

Sven Bittner, Stefan Hornbostel, Frank Scholze (Hg.)

FORSCHUNGSINFORMATION IN DEUTSCHLAND: ANFORDERUNGEN, STAND UND NUTZEN EXISTIERENDER FORSCHUNGSINFORMATIONSSYSTEME

WORKSHOP FORSCHUNGSINFORMATIONSSYSTEME 2011

iFQ-Working Paper No.10 | Mai 2012



Institut für
Forschungsinformation
und Qualitätssicherung

**Barbara Ebert, Alexander Kujath, Joachim Holtorf, Karsten Holmberg,
Thomas Rupp**

Erfahrungen aus der Einführung des Forschungsinformationssystems Pure an der Leuphana Universität Lüneburg

1 Einführung

Die Leuphana Universität Lüneburg betreibt seit 2011 ein integriertes Forschungsinformationssystem. Pure steht für „Publications and Research“ und führt die bisher fragmentierten Informationen über Projekte, Publikationen und Expertise der Universität zusammen. So soll eine einheitliche, zwischen Wissenschaft und Verwaltung abgestimmte Datenbasis entstehen, die die bisherige Praxis der dezentralen, teils redundanten Datensammlungen ersetzt und Administration wie Wissenschaft in der Forschungsberichterstattung entlastet. Die Einrichtung einer Forschungsdatenbank ist Teil der 2007 vom Senat verabschiedeten Forschungspolitik.

Die Leuphana Universität ging 2005 aus der Fusion der Fachhochschule Nordostniedersachsen und der Universität Lüneburg hervor. Mit 7.500 Studierenden, 155 Professorinnen und Professoren sowie 300 wissenschaftlichen Mitarbeiterinnen und Mitarbeitern gehört sie zur Gruppe der mittelgroßen Universitäten in Deutschland.

In den vergangenen sieben Jahren hat sich am Standort eine Modelluniversität entwickelt, die besonders in der Lehre Neuland beschreitet und mehrfach für ihre visionären Konzepte ausgezeichnet wurde. Das wissenschaftliche Profil der Leuphana bilden vier fachübergreifende, transdisziplinäre Wissenschaftsinitiativen, die auch den thematischen Rahmen für die vier im Oktober 2010 neu gegründeten Fakultäten abstecken: Bildung, Kulturwissenschaften, Wirtschaftswissenschaften und Nachhaltigkeit.

Ein wichtiger Meilenstein in der Neuausrichtung war 2009 die Genehmigung des EU-Großprojekts „Innovations-Inkubator Lüneburg“. Für sechs Jahre werden aus dem Europäischen Fonds für regionale Entwicklung sowie vom Land Niedersachsen Gesamtinvestitionen von 100 Mio. Euro zur Verfügung gestellt, um aus der Universität heraus in einem europaweit einmaligen Vorhaben die regionale Wirtschaftsentwicklung substanziell zu stärken.

1.1 Besondere Anforderungen an die Forschungsberichterstattung

Die fusionierte Universität hatte in den vergangenen Jahren eine Reihe von Herausforderungen zu bewältigen und stand bzw. steht im besonderen Fokus des Landes. Eines der Ziele, an denen der Erfolg der Neuausrichtung gemessen wird, ist die Entwicklung des akademischen Renommées und eine spürbare Steigerung der zentralen Kennzahlen für die Forschung. Der Forschungsberichterstattung kommt an der Leuphana daher eine besondere Bedeutung zu.

Forschung gehört neben der Lehre zu den wichtigsten Aufgaben der Universität. Informationen über Aufgabenerfüllung und Leistungen in diesem Bereich sollten unaufwändig und verlässlich zur Verfügung stehen. In der Realität sind diese Informationen jedoch über viele Stellen und Systeme verteilt – teils inneruniversitär, teils in externen Publikations- und Projektdatenbanken. Intern wurden vorab rund zwei Dutzend Prozesse identifiziert, die regelmäßig Angaben zu Publikationen, Projekten, wissenschaftlichen Aktivitäten, Kooperationen usw. erheben – sei es auf Organisationsebene (Internetseiten, Mittelvergabe), sei es thematisch (Fächerevaluationen des Landes Niedersachsen, Teilnahme an Rankings) oder individuell, wenn Wissenschaftlerinnen und Wissenschaftler Anträge einreichen oder persönliche Webseiten pflegen.

Die benötigten Informationen sind in der Regel ähnlich, jedoch sind sie aufgrund der individuellen

Formate oft nur mit großem Aufwand wieder verwendbar und durch die dezentrale Organisation schlecht verfügbar. Daraus resultieren unweigerlich Mehrfachabfragen und erhöhter administrativer Aufwand für die einzelnen Berichtsanlässe. Daher werden im Projekt Forschungsdatenbank vorrangig zwei Einzelziele verfolgt:

1. Schaffung einer einheitlichen, zwischen Wissenschaft und Verwaltung abgestimmten Datenbasis für die Forschungsberichterstattung
2. Schaffung von Mehrwerten für die Wissenschaftlerinnen und Wissenschaftler durch verbesserte Verfügbarkeit der Informationen für eigene Zwecke

In der Implementierungsphase wurden u.a. mehrere Tausend Publikationen und Hunderte von Autorinnen und Autoren der seit 2007 geführten Hochschulbibliographie in das neue System übernommen. Informationen über Organisationen und Personal sowie Drittmittel werden täglich aus dem SAP-System zur Verfügung gestellt. Wissenschaftlerinnen und Wissenschaftler steuern über ihr individuelles Nutzerkonto die Informationen über ihre Projekte, Publikationen und Aktivitäten. Zur Unterstützung können sie auf ein Netzwerk geschulter Redakteurinnen und Redakteure in den Instituten und Zentren zurückgreifen. Einmal registriert, stehen die Informationen regelhaft für interne Zwecke zur Verfügung, so dass die bislang praktizierten Mehrfachabfragen reduziert werden können. Nutzerinnen und Nutzer können für jeden Datenbankeintrag gesondert festlegen, ob er auch für den öffentlichen Forschungskatalog verfügbar sein soll.

Abbildung 1: Ziel der Datenbankeinführung – eine autorisierte Sammlung von Forschungsinformationen schaffen, Mehrfachmeldungen und redundante Datenhaltung verringern

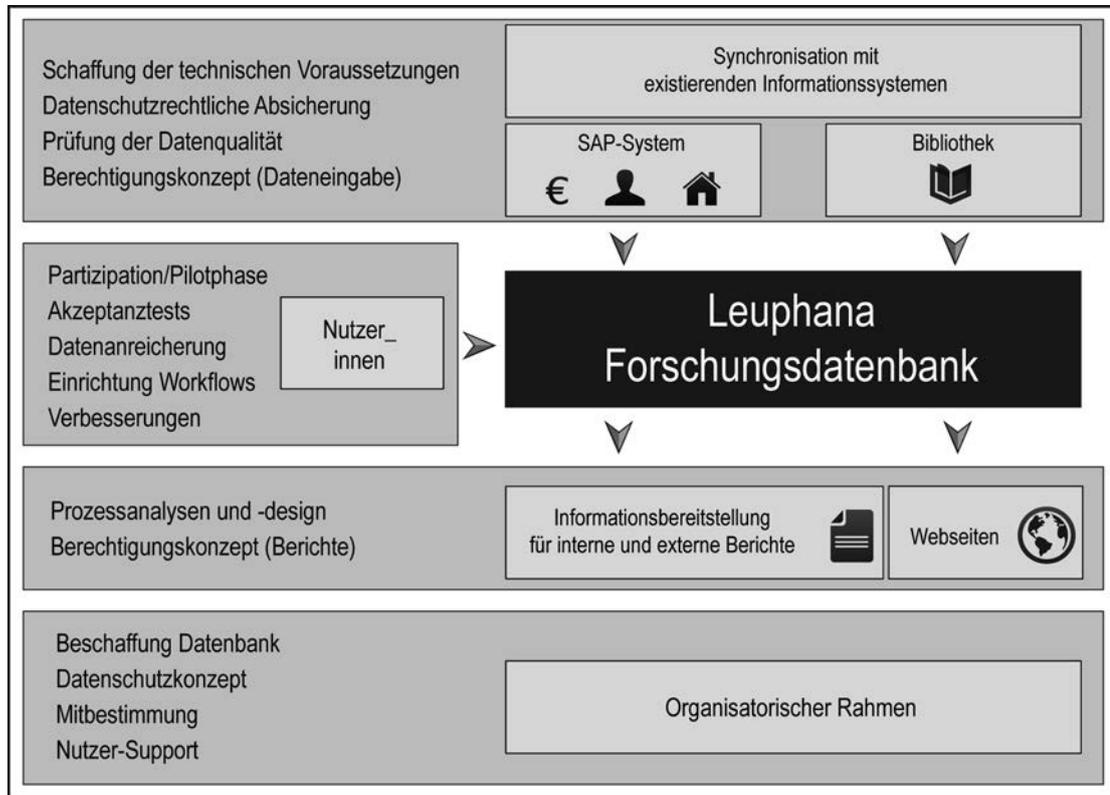


Für unseren Beitrag in diesem Tagungsband berichten wir auszugsweise über den bisherigen Verlauf des Projektes Forschungsdatenbank und ziehen eine vorläufige Bilanz auf Basis der gewonnenen Erfahrungen.

2 Das Datenbankprojekt

Eine Übersicht der Aufgabenpakete für die Einführung und Betrieb gibt Abbildung 2.

Abbildung 2: Projekt Forschungsdatenbank



2.1 Verwendetes System

Pure ist ein integriertes Forschungsinformationssystem (*Current Research Information System – CRIS*), in dem Expertenprofile, wissenschaftliche Aktivitäten, Projekte und Publikationen gepflegt werden. Diese werden verknüpft mit Datensätzen zu Personen (Autorinnen/Autoren, Projektleitungen etc.) und Organisationseinheiten der Leuphana. Das System übernimmt ausgewählte Informationen aus anderen internen Informationssystemen und wird dezentral durch berechtigte Nutzerinnen und Nutzer gepflegt. Das Datenmodell orientiert sich an dem europäischen Datenformat CERIF (Common European Research Information Format).

Pure wurde vor neun Jahren von dem dänischen Start-up Unternehmen Atira entwickelt. In Dänemark wird Pure auf Beschluss der Hochschulrektoren an allen staatlichen Universitäten eingesetzt und speist die nationale Forscherdatenbank. Neben der Leuphana gehören in Europa außerdem die Universität Helsinki, 17 Universitäten in Großbritannien sowie die flämische Regionalregierung zu den Kunden. Pure wird von Atira gemeinsam mit den beteiligten Forschungseinrichtungen laufend weiterentwickelt und ist dadurch konsequent an den Erfordernissen der Wissenschaft orientiert.

2.2 Das Ziel: Eine autorisierte Sammlung von Informationen schaffen

In Pure werden Informationen zu Forschungsleistungen aus mehreren Quellen zusammengeführt. Um Mehrfacheingaben und Unstimmigkeiten zu vermeiden, wurde Pure von Anfang an mit vor-

handenen Informationssystemen der Verwaltung synchronisiert: Ausgewählte Aufträge der Drittmittelverwaltung - SAP FI/CO

- Informationen über personelle Wechsel und organisatorische Veränderungen – SAP HR
- Hochschulbibliographie - PICA

Hochschulmitglieder melden sich aktiv zur Nutzung der Forschungsdatenbank an und geben damit ihre Einwilligung zur Datenverarbeitung. In Pure registrierte Hochschulmitglieder erhalten ein Nutzerkonto, so dass sie die mit ihrer Person verknüpften Daten einsehen und bearbeiten können. Zur Unterstützung können sie auf geschulte Redakteurinnen und Redakteure in den Instituten und Zentren zurückgreifen, die lokal mit der Pflege der Daten betraut werden.

2.3 Umgang mit personenbezogenen Daten

Die Natur eines Forschungsinformationssystems und der Anschaffungsgrund für Pure an der Leuphana Universität legen es nahe, dass nicht nur aggregierte oder öffentliche Informationen in der Forschungsdatenbank gespeichert und verarbeitet werden. Grundsätzlich werden alle in Pure eingetragenen Daten mit Personen und Organisationseinheiten verknüpft. Dies ermöglicht es beispielsweise dem wissenschaftlichen Personal in Pure ein persönliches Archiv der wissenschaftlichen Aktivitäten wie Vorträge, Posten und Ämter, Preise und dergleichen zu pflegen. Pure bietet eine Reihe technischer Funktionalitäten zur Kontrolle der eigenen Daten. Nutzerinnen und Nutzer können die Verwendung der personenbezogenen Einträge in Pure auf mehreren Ebenen steuern:

1. Umfang der Dateneingabe – registriert wird nur, was zur Berichterstattung vorgesehen ist.
2. Sichtbarkeitseinstellung für Einzeleinträge (entscheidet über die Verwendbarkeit des Eintrags):
 - Datenbank – nur für autorisierte Nutzerinnen und Nutzer sichtbar
 - Campus – Darstellung auf Intranet und Campus-Systeme beschränkt
 - Öffentlich – Nutzung für die Öffentlichkeitsarbeit, z.B. Leuphana Webseite

Über neue Einträge und Änderungen der mit ihrer Person verknüpften Daten werden die jeweiligen Nutzerinnen und Nutzer informiert. Dies geschieht auch, wenn die Nutzerinnen und Nutzer das Eintragen von Informationen an andere Personen delegieren.

2.4 Regeln für die Auswertungen von Daten

Um Mehrfachabfragen bei Forschenden und redundante Datenhaltung zu reduzieren, werden die in Pure registrierten Datensätze autorisierten Nutzerinnen und Nutzern für anerkannte und legitime Berichtsverfahren zur Verfügung gestellt. Art und Umfang der Berichte werden bestimmt durch die Beschlüsse, Richtlinien, Ausführungspapiere, Rechtsgrundlagen oder sonstige legitime Anforderungen der jeweiligen Verfahren. Ein Berichtsleitfaden legt fest, dass Nutzerinnen und Nutzer vorab über die Auswertungen zu informieren sind und ausreichend Gelegenheit zur Aktualisierung erhalten müssen. Personen mit Reporting-Rechten und ihre Vorgesetzten werden schriftlich auf die Einhaltung der Richtlinie verpflichtet.

Der Begriff „Bericht“ umfasst alle Darstellungen von Forschungsleistungen auf den verschiedenen Organisationsebenen der Leuphana Universität:

- Öffentlichkeitsarbeit, insbesondere
 - die Darstellung von Forschungsleistungen auf Webseiten
 - Tätigkeitsberichte
- Erstellung aggregierter Statistiken und Auswertungen für die Leitungsebenen der Universität
- Evaluation interner Fördermaßnahmen und –programme

- Leuphana Forschungspreise Erstellung von Forschungs- und Leistungsberichten für organisations- und fachbezogene Evaluationsverfahren

2.5 Arbeitsorganisatorische Auswirkungen

Während der Datenbankeinführung entsteht projektbezogene Mehrarbeit a) für die Organisation der Datenbereitstellung, b) für die Entwicklung der Dokumentation und c) für die Organisation der Prozesse zum Betrieb des Systems. Im Forschungsservice der Leuphana, der die System-einführung koordiniert, wurde zur Unterstützung der Projektleitung eine temporäre Stelle geschaffen und mit einem Informatiker besetzt. Für den Regelbetrieb entstehen durch die neu eingeführte Qualitätssicherung (Vier-Augen-Prinzip zur Validierung von Projekten und Publikationen) in Bibliothek und Drittmittelbetreuung neue Aufgaben. In den Instituten, Arbeitsgruppen und Zentren übernehmen v.a. die Sekretariate neue Aufgaben in der Datenpflege (Redakteursnetz, s. Abschnitt 4). In der Forschungsberichterstattung wirkt sich der Minderaufwand für das Erstellen von Statistiken und Berichten, bedingt durch bessere Datenverfügbarkeit, entlastend aus.

Tabelle 1 gibt eine Übersicht der beteiligten Organisationseinheiten und ihrer Aufgaben.

Tabelle 1: Übersicht der internen Schnittstellen und ihrer Aufgaben

Organisation	Pilotphase	Regelbetrieb (Neue Aufgaben)
Forschungsservice	Projektkoordination, Schulung, Dokumentation, Anschaffung	Datenbankadministration, Nutzersupport, Qualitätssicherung
Universitätsbibliothek	Übertragung der PICA-Hochschulbibliographie, bibliographische Fachkompetenz, (manuelle) Nachpflege der Publikationsdaten	Qualitätssicherung durch Validierung der Publikationseinträge (Vier-Augen-Prinzip)
Personal/Finanzen	Begleitung der Schnittstellenimplementierung, Beratung bei Fragen, Kontrolle der importierten Daten	Qualitätssicherung der laufenden Synchronisation
Universitätsmarketing	Entwicklung der Webintegration von Forschungsdaten aus Pure (u.a. Typo3-Plugin)	Technische Betreuung der Webintegration, inkl. der persönlichen Webseiten und des Online-Forschungskatalogs, Weiterentwicklung von Webanwendungen
Medien- und Informationszentrum	Bereitstellung einer Systemumgebung	Technische Betreuung und Administration, Backups
Wissenschaftliche Bereiche	Mitwirkung einer Pilotnutzergruppe an der Erprobung des Systems	Datenpflege durch Sekretariate bzw. andere Mitarbeiterinnen und Mitarbeiter (Redakteursrolle), Datenerhebung und Reporting mit Pure (Dekanate)
Personal- und Organisationsentwicklung	Projektsupervision und Koordination mit Einführung SAP-HR Abschluss Dienstvereinbarung mit Personalrat	
Datenschutzbeauftragter	Beratung und Begleitung	
Personalrat	Beratung und Begleitung	

3 Implementierung und Anbindung in der IT-Landschaft der Universität

Für den Einsatz an der Leuphana Universität Lüneburg wurde das Pure Datenmodell durch Atira weiterentwickelt und an die lokalen Anforderungen angepasst. Nach der formalen Abnahme erfolgte unter Mitwirken des Medien- und Informationszentrums (MIZ) die grundständige Inbetriebnahme des Forschungsinformationssystems. Dem gesteckten Ziel „vorhandene Daten nutzen“ entsprechend wurden in der Implementierungsphase die vorab identifizierten Datenquellen für die Integration in die Pure Forschungsdatenbank vorbereitet und an die Datenbank angeschlossen (vgl. Abschnitt 2.2).

Aus dem SAP-System wurden die Organisationsstruktur der Leuphana mit 170 Organisationen sowie ein ausgewählter Bestand von rund 650 Drittmittelaufträgen übernommen. Für die Datenübernahme aus dem Bibliothekssystem PICA wurden rund 400 Personendatensätze von Autorinnen und Autoren mit ihren PICA-Normnummern in Pure erfasst. Damit gelang die Übernahme und Zuordnung von ca. 6.000 Publikationseinträgen der Hochschulbibliographie zu 95%. Die aus den verbundenen Systemen übernommenen Daten wurden durch vorab definierte Qualitätsstandards bewertet und der jeweilige Nachbearbeitungsbedarf formuliert.

Nachbearbeitung war insbesondere für Projekteinträge nötig, um die Datensätze für die Forschungsdatenbank zu erschließen, da die importierten SAP-Datensätze nur Finanzinformationen, eine Organisationszuordnung sowie eine für finanztechnische Zwecke abgekürzte Vorhabensbezeichnung enthielten.

Auch der Bestand an Publikationseinträgen in der Hochschulbibliographie musste nach der Übertragung in die Forschungsdatenbank nochmals nachbearbeitet werden, um Übertragungsfehler zu korrigieren und die Typisierung der Einträge als Monographie, Sammelwerke, Arbeits- oder Projektberichte anzupassen. Zuvor hatten bereits umfassende Arbeiten am Datenbestand in PICA stattgefunden, um die Datenübertragung vorzubereiten (vgl. Abschnitt 3.1).

3.1 Hochschulbibliographie: Von PICA zu Pure

Vor der Einführung von Pure als zentrale Publikationsdatenbank an der Leuphana Universität Lüneburg wurden an ihren Vorläufereinrichtungen bereits verschiedene Systeme für Hochschulbibliographien verwendet. Dabei handelte es sich um Eigenentwicklungen und OpenSource-Systeme, die nicht mit bibliographischen Standardformaten arbeiteten.

Bei Aufkommen des Wunsches nach einem einheitlichen System zur Speicherung von Hochschulschriften im Jahr 2007 wurde die Schaffung eines einheitlichen Datenmodells als wesentliche Voraussetzung für die Konsolidierung der unterschiedlichen Datenbestände erkannt. Dieses Ziel sollte über die Nutzung des lokalen Bibliothekssystems PICA erreicht werden. In PICA werden die Datensätze mithilfe von Normnummern und sogenannten Kategorien anstelle von Eingabefeldern organisiert. Damit wird das System flexibel gehalten und ist für die Erfassung eines breiten Spektrums an Medien geeignet. Die Standardisierung erfolgt durch ein komplexes Regelwerk für die Katalogisierung, das auch Spielraum für Auslegungen und lokale Nutzung einzelner Kategorien zulässt.

In 2007 wurden die Mitglieder der Hochschule aufgerufen, alle Publikationen ab dem Jahr 2000 an die Bibliothek zur Eingabe zu geben. Nach der Eingabe wurde ab August 2008 die Hochschulbibliographie der Universität in PICA veröffentlicht. Hauptanwendungszweck war die Präsentation der Publikationen für lokale Studierende und der Nachweis im Gesamtverbundkatalog. Darüber hinaus wurden die Daten auch zur Erstellung fakultätsbezogener Statistiken, u.a. für die leistungsbezogene Mittelvergabe und die Verleihung der Forschungspreise der Universität genutzt. Für diesen Zweck wurden zusätzlich benötigte Informationen wie Fakultätszugehörigkeit, Fächer, Dokumententypen, Verlagsinformationen, Rolle der Autoren und Medienformen in lokale Schlag-

wortkategorien des Bibliothekskatalogs aufgenommen. Der Export erfolgte über einen Webservice, der die Daten in einem standardisierten Format ausgab. Dieses konnte mit gängigen Tabellenkalkulationsprogrammen weiterverarbeitet werden.

Nach der Entscheidung der Universitätsleitung für die Einführung eines integrierten Forschungsinformationssystems wurde überlegt, die PICA-Hochschulbibliographie als Datenquelle für Pure zu verwenden, um den etablierten Prozess der Meldung und Erfassung von Forschungsschriften zu erhalten und von Anreicherungen der Datenbestände im Gesamtverbundkatalog profitieren zu können. Dafür wurde von Atira in Zusammenarbeit mit der Bibliothek eine Schnittstelle programmiert, um das PICA+ Format in das von Pure verwendete CERIF-basierte Datenmodell zu übersetzen. Über diese Schnittstelle sollte durch einen kontinuierlichen Synchronisationsprozess ein Abgleich von Neuerfassungen und Änderungen in der PICA-Datenbank mit den Daten in Pure durchgeführt werden.

Zudem wurde der auf 7.000 Literaturnachweise angewachsene Bestand der Hochschulbibliographie in PICA grundlegend überprüft und überarbeitet. Unter anderem wurden doppelt aufgenommene elektronische Ausgaben von Printwerken entfernt und mit den Schwestereinträgen zusammengeführt. Neben den Korrekturen wurden zahlreiche Attribute nachgepflegt, um die reibungslose Übertragung in die Forschungsdatenbank abzusichern.

Zur Übernahme der Publikationsdaten haben der Hersteller Atira und das Schnittstellenteam der Bibliothek erheblichen Aufwand investiert, um die Belegung der PICA-Kategorien zu dokumentieren und Regeln für die Datenübernahme zu entwickeln. Für viele Sonderfälle konnte nur unter zu Hilfenahme bibliothekarischer Fachkompetenz eine Lösung entwickelt werden.

Nach der initialen Übertragung der rund 6.000 verbliebenen Publikationseinträge von PICA nach Pure machten sich grundlegende Inkompatibilitäten zwischen dem bibliographischen Datenformat PICA+ und dem relationalen, an CERIF orientierten Datenmodell von Pure bemerkbar. Sie führten zu manuellen Nacharbeiten in erheblichem Umfang. Schlussendlich wurde deutlich, dass der Synchronisationsprozess den Anforderungen der Anwender nicht gerecht werden konnte.

Insbesondere zwei Attribute konnten über die Integration nicht aufgelöst werden. Eine ist die PICA-Kategorie "Umfang" einer Publikation, in der sowohl Seitenzahl, Seitenumfang als auch Dateigröße oder Gewicht eines Druckwerks erfasst werden. Für 18 Standardbelegungen dieser Kategorie wurden Integrationsregeln entwickelt. Trotzdem konnte eine korrekte Identifikation der Seitenangabe durch die Pure Integration nicht sicher gewährleistet werden. Die zweite Besonderheit betrifft Sammelwerke, bei denen in PICA nicht nur die Herausgeber, sondern auch der erste beitragende Autor erfasst sein können. Dies wird von den Verbundbibliotheken nicht einheitlich gehandhabt, und leider ist der Status der Personen anhand der Kategorienbelegung nicht eindeutig entscheidbar. Eine Berichtigung erfolgte manuell in Pure, war aber im Hinblick auf eine laufende Synchronisation problematisch.

Da für eine integrationsfähige Erfassung der Literaturnachweise in PICA letztendlich in beiden Systemen erheblicher Aufwand betrieben werden muss, entschied die Steuerungsgruppe sich für eine Ersterfassung mit allen für Statistiken notwendigen Attributen in Pure. Auf den Nachweis der Leuphana Hochschulbibliographie im Gesamtverbundkatalog (GVK) wird aktuell verzichtet, da die direkte Einbindung lokaler Systeme wie Pure in die GVK-Suche technisch nicht möglich ist.

3.2 Projekte, Organisationen und Personen – Anbindung an SAP

Für die Anbindung und Belieferung des Systems Pure mit personalwirtschaftlichen und projektspezifischen Daten aus dem SAP-System wurde das Niedersächsische Hochschulkompetenzzentrum für SAP (CCC) in das Projekt mit eingebunden. Das CCC unterstützt und betreut 18 niedersächsische Hochschulen beim Einsatz und der Weiterentwicklung des SAP-Systems in den

Bereichen Rechnungswesen, Personalwesen, Logistik, Business Warehouse und Folders Management (elektronische Drittmittelakte).

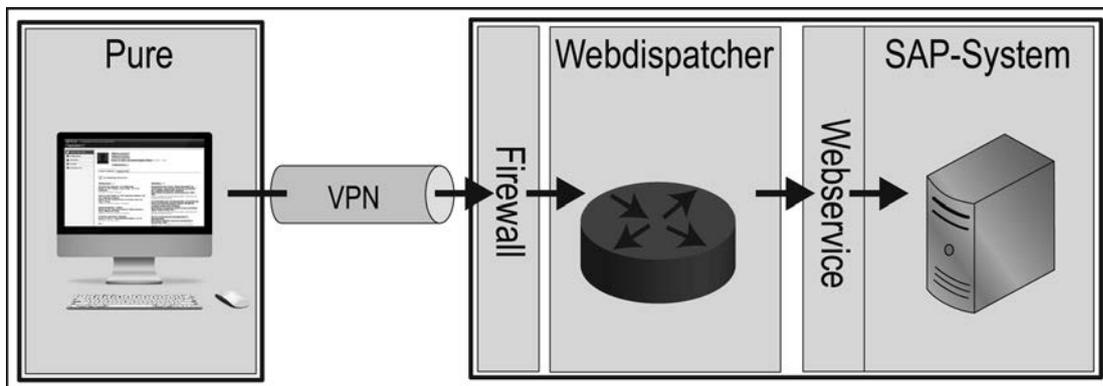
Im gemeinsam mit Universität und Datenbankanbieter entwickelten Fachkonzept wurden die zu übergebenen Datenstrukturen abgestimmt und die technische Infrastruktur festgelegt. Der Betrieb von SAP ist aufseiten der Leuphana durch eine Dienstvereinbarung geregelt. Zur Nutzung ausgewählter SAP-Daten für den Anwendungsfall Pure wurde an der Leuphana eine Ergänzung zur Dienstvereinbarung abgeschlossen.

Die Anforderungen aus dem Projekt Pure wurden auf SAP-Seite im CCC über die Bereitstellung zweier Webservices realisiert, über welche ausschließlich die von Pure benötigten Attribute abgerufen werden. Die Datenübermittlung erfolgt über eine gesicherte VPN-Verbindung. Eine zusätzliche Absicherung des SAP-Systems wurde durch den Einsatz des SAP Webdispatchers mit URL-Filterung und einer restriktiven Einschränkung der Berechtigungen des SAP-Nutzers, über den der Zugriff vorgenommen wird, erreicht. Eine schematische Darstellung des Zugriffs ist in Abbildung 3 dargestellt.

Der Webservice wurde so gestaltet, dass die spezifizierten Informationen entweder für eine einzelne Person oder aber für alle Personen übertragen werden können. Durch die Aufteilung der Datenlieferungen in diese drei Tabellen konnte die häufig vorkommende Zuordnung einer Person zu verschiedenen Organisationseinheiten transparent bereitgestellt werden.

Für die Bereitstellung von Projektinformationen wurde ein weiterer Webservice entwickelt, der zum Stichtag ausgewählte Stammdaten zu den Drittmittelprojekten aus dem Modul SAP-PSM selektiert und diese ebenfalls als Tabelle übergibt.

Abbildung 3: Infrastruktur der Pure-SAP-Schnittstelle



Der Webservice, der die Daten zum Stichtag aus dem Bereich Personalwirtschaft bereitstellt (SAP HR), liefert als Ergebnis drei Tabellen mit folgenden Inhalten an Pure zurück:

- eine Tabelle mit der im SAP-HCM-System hinterlegten hierarchischen Organisationsstruktur
- eine Tabelle mit den Zuordnungen von Personen zu Organisationseinheiten und
- eine Tabelle mit ausgewählten Personalinformationen.

Um es den SAP-Key Usern zu ermöglichen, die über die Webservices bereitgestellten Daten zu plausibilisieren und zu kontrollieren, wurde eine Transaktion für den SAP GUI entwickelt, mit der auf die selben Daten zugegriffen werden kann, wie sie auch über die Webservices bereitgestellt werden.

Zur Implementierung des Webserviceaufrufs in Pure wurden den Entwicklern von Atira entspre-

chende WSDL-Dateien bereitgestellt. Hier zeigte sich auch der Vorteil eines webserviceorientierten Schnittstellenansatzes, da bereits durch die Architektur eindeutige Verantwortungsbereiche abgegrenzt werden: Zum einen die Bereitstellung der Webservices durch das CCC und zum anderen die Implementierung des Webserviceaufrufs und der weiteren Verarbeitung in Pure durch den Datenbankanbieter Atira. In der Testphase wurden noch kleinere Anpassungen und Weiterentwicklungen an der Schnittstelle vorgenommen, nach positivem Testabschluss schloss sich die Produktivsetzung der Schnittstelle an. Die erfolgreiche Implementierung der Schnittstelle zeigt sich vor allem auch daran, dass es im produktiven Betrieb bisher keine Probleme oder Störungen gegeben hat.

3.3 Rechner und Routen: Einbettung im Medien- und Informationszentrum

Das Hosting für das Forschungsinformationssystem der Hochschule erfolgt im Leuphana Medien- und Informationszentrum (MIZ). Server-Maintenance und Updates der Anwendung erfolgen durch den Datenbankanbieter Atira, Hardware, Betriebssystem und Backups durch das MIZ. Teile der Daten werden aus dem Niedersächsischen Hochschulkompetenzzentrum für SAP (CCC) bezogen. Mitarbeiterinnen und Mitarbeiter der Leuphana können über eigene Zugänge mit der Datenbank arbeiten. Der Pure Server ist unter anderem durch eine Firewall umfassend geschützt.

Das Miteinander aller am Betrieb Beteiligten koordiniert ein Regelwerk, die “Pure Policy”, in der die Zuständigkeiten, die Zuverlässigkeit, die Sicherheit, die Performanz und die Wiederherstellbarkeit geregelt werden. Enthalten sind die Punkte:

- Redundanzsystem für Updates und Bugfixes sowie als Fallback für Notfälle
- nächtliches Backup mit ADSM Tivoli Storage Manager
- Monitoring mit Nagios (Server, Dienst und Datenbank)
- Datenschutzbestimmungen
- Wartungsfenster und Informationspolitik bei Störungen
- Physikalischer Schutz
- Firewall-Schutz und Administrationszugänge

3.3.1 Überführung in eine virtuelle Infrastruktur

Als Basissystem wurde ein Debian System mit Festplattenspiegelung installiert, der Datennetzanschluss erfolgt über die Cisco Catalyst- und ASA-Infrastruktur. Zur IT-Infrastruktur Strategie des MIZ gehört es, dass Pure in 2012 in eine vmWare-Virtualisierungsumgebung überführt wird. Bei Ausfall eines vmWare-Gastgebers werden im Hochverfügbarkeitsmodus virtualisierte Server automatisch migriert, Pure wird also ohne spürbare Unterbrechung für die Nutzer auf anderer Hardware weiter ausgeführt. Die Datenspeicherung erfolgt in einem redundanten SAN-Speichersystem, an das alle Gastgeber angeschlossen sind. Das MIZ fördert mit dieser IT-Infrastruktur-Strategie die nachhaltige Entwicklung der Universität, indem sie die optimale Nutzung der Ressourcen ermöglicht und gleichzeitig die Ausfallsicherheit des ganzen Systems signifikant erhöht.

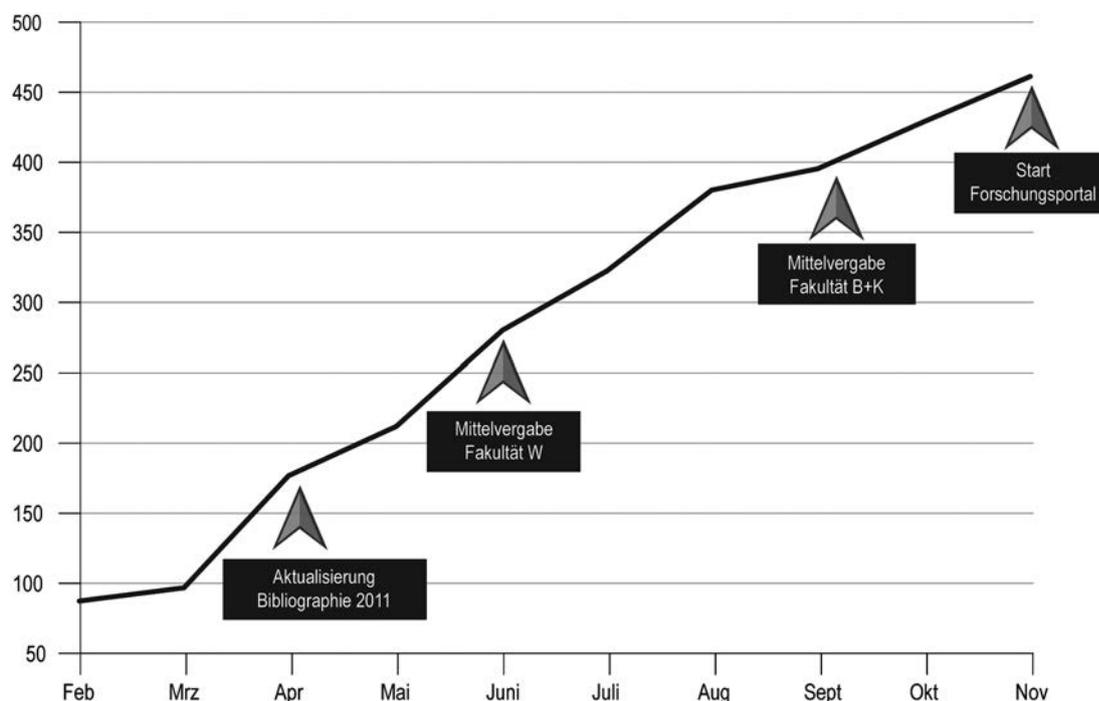
4 Going Live: Einführung in der Universität

Als Auftakt für die Einführung der neuen Forschungsdatenbank wurde Anfang 2011 die jährliche Aktualisierung der Hochschulbibliographie gewählt. Dieser etablierte und weithin akzeptierte Prozess sieht vor, dass die Mitglieder der Universität ihre Publikationen zur Erfassung in der Universitätsbibliothek einreichen. Dieses Verfahren wurde in Grundzügen beibehalten, so dass sich die Neuerung zunächst auf den Systemwechsel von PICA nach Pure beschränkte. Die Autorinnen und Autoren wurden informiert, dass sie bereits in PICA erfassten Publikationen in das neue System Pure überführt wurden und erhielten die Gelegenheit, sich zur Datenbank anzumelden, um die Einträge prüfen und falls nötig korrigieren zu können.

Im zweiten Quartal 2011 begann die Fakultät Wirtschaft mit der Vorbereitung der internen Mittelvergabe. Für den Indikator Publikationen sollten Einträge der Hochschulbibliographie zentral ausgewertet werden. Forschungsservice und Dekanat organisierten gemeinsam eine Informationskampagne zur Durchsicht der Datenbestände und die Meldung noch fehlender Publikationen. Diese Kampagne führte noch einmal zum Anstieg der Pure Anmeldungen, zur Vervollständigung des Redakteursnetzes und zu einer Verbesserung des Datenbestands in Pure aufgrund zahlreicher Korrekturen, die die Nutzerinnen und Nutzer vornahmen.

Die beiden Auftaktprojekte führten innerhalb von fünf Monaten zu einer Steigerung der Nutzerzahlen von 93 (Pilotphase) auf 370. Im Lauf des Jahres sorgten weitere Anwendungsprojekte für einen kontinuierlichen Anstieg der Anmeldungen (s. Abbildung 4).

Abbildung 4: Regelmäßige Kampagnen zur Nutzung von Pure lassen die Anmeldezahlen stetig steigen. Dargestellt ist die Entwicklung der Nutzerzahlen im Jahr 2011 mit den Meilensteinen Hochschulbibliographie, Mittelvergabe und Start Forschungsportal.



Mit der Einführung von Pure begann der systematische Aufbau eines Redakteursnetzes in der Universität. Die Redakteursrolle in Pure ermöglicht die Vergabe von Berechtigungen zur Datenpflege in Instituten, Arbeitsgruppen und Zentren. Diese Rolle ist in der Regel bei den Sekretariaten angesiedelt, wobei die Unterstützung durch einen Wissenschaftler oder eine Wissenschaftlerin als gute Praxis empfohlen wird. Das Redakteursnetz umfasst ein Jahr nach Einführung knapp 80 Personen und deckt 74 % der Professuren ab. Zu den Unterstützungsmaßnahmen für Redakteure zählen eine individuelle Einführung am Arbeitsplatz, wöchentliche Redakteurstreffen mit kurzen Tutorials (30 min) und anschließenden betreuten Arbeiten in der Gruppe.

4.1 Anwendungsprojekte – Nutzung des Datenbestands

In Vorbereitung der Datenbankeinführung war bereits eine Liste von knapp zwei Dutzend internen Prozessen erstellt worden, die regelmäßig Forschungsinformationen benötigen – von der Mittelvergabe über das Forschungsmarketing bis hin zu Tätigkeitsberichten und externen Evaluationen. Mithilfe der Forschungsdatenbank Pure soll eine autorisierte Sammlung entstehen, aus der

die benötigten Informationen entweder von den beteiligten Individuen selbst oder bei entsprechender Legitimation von den zuständigen Organisatoren abgerufen werden können.

Anwendungsprojekte im ersten Jahren waren neben den Mittelvergabeverfahren der Fakultäten z.B. die Bewerbungen für die Leuphana Forschungspreise. Für den vom Universitätsmarketing betreuten Internetauftritt wurde ein Online-Forschungskatalog entwickelt sowie ein Typo3-Plugin zur Integration von Publikations- und Projektlisten auf die persönlichen Webseiten der Wissenschaftlerinnen und Wissenschaftler.

Die Erstellung eines Querschnittsberichts für die Evaluation der Genderforschung in Niedersachsen sei hier exemplarisch vorgestellt, um den Mehrwert des Instruments Forschungsdatenbank zu illustrieren:

Die Wissenschaftliche Kommission Niedersachsen (WKN) begann in 2011 mit einer Evaluation der Genderforschung in Niedersachsen. In diesem Zuge waren die zur Genderforschung beitragenden Forschungseinheiten zu benennen und Fachgebiete jenseits der Geistes- und Gesellschaftswissenschaften einzubeziehen. Dies bedeutete faktisch eine Datenerhebung in der Universität über alle Organisationseinheiten hinweg.

Zunächst wurden Forschungseinheiten und Personen identifiziert, die in den letzten fünf Jahren Forschungsarbeiten zu Geschlechterfragen durchgeführt haben. Dieser Personenkreis wurde gebeten, relevante Publikationen, Projekte, Beratungs- und Beiratstätigkeiten sowie ihre Expertenprofile in der Forschungsdatenbank zu vervollständigen und sie dem Fachgebiet „Gender-Diversity“ zuzuordnen. Auf diese Weise wurden in einem sechswöchigen Zeitraum über 400 Publikationen sowie mehr als 60 Projekte und Expertenprofile erschlossen.

Für die Ausgabe der Informationen wurde mithilfe des Pure Berichtsmoduls eine spezielle Berichtsvorlage konfiguriert. Filterkriterium war wiederum das Fachgebiet „Gender-Diversity“. So konnten bereits während der Dateneingabe wöchentliche Zwischenbilanzen erstellt und Korrekturfassungen an die teilnehmenden Wissenschaftlerinnen und Wissenschaftler versandt werden. Durch die standardisierten Berichtsvorlagen sind diese Berichtsversionen gut kontrollierbar (Datum und Erzeuger werden jeweils vermerkt).

Vorteile des Instruments Forschungsdatenbank gegenüber der herkömmlichen Datenerhebung waren zum einen die strukturierte Erfassung über die Eingabeformulare der Datenbank, zum anderen entfiel der sonst übliche Aufwand zur Vereinheitlichung von Publikationslisten usw. Die Erzeugung einer einheitlich formatierten „Gender-Diversity-Bibliographie“ aus der zugehörigen Berichtsvorlage im Word-Format ist beispielsweise innerhalb einer Minute möglich. Darüber hinaus stehen die so erschlossenen Einträge unmittelbar für alternative Anwendungen zur Verfügung, z.B. für die Fachgebietssuche im Online-Forschungskatalog¹. Die Redaktion des „Gender-Diversity Portals“ plant derzeit, auf dieser Datenbasis die Leistungen seines Forschungsnetzwerks besser zu präsentieren.

5 Lessons learned und Ausblick

5.1 Fazit nach dem ersten Jahr der Einführung

In der Vorbereitungsphase war insbesondere die Altdatenübernahme zeitintensiv. Beispielsweise war der Aufwand für die Übernahme der PICA-Daten trotz sorgfältiger Vorbereitung kaum planbar, und die Korrekturen am Datenbestand führten zu erheblichen Zeitverzögerungen. Im Nachgang bildet die komplette Durchsicht und Korrektur der Hochschulbibliographie eigentlich ein eigenes Teilprojekt im Gesamtkonzept. Weitere Nebenprodukte der Datenbankeinführung

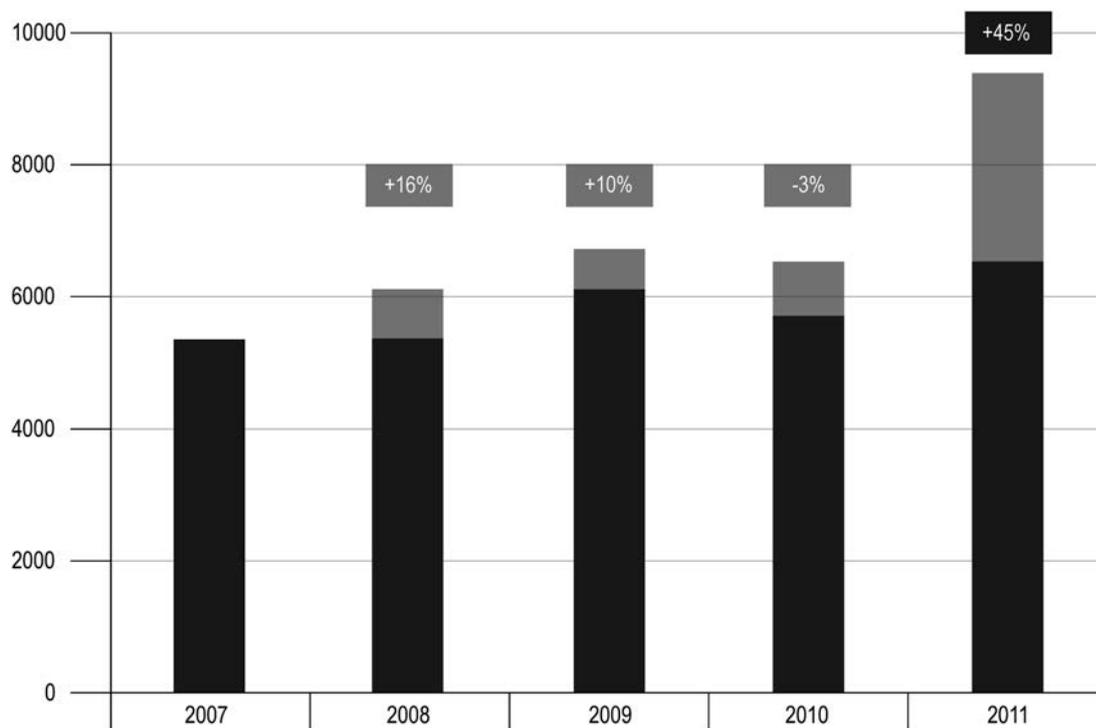
¹ <http://www.leuphana.de/forschungsdatenbank> (nur Einträge, die in Pure als „öffentlich“ gekennzeichnet sind)

waren die Einrichtung der ersten Webservice-Schnittstelle im niedersächsischen SAP-System des CCC, der Aufbau eines internen Verzeichnisses der Mittelgeber und Kooperationspartner, die Weiterentwicklung des Fachgebietskatalogs aus der Hochschulbibliographie und die Verschlagwortung des Projektbestands.

Wie steht es um die Zielerreichung? Mit der Einführung eines integrierten Forschungsinformationssystems wurden vorrangig zwei Ziele verfolgt: Die Schaffung einer einheitlichen, zwischen Wissenschaft und Verwaltung abgestimmten Datenbasis für die Forschungsberichterstattung und die Erzeugung von Mehrwerten für die Wissenschaftlerinnen und Wissenschaftler durch verbesserte Verfügbarkeit der Informationen für eigene Prozesse.

Nach dem ersten Jahr des Produktivbetriebs kann konstatiert werden, dass die Schaffung einer soliden Datenbasis und die Motivation der Datenbanknutzung sich gegenseitig bedingen: Durch die systematische Korrektur und Anreicherung der übernommenen Publikationen im Pure Redaktionsteam konnte die Datenbank relativ schnell ersten Nutzen für Mittelvergabeverfahren und andere Prozesse generieren. Den Wissenschaftlerinnen und Wissenschaftlern wurden die bereits vorhandenen Daten zur Verfügung gestellt, für umfangreiche Nacherfassungen steht ein Eingabeservice zur Verfügung.

Abbildung 5: Meldung von Publikationen für die Leuphana Hochschulbibliographie. Der konsequente Einsatz von Pure für die Forschungsberichterstattung hat die Meldebereitschaft 2011 stark erhöht.



Wie Abbildung 5 zeigt, ist im Jahr der Pure Einführung beispielsweise die Meldebereitschaft für die Hochschulbibliographie stark gestiegen. Wurden in den Vorjahren 800-1.000 Publikationen an die Bibliothek gemeldet, so belief sich der Zuwachs 2011 auf 3.000 neue Einträge. Die Forschungsdatenbank hat sich hier schnell als zentrales Instrument etabliert, nicht zuletzt wegen der besseren Zugänglichkeit für die Beteiligten und einer Selbststeuerung der Einträge.

Die gute Abdeckung im Bereich Publikationen und die deutlich einfacheren Auswertungsmöglichkeiten überzeugten Dekane, ihre zuvor praktizierten Datenerhebungsverfahren für die Mittelvergabe zugunsten einer Nutzung der Forschungsdatenbank aufzugeben. Dies ist nicht für jeden

Prozess selbstverständlich. Für einzelne Kolleginnen und Kollegen erscheint es zunächst einfacher, bei den bewährten Word- und Excel-Listen zu bleiben und diese Jahr für Jahr in den gewohnten Umlaufverfahren zu aktualisieren. Demgegenüber erfordert die Nutzung des integrierten Forschungsinformationssystems für komplexere Berichte eine gewisse Systemkompetenz und eine gute Kenntnis der Daten. Insbesondere im ersten Jahr entsteht teils erheblicher zusätzlicher Aufwand durch Prüfungen und Korrekturen der Einträge sowie durch die Schulungen bzw. Beratungen der Nutzerinnen und Nutzer.

In der Erstellung von Berichten sind jedoch bereits deutliche Effizienzgewinne realisierbar. Da jede Korrektur die Datensätze dauerhaft verbessert und die Nutzerkompetenz laufend steigt, werden die Vereinfachungen beim zweiten Mal noch deutlicher ausfallen.

Das Pure Redaktionsteam arbeitet daran, für viele der vorhandenen Berichtsansätze Effizienzgewinne zu realisieren. Dies erfordert eine intensive inhaltliche und prozessuale Begleitung. Daher investiert die Leuphana auch im zweiten Jahr nach der Einführung in zusätzliche Kapazitäten für die Forschungsberichterstattung.

Insgesamt beobachten wir, dass die Idee des integrierten Forschungsinformationssystems bereitwillig aufgegriffen wird. Allerdings begegnen wir auch überhöhten Erwartungen an die à la carte Verfügbarkeit der Einträge aus dem Forschungsinformationssystem. Datenqualität wie Nutzerkompetenz werden ebenfalls gern überschätzt. Die Botschaft, dass die an Forschungsberichten interessierten Einrichtungen sich aktiv an der Motivation der Wissenschaftlerinnen und Wissenschaftler und der Meldung von Korrekturbedarfen beteiligen müssen, ist ein wichtiger Baustein in der Kommunikation.

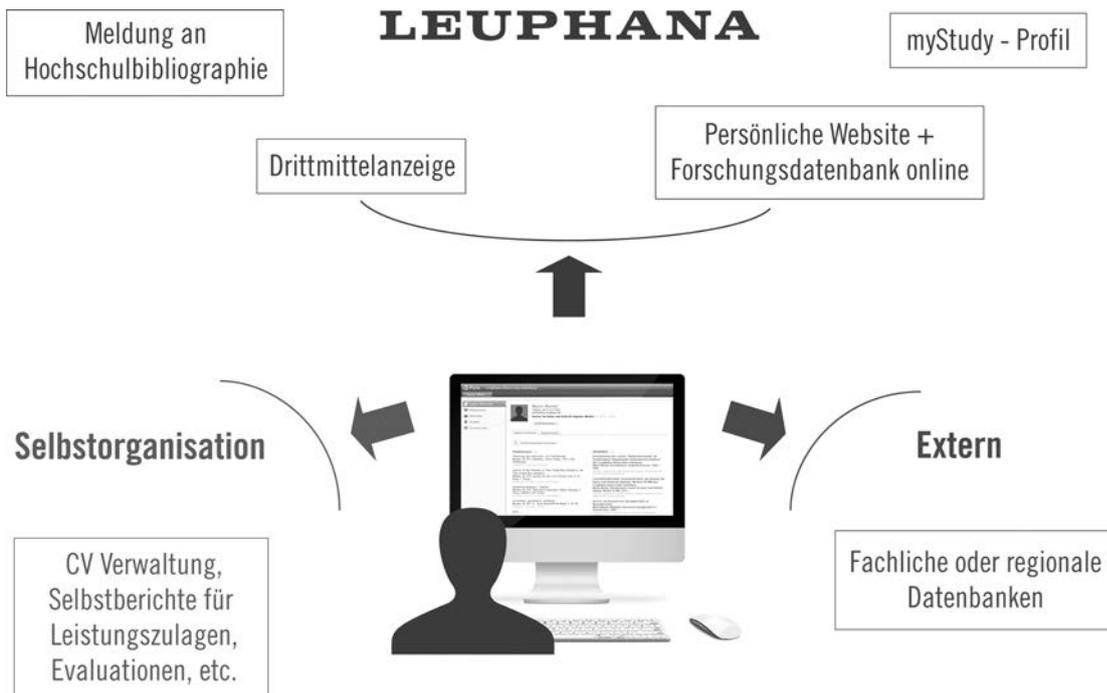
5.2 Aus unseren Erfahrungen: Anregungen für die Entwicklung von universitären Forschungsinformationssystemen in Deutschland

Eines der wichtigen Ziele für die Leuphana war es, die über viele Stellen und Systeme verteilten Informationen zu sichten und aus ihnen eine Datenbasis für die Forschungsberichterstattung zu schaffen. Dazu gehören nicht zuletzt Informationen, die exklusiv oder mit höherer Qualität in externen Datenbanken gepflegt werden. Wichtig war hier vor allem der Abgleich mit den Daten der Förderer – z.B. mit den Datenbanken GEPRIIS (DFG) und CORDIS (europäische Forschungsrahmenprogramme). Die Übernahme und der Abgleich erfolgten manuell – hier wäre es insbesondere für größere Universitäten von Vorteil, wenn die Daten in elektronischer Form ausgetauscht werden könnten.

Inspiration für Services, die beim Aufbau lokaler Systeme hilfreich wären, bietet ein Blick in andere Länder. Beispielsweise nutzen die lokalen Systeme in Norwegen ein nationales Verzeichnis von Universitäten und Forschungseinrichtungen als Datenquelle. So können die institutionellen Kooperationsbeziehungen auf einer einheitlichen Datenbasis analysiert werden. In Dänemark können lokale Forschungsinformationssysteme ein Verzeichnis mit mehreren Tausend Zeitschriftentiteln nutzen, und Vergleichsdaten aus der Forschungsevaluation stehen für Benchmarking zur Verfügung.

Als Universität verfolgen wir das Ziel, Informationen über die Organisation zunächst vor Ort zu erfassen und die Forschungsdatenbank als lokalen „One stop shop“ für die Wissenschaftlerinnen und Wissenschaftler zu etablieren (Abbildung 6). Erstes Projekt für einen Export von Forschungsdaten ist die Repräsentation von Profilen der Forschungseinheiten im Portal „Forschung in Niedersachsen“.

Abbildung 6: Vision - Forschungsdatenbank als lokaler „One stop shop“ für die Wissenschaftlerinnen und Wissenschaftler



Für die zukünftige Entwicklung wünschen wir uns vor allem eine bessere Wiederverwendbarkeit vorhandener Daten und eine bessere Interoperabilität für institutionelle Systeme und Fachdatenbanken – vor allem um Barrieren für Wissenschaftlerinnen und Wissenschaftler zu senken. Sie kritisieren zu Recht die Vielfalt der Abfragen, die seitens der Arbeitgeber, der Fachgemeinschaften und der Förderer an sie gestellt werden.

Der europäische Datenstandard CERIF bietet für die (Weiter-)entwicklung von Forschungsinformationssystemen wertvolle Anregungen. CERIF wird von engagierten Wissenschaftlerinnen und Wissenschaftlern sowie kommerziellen Datenbankanbieter im Verband euroCRIS entwickelt.² Wertvoll ist auch der Austausch mit den vielfältigen Mitgliedseinrichtungen, die von renommierten Forschungseinrichtungen über Bibliotheken bis hin zu Forschungsförderern reichen.

² www.eurocris.org