Trends sind dabei besonders signifikant: Die „Plattform-Ökonomie", mit Crowdworking-Plattformen als einer ‚Spielart', die in den letzten Jahren starken Zulauf erfahren hat, und die „Sharing-Economy". In diesem Beitrag zeigen wir am Beispiel der Crowdworking-Plattform Mila und ihres Kunden und 51-prozentigen Eigentümers Swisscom AG, wie Unternehmen die Crowd für Wertschöpfung einsetzen, ihren Kundensupport ausbauen und damit Vorteile für alle Beteiligten generieren können. Die Crowdworking-Plattform Mila verbindet in einem neuartigen Ansatz das Paradigma der Crowd, insbesondere die Nutzung einer Vielzahl von Leistungserbringern, mit den Prinzipien der Sharing-Economy, vor allem dem Einbezug der Kunden als Produzenten von Dienstleistungen. Basierend auf Experten-Tiefen-Interviews mit Verantwortlichen von Mila und Swisscom analysieren wir in diesem Beitrag dieses neuartige Arbeitssystem inklusive des Zusammenspiels aller Beteiligten und leiten daraus Handlungsempfehlungen für Unternehmen ab.

Keywords: Crowdworking-Plattformen, Crowd Services, Kundensupport, Plattform-Ökonomie, Sharing Economy

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany