



Wissenschaftskommunikation der Zukunft

4. Konferenz der Zentralbibliothek
Forschungszentrum Jülich

Rafael Ball (Hrsg.)

6. – 8. November 2007

Beiträge und Poster

Mit einem Festvortrag von Ernst Pöppel

WissKom 2007

Forschungszentrum Jülich GmbH
Zentralbibliothek

Wissenschaftskommunikation der Zukunft

4. Konferenz der Zentralbibliothek Forschungszentrum Jülich

Rafael Ball (Hrsg.)

6. – 8. November 2007

Beiträge und Poster

Schriften des Forschungszentrums Jülich
Reihe Bibliothek/Library

Band/Volume 18

ISSN 1433-5557

ISBN 978-3-89336-459-6

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte Bibliografische Daten sind im Internet
über <<http://dnb.d-nb.de>> abrufbar.

Herausgeber Forschungszentrum Jülich GmbH
und Vertrieb: Zentralbibliothek, Verlag
 52425 Jülich
 Telefon (02461) 61-5368 · Telefax (02461) 61-6103
 E-Mail: zb-publikation@fz-juelich.de
 Internet: <http://www.fz-juelich.de/zb>

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2007

Der Herausgeber übernimmt für die Richtigkeit der Beiträge und deren Inhalt keine
Verantwortung oder Haftung.

Zusammenstellung: Beatrix Küven,
 Roswitha Moes
 Zentralbibliothek, Forschungszentrum Jülich GmbH

Schriften des Forschungszentrums Jülich
Reihe Bibliothek / Library Band / Volume 18

ISSN 1433-5557
ISBN 978-3-89336-459-6

Vollständig frei verfügbar im Internet auf dem Jülicher Open Access Server (JUWEL)
unter <http://www.fz-juelich.de/zb/juwel>

Alle Rechte vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie oder
in einem anderen Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder
unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

INHALTSVERZEICHNIS

Konferenzkomitees.....	3
Vorwort.....	5
FESTVORTRAG	7
Ernst Pöppel: Wissen - und wie es kommuniziert werden kann.....	9
E-SCIENCE: NEUE FORMEN WISSENSCHAFTLICHER ZUSAMMENARBEIT	23
Anne-Katharina Weilenmann: Von Cyberscience zu e-Science.....	25
Katrin Weller, Indra Mainz, Ingo Paulsen, Dominic Mainz: Semantisches und vernetztes Wissensmanagement für Forschung und Wissenschaft.....	33
Ute Rusnak, Matthias Razum, Leni Helmes: Wissensvernetzung im Forschungsprozess	47
WISSENSCHAFTSKOMMUNIKATION IN LEHRE UND FORSCHUNG	59
Elena Semenova, Martin Stricker: Eine Ontologie der Wissenschaftsdisziplinen. Entwicklung eines Instrumentariums für die Wissenskommunikation.....	61
Peter Haber, Jan Hodel: Historische Fachkommunikation im Wandel. Analysen und Trends	71
Lydia Bauer, Nadja Böller, Josef Herget, Sonja Hierl: Konzepte zur Förderung der Wissenschaftskommunikation: Der Churer Ansatz zur Vermittlung von kollaborativen Kompetenzen	81
WEB 2.0: BEISPIELE AUS DER PRAXIS	93
Susanne von Itter: Wissenschaftskommunikation in der Entwicklungsforschung/Entwicklungszusammenarbeit Web2.0 und Communities of Practice - ein Beitrag aus der Praxis	95
Steffen Leich-Nienhaus: Wissenschaftliche Informationsversorgung am modernen digitalen Arbeitsplatz	107
Christian Hänger, Christine Krätzsich: Collaborative Tagging als neuer Service von Hochschulbibliotheken.....	123
Christoph Bläsi: Mobile IT in kleinen und mittleren Medienunternehmen zur Prozess- und Serviceverbesserung	135

HERAUSFORDERUNG PRIMÄRDATENMANAGEMENT145

Michael Diepenbroek, Hannes Grobe: PANGAEA® als vernetztes
Verlags- und Bibliothekssystem für wissenschaftliche Daten..... 147

Jan Brase, Jens Klump: Zitierfähige Datensätze: Primärdaten-
Management durch DOIs 159

Harald Krottmaier: Die Systemarchitektur von PROBADO: Der
allgemeine Zugriff auf Repositorien mit nicht-textuellen Inhalten 169

Annette Holtkamp: Open Access Publishing in der Hochenergiephysik:
Das SCOAP³ Projekt 177

Wiebke Oeltjen: Metadaten-Management mit MyCoRe 185

Karin Weishaupt: Open-Access-Zeitschriften als neue Form
wissenschaftlicher Kommunikation: Vorbehalte und Vorschläge für
Maßnahmen zur Akzeptanzsteigerung..... 193

WISSENSCHAFTSINDIKATOREN: NEUE ASPEKTE UND ENTWICKLUNGEN .207

Wolfgang Glänzel, Koenraad Debackere: Bibliometrie zwischen
Forschung und Dienstleistung 209

Patrick Vanouplines, Ronald Beullens: Merging information sources to
obtain the impact factor of open access journals 223

Dirk Tunger: Bibliometrie als Teil eines Trenderkennungs-Systems in der
Naturwissenschaft 235

Miloš Jovanović: Indicators for Footprints through science – Designing
the Journal Application Level (JAL)..... 247

WISSENSCHAFTSINDIKATOREN ALS MANAGEMENT TOOL251

James Pringle: The ISI Web of Knowledge as a Management Tool253

Show-Ling Lee-Müller, Gerd Schumacher: Einsatz bibliometrischer
Analysen im EU-Projekt zur Technologiefrüherkennung SMART263

Henning Möller: Messen, Steuern, Regeln – zum Controlling der
Helmholtz-Forschung273

LISTE DER AUTOREN283

REGISTER293

SPONSOREN301

Konferenzkomitees

Programmkomitee

Dr. Rafael Ball	Forschungszentrum Jülich, Zentralbibliothek
Dr. Sigrun Eckelmann	Deutsche Forschungsgemeinschaft, Abt. III Programm- und Infrastrukturförderung
Dr. Felix Grützner	Chefredakteur der Fachzeitschrift „Wissenschaftsmanagement“
Prof. Dr. Josef Herget	Hochschule für Technik und Wirtschaft, Chur, Leiter des Studiengangs Informationswissenschaft
Prof. Dr. Dr. Thomas Lippert	Forschungszentrum Jülich, Zentralinstitut für Angewandte Mathematik
Prof. Dr. Anthony van Raan	Center for Science and Technology Studies, University of Leiden

Organisationskomitee

Dr. Rafael Ball	Forschungszentrum Jülich, Zentralbibliothek (Konferenzleitung)
Dr. Christoph Holzke	Forschungszentrum Jülich, Zentralbibliothek
Dr. Bernhard Mittermaier	Forschungszentrum Jülich, Zentralbibliothek
Anne Otto	Forschungszentrum Jülich, Zentralbibliothek
Edith Salz	Forschungszentrum Jülich, Zentralbibliothek

Wissenschaftskommunikation der Zukunft

Die Wissenschaftskommunikation der Zukunft hat bereits begonnen. Mit der Digitalisierung der Informationsressourcen einerseits und mit dem Einzug von elektronischen Kommunikationstools in den wissenschaftlichen Prozessen andererseits hat die wissenschaftliche Informationsversorgung in den letzten Jahren einen qualitativen Sprung gemacht. Dies betrifft nicht nur die Medienformen der Informationen und Quellen für wissenschaftliches Arbeiten, sondern auch den gesamten wissenschaftlichen Erkenntnisprozess vom Beginn bis zu seinem Ende.

Dabei wird die Wissenschaftskommunikation der Zukunft geprägt von einem massiven Wandel der Kommunikationsinfrastruktur und von der Art und Weise, wie wissenschaftliche Information kommuniziert und verbreitet wird. Die Durchdringung und Bearbeitung des Themas Wissenschaftskommunikation ist längst zu einem integralen Prozess geworden, bei dem neben Bibliotheken, Informations- und Kommunikationswissenschaften sowie ICT auch alle anderen Wissenschafts-Disziplinen beteiligt sind. Im Zeitalter von E-Science, der „Wissenschaft aus der Steckdose“, ist Wissenschaftskommunikation zu einem ganzheitlichen Ansatz aus Informationsversorgung, wissenschaftlichen Inhalten, Aufbereitung und Publikation von wissenschaftlichen Erkenntnissen sowie der Implementierung von Strukturen für die Langzeit-Verfügbarkeit geworden.

Die Konferenz „Wissenschaftskommunikation der Zukunft“ ist bereits die vierte Konferenz der Zentralbibliothek zu aktuellen Themen im Umfeld von Wissenschaft, Publikation und Bibliothek und hat sich in der Community längst zu einem festen Punkt in der Tagungsstruktur etabliert. Die Beiträge im Konferenzband gruppieren sich um die Themen-Schwerpunkte der Konferenz: „E-Science“, „Wissenschaftsindikatoren“, „Web 2.0“ und „Primärdaten-Management (Open-Access)“. Mit dieser Themenauswahl wird der integrative Ansatz der Wissenschaftskommunikation deutlich. Die anstehenden Fragen und Diskussionsfelder werden von allen Seiten beleuchtet und interdisziplinär durchdrungen. Der Konferenzband enthält neben den Ausführungen der Referenten zudem die Beiträge der Postersession sowie den Festvortrag von Professor Ernst Pöppel zum Thema „Wissen und wie es kommuniziert werden kann“.

An dieser Stelle gilt mein herzlicher Dank allen, die zum Gelingen der Konferenz beigetragen haben, sei es als Vortragende, Moderatoren, Organisatoren oder Teilnehmer. Besonderer Dank gilt dem Forschungszentrum Jülich und den Sponsoren, die diese Konferenz erst möglich gemacht haben.

Dr. Rafael Ball
Forschungszentrum Jülich, Zentralbibliothek

Festvortrag

Wissen - und wie es kommuniziert werden kann

Ernst Pöppel

Ist es nicht eigentlich verwunderlich zu glauben, dass mit gezielter Planung, auf der Grundlage von Wissen, das man hat, und wenn man es nicht hat, das man sich besorgen kann, (was man dann Wissenschaft nennt), alle Probleme der Welt gemeistert werden können? Es ist wohl die Erwartung vieler, die auf die Wissenschaften schauen, (aber die nicht selber in der Forschung tätig sind), dass Wissensvermehrung in erster Linie der Problembewältigung gilt. Schaut man aber um sich (und auch in sich), dann muss man feststellen, dass wir mit dem angereicherten Wissen, das immer mehr wird, (und das niemand mehr überblicken kann), bei der Lösung von Problemen nicht weit gekommen sind; im Gegenteil: man gewinnt den Eindruck, dass durch mehr Wissen mehr Probleme geschaffen werden. Ein Grund (nur ein Grund) dafür, mit den Problemen der Welt und mit unseren eigenen Problemen nicht zurecht zu kommen, ist ein falsches oder besser: ein zu eingeschränktes Wissen über das Wissen selbst, sich nämlich auf nur einen Teil dessen, was Wissen umfasst, zu beziehen. Wir wissen mehr über das Wissen, wenn wir in die Vergangenheit schauen. Manchmal ist es gut, in der Bibel zu lesen; sie beginnt nach der Übersetzung von Martin Luther mit den Worten:

„1 Am Anfang schuf Gott Himmel und Erde. 2 Und die Erde war wüst und leer, und es war finster auf der Tiefe; und der Geist Gottes schwebte auf dem Wasser. 3 Und Gott sprach: Es werde Licht! Und es ward Licht. 4 Und Gott sah, dass das Licht gut war. Da schied Gott das Licht von der Finsternis 5 und nannte das Licht Tag und die Finsternis Nacht. Da ward aus Abend und Morgen der erste Tag.“

Dieser Text der Genesis bezieht sich auf drei Formen des Wissens, nämlich begriffliches oder explizites Wissen (Nennen, Sagen), implizites oder Handlungs-Wissen (Schaffen, Tun) und bildliches oder Anschauungs-Wissen (Sehen, Erkennen). Am Anfang der geistigen (und für viele auch geistlichen) Geschichte unseres Kulturkreises steht ein Text, in dem ein Rahmen für jene Formen des Wissens bestimmt wird, der auch dem Menschen, nicht nur dem Schöpfer, auf den die Schrift sich bezieht, gemäß ist, der aber in den Wissenschaften wenn nicht vergessen so doch oft vernachlässigt wurde. Die dreifache Begründung des Wissens liegt in unserer Natur; sie ist durch die Weise unserer Welterfahrung, von den Verarbeitungsprinzipien unserer Sinnessysteme und unseres Gehirns, vorgegeben. Obwohl diese dreifache Begründung ein Wesensmerkmal ist, müssen wir uns dieser Tatsache trotzdem immer wieder versichern. Was selbstverständlich ist, wird in seiner Bedeutung oft verkannt, an was wir gewöhnt sind, wird übersehen; erst, wenn das Selbstverständliche verloren gegangen ist, wird es rückblickend als wesentlich erkannt. Die drei Formen menschlichen Wissens sind so grundlegend, sie

bestimmen derart stabile Koordinaten, dass eine Gesellschaft des Wissens und Lernens (wenn es die je geben sollte) nur dann wohl verortet, eine Wissenswelt ist nur dann fest gefügt, wenn die Bewohner dieser Welt ihr Wissenspotential gemäß ihrer Ausstattung, das von der Natur mitgegeben wurde, dreifach gestalten, also als explizites, als implizites und als bildliches Wissen.

Wie kann man diese drei Formen des Wissens genauer kennzeichnen? Es sollen keine präzisen Definitionen gegeben werden, sondern mit den Umschreibungen der einzelnen Wissensformen wird auf das jeweils Gemeinte hingewiesen. (Eine Definition bestimmt immer einen sehr engen Rahmen, damit sich alle in einer expliziten Weise auf Sachverhalte innerhalb des Rahmens beziehen können; was durch die Definition nicht abgedeckt wird, bleibt außerhalb des Rahmens; eine Konsequenz solcher notwendigerweise ausgrenzender Definitionen ist, daß die Kreativität eingeschränkt wird, denn kreative Prozesse gehen häufig über einen vorgegebenen Rahmen hinaus).

Explizites Wissen bedeutet, Auskunft erteilen zu können, also Bescheid zu wissen. Explizites Wissen ist Information mit Bedeutung. Explizites Wissen ist einem bewusst, und wenn man es vergessen hat, kann man es sich wiederholen. Explizites Wissen ist katalogisiert und katalogisierbar; es steht in Enzyklopädien und Lehrbüchern; man eignet es sich als jene Kenntnisse an, die man dann hat. Es ist jenes Wissen, das uns in unserer Geschichte der Neuzeit dominiert hat, und das manche als das eigentliche Wissen ansehen. Explizites Wissen, das uns begrifflich zur Verfügung steht, wird durch Lernen erworben, das manchmal mühsam ist; (aber Lernen kann auch ein Vergnügen sein). Es ist jenes Wissen, auf das sich bei Goethe der Famulus Wagner im Dialog mit Faust bezieht, wenn er sagt:

„Wie schwer sind nicht die Mittel zu erwerben,
Durch die man zu den Quellen steigt!
Und eh man nur den halben Weg erreicht,
Muss wohl ein armer Teufel sterben.“

Worauf ihm Faust antwortet:

„Das Pergament, ist das der heil'ge Bronnen,
Woraus ein Trunk den Durst auf ewig stillt?“,

und mit seiner Bemerkung das alleinige Quellenwissen und die alleinige Sehnsucht nach solchem sekundären Wissen in Frage stellt.

Explizites Wissen ist besonders in der neuzeitlichen Tradition des Rationalismus herausgehoben worden, wie es beispielsweise René Descartes im „Discours de la méthode“ („Von der Methode“) getan hat. Mit explizitem Wissen als Orientierung wird der Anspruch erhoben, jedes Problem klar und deutlich formulieren und damit auch lösen zu können. Wenn Sokrates sagt: „Ich weiß, dass ich nichts weiß“, dann bezieht

er sich auf explizites Wissen. Die veröffentlichten Erkenntnisse der Wissenschaften, insbesondere der Naturwissenschaften, repräsentieren explizites Wissen. Der Versuch, menschliches Wissen nur als explizites Wissen zu begreifen, kann jedoch in die Irre führen. In der Forschung zur „Künstlichen Intelligenz“ ging man anfänglich davon aus, daß explizites Wissen für das gesamte menschliche Wissen stehe. Da dieses Wissen sich präzise in der Sprache abbilde, könne es auch formal erfasst und mathematisch beschrieben werden, und man könne Algorithmen entwickeln, um den menschlichen Geist als Programmablauf in einem Computer festzuhalten. Dieses Projekt, die explizite Beschreibung und eine darauf aufbauende Algorithmisierung des menschlichen Geistes, kann nie gelingen. So meine ich, doch sei gesagt, dass manche Wissenschaftler, wie Informatiker, Computerwissenschaftler, Neurobiologen oder Psychologen, glauben, und viele offenbar auch hoffen, der Versuch der vollständigen Algorithmisierung des menschlichen Geistes könne gelingen. Dahinter mag der Traum von der Unsterblichkeit stehen: Wenn es gelingt, die Seele von der fleischlichen Körperlichkeit abzuziehen und in immer währendes Silizium zu übertragen, dann wäre ein solches Maschinen-Wesen wohl unsterblich. Diese Vision kann nie verwirklicht werden, da die Komplexität unseres Gehirns zu groß ist, und damit einzelne neuronale Zustände, die für bestimmte subjektive Zustände stehen, nicht berechenbar sind.

Das explizite Wissen kann man mit einem Bild der Hirnforscher auch als „linkshemisphärisch“ bezeichnen. Dieses Bild rührt daher, dass nach neuronalen Störungen in der linken Gehirnhälfte, beispielsweise nach einem Schlaganfall, die Fähigkeit zu sprechen verloren gehen kann. Es sieht dann so aus, als habe der Patient sein explizites Wissen verloren. Das mag aber nur so aussehen, da ein solcher Sprachverlust auch bedeuten kann, dass die Ankopplung des expliziten Wissens, das im Gehirn noch vorhanden ist, an die begriffliche Repräsentation in der Sprache unterbrochen ist, so als würden die Gedanken die Wörter oder die Worte nicht mehr erreichen.

Die zweite Form des Wissens ist implizit, und dieses Wissen bezieht sich auf unser Können und auf unsere Handlungen, ohne dass wir Worte hierfür haben oder haben müssen. Wenn das explizite Wissen mit dem sokratischen Satz „Ich weiß, dass ich nichts weiß“ gekennzeichnet werden kann, so gilt für das implizite Wissen der Satz: „Ich weiß nicht, dass ich weiß.“ Der Unterschied zwischen explizitem und implizitem Wissen lässt sich auch an der klassischen Frage des Augustinus verdeutlichen, wenn er sagt: „Was also ist Zeit? Wenn mich niemand danach fragt, weiß ich es; will ich einem Fragenden es erklären, weiß ich es nicht.“ Hier wird das Wort „wissen“ in zwei Bedeutungen verwendet, nämlich zuerst als implizites Wissen, und dann als explizites Wissen. Ungefragt und ungesagt weiß man Bescheid; mit klärenden Worten verwirrt man sich.

Doch implizites Wissen ist auch körperliches Wissen, nämlich jenes Wissen über bestimmte Bewegungsabläufe, wie ein Fahrrad zu fahren, ein Musikinstrument zu spielen oder mit dem Federhalter zu schreiben, was wir als Kind gelernt haben, und was dann selbstverständlich geworden ist. Nie können wir im Detail beschreiben, wie wir etwas machen, welches die Komponenten waren, die eine Bewegung als gelungen oder eine Handlung als erfolgreich erscheinen lassen. Wenn wir einen Golfschwung oder einen Tennisaufschlag beherrschen, dann geschieht die Bewegung mit uns, sie ist ein Teil von uns, die unreflektiert aus uns heraus entsteht. Wenn wir etwas wirklich können, dann beherrschen wir anstrengungslos den Ablauf einer Bewegung, ohne dass wir uns darauf konzentrieren müssen.

Neben dem impliziten motorischen Wissen wird unser Alltag bestimmt vom impliziten heuristischen Wissen. Heuristisches Wissen ist das Gewohnheitswissen des Tages, es ist das Eingebettetsein in Abläufe und auch Rituale, die nicht mehr hinterfragt werden. Wenn wir etwas tun oder erschaffen, mit den Händen etwas formen und manchmal kreativ gestalten, wenn wir kochen oder uns anziehen, wenn wir etwas „erledigen“, dann verwirklicht sich in diesen Tätigkeiten die heuristische Form des impliziten Wissens, die im Augenblick der Tat stumm ist, und deren Ergebnis erst im Rückblick bewusst wird oder dies werden kann. Diese Form des impliziten Wissens ist vor allem auch das experimentelle Wissen des Forschers. Wenn man ein Experiment macht, dann hat man üblicherweise soviel Erfahrungswissen in sich gesammelt, das man explizit gar nicht vermitteln kann. Die einzelnen Arbeitsschritte sind einem nicht mehr bewusst, doch ohne deren Kenntnis kann man kein erfolgreiches Experiment machen. Doch wie kann man solches „wissenschaftliches Wissen“ kommunizieren? Da es trotz aller Versuche nicht aufschreibbar ist, auch wenn man immer wieder in SOPs (standard operating procedures) diese stumme Wissen festzuhalten versucht, so bleibt nur die Möglichkeit des impliziten Lernens, dass also der Schüler vom Meister instruiert wird, und dass durch die Praxis des Lernens Kompetenz erworben wird.

Eine dritte Form des impliziten Wissen drückt sich in unseren Intuitionen aus, ohne die ein Künstler oder ein Wissenschaftler (und natürlich auch ein Handwerker, ein Politiker, ein Unternehmer oder eine Hausfrau) nicht wirken und nichts erreichen kann. Die Fülle und der Reichtum des intuitiven Wissens jedes einzelnen ist explizit nicht berechenbar ist, weil zu viele Faktoren zu berücksichtigen wären, die zum großen Teil nicht bekannt sind, und die auch nicht bekannt sein können. Diese Tatsache der Nichtberechenbarkeit unserer Innenzustände kann sich in Unberechenbarkeit äußern; niemand kann das Handeln eines anderen oder sein eignes Handeln voraussagen. Diese Offenheit des intuitiven Erfahrungswissens garantiert absolute Individualität; wir können nie kopiert werden, denn selbst wenn wir Klone wären, würden uns einzelne Prozesse in unserer impliziten Wissensbasis hinreichend verschieden machen.

Das intuitive Wissen spiegelt sich in unseren Entscheidungen, die oft „aus dem Bauch heraus“ erfolgen. Entscheidungen sind immer auch emotional gefärbt oder begründet, auch wenn diese emotionale Tönung in der Intuition nicht bewusst ist. Das intuitive Wissen ist jedoch nicht irrational, denn retrospektiv können wir uns in der Reflexion der Sinnhaftigkeit des Handelns versichern. Wenn Goethe darauf hinweist, daß sein künstlerisches Schaffen „mit einer gewissen Bewusstlosigkeit und gleichsam instinktmäßig“ ablaufe, bezieht er sich auf das implizite Wissen als intuitives Wissen. Intuition kennzeichnet den Experten, der ohne notwendige Reflexion handelt, und dennoch richtig handelt.

Die dritte Form des Wissens ist bildliches Wissen, und dieses erscheint uns ebenfalls in dreifacher Form, nämlich als Anschauungswissen, als Erinnerungswissen (oder auch episodisches Wissen) und als abstrahierendes Wissen. Das sinnliche Anschauungswissen ist so selbstverständlich, dass wir es erst erkennen, wenn es verloren gegangen ist. Wir müssen nur die Augen öffnen, um vom Anschauungswissen Kenntnis zu nehmen. Die Welt stellt sich uns bildlich vor in Formen und Gegenständen, in ruhenden und bewegten Gestalten. Der Aufbau der visuellen Welt erfolgt mühelos, indem unser Auge Umrisse, Figuren vom Hintergrund abhebt und als gesehenen Gegenstand in das Bewusstsein setzt. Es ist immer etwas Bestimmtes, was wir sehen, und in diesem Wahrnehmungsakt wird das Gesehene für wahr genommen. Beim Aufbau des visuellen Wissens unterliegen wir einem kategorialen Zwang; das Gehirn mit seinen Sinnessystemen kann gar nicht anders, als gestaltend zu wirken; es wird immer etwas Bestimmtes erkannt. Dass es sich hier um eine aktive Leistung des Gehirns handelt, die konstitutiv für unser Anschauungswissen ist, erkennt man an Patienten mit Agnosien, bei denen das Fürwahrnehmen des Gesehenen nicht mehr möglich ist; diese Patienten erkennen zwar noch etwas, doch wissen sie nicht mehr, was es ist. Die Gliederung des Sehraumes und die Gestaltung der visuellen Welt, das Wahrnehmen von Gegenständen und damit ihr Fürwahrnehmen, ist ein Ausdruck unseres bildlichen Wissens, das unser gegenwärtiges Erleben erfüllt.

Die zweite Form des bildlichen Wissens spiegelt sich in den sinnlichen Erfahrungen und den Episoden wider, also in den Erinnerungen, die wir in uns tragen. Erinnerungswissen aus der Vergangenheit ist mit Orten verbunden, die sich bleibend in unser Gedächtnis eingepägt haben. Die Bilder dieser Orte beziehen sich auf hervorstechende Episoden unserer Lebensgeschichte, mögen sie beglückend oder verletzend gewesen sein. Diese Bilder bestimmen unser Selbst, und sie verbinden uns mit der Welt. Wenn wir uns fragen, welches unsere erste Erinnerung ist, dann tritt ein Bild in das Bewusstsein, und dieses Bild bezieht sich auf einen bestimmten Ort und ein bedeutsames Ereignis, das uns nicht mehr loslässt. Bildliches Wissen als Erinnerungswissen ist grundlegend für die Gestaltung eines Rahmens, in dem das Selbstwissen gefasst ist. Unsere persönliche Identität wird maßgeblich durch dieses episodische Wissen über uns selbst bestimmt, und es ist im Grunde nicht mitteilbar.

Wie könnte ich anderen die Bedeutung jener Bilder vermitteln, die in mir ein Leben lang aufbewahrt werden? Man muss in eine ganz andere Dimension gehen, nämlich in die Künste, wenn man eine Vorstellung gewinnen will über die mögliche Kommunizierbarkeit dieser Form des Wissens. In Werken mancher Künstler mag man plötzlich Anklänge finden für das, was man in sich trägt.

Bildliches Wissen ist uns aber noch in einer dritten Form gegeben, nämlich als abstrahierendes Wissen, ohne das die Wissenschaften nicht auskämen. Während das Erinnerungswissen aus der persönlichen Vergangenheit durch Ich-Nähe und individuelle Bedeutung geprägt ist, das sich durch jeweils nur ein einziges Ereignis, ein einmaliges Erleben in uns einprägt, bezieht sich bildliches Wissen als abstrahierendes Wissen auf Strukturen, geometrische oder topologische Anordnungen, auf Wissen, das in Diagrammen oder Histogrammen repräsentiert werden kann, und das wir aus der Distanz betrachten. Diese Form des bildlichen Wissens ist Gegenstand der Geometrie, wie sie in der Antike durch Euklid begründet wurde, und Vorstellungswissen ist in der analytischen Geometrie thematisiert, wie sie durch René Descartes entwickelt wurde. Wenn wir eine einfache funktionelle Beziehung in einem Diagramm veranschaulichen, dann wird in einem zweidimensionalen Koordinaten-System die Abhängigkeit einer Variablen y (Ordinate) von einer unabhängigen Variablen x (Abszisse) ins Bild gesetzt. Über den im Bild veranschaulichten Zusammenhang kann dann deutlich werden, dass in verschiedenen Bereichen menschlicher Erfahrung funktionelle Zusammenhänge identisch erscheinen, obwohl die Variablen aus verschiedenen Kontexten stammen. Ein typisches Beispiel ist die sigmoide oder S-förmige Beziehung zwischen zwei Variablen, die charakteristisch ist für Phänomene in der Mathematik, Physik, Chemie, Biologie, Ökologie oder Psychologie, und die Übergänge zwischen verschiedenen Zuständen beschreibt, die bekanntlich in allen Bereichen vorkommen können. Ähnliches gilt für die statistische Normalverteilung, die so genannte Glockenkurve, die von dem Mathematiker Carl Friedrich Gauß beschrieben wurde, (und die früher auf den 10-Mark-Scheinen abgebildet war), die beispielsweise beschreibt, wie sich die Größe oder das Gewicht aber auch die Intelligenz bei Menschen verteilt; wenige sind klein oder leicht oder dumm bzw. groß oder schwer oder gescheit, und die meisten liegen in der Mitte. Hier wird uns erst durch das Bild deutlich, dass es in der Natur einheitliche und sehr einfache Prinzipien gibt, die uns nicht „ins Auge fallen“, beschrieb man diese Phänomene nur mit Worten. Mühelos informiert uns das abstrahierende Wissen über die Welt und seine Gesetze mit einer einfachen mathematischen Kurve, und es erzeugt eine anschauliche Vorstellung. Die Wissenschaft konnte nur deshalb erfolgreich sein, weil das bildliche Wissen in dieser Form anderen auf einfache Weise mitgeteilt werden konnte, und so eine Gemeinschaft der Wissenden entstand.

Die drei Formen des Wissens, das explizite, das implizite und das bildliche Wissen, sind an unterschiedliche Mechanismen des Gehirns gebunden, was aber nicht

bedeutet, dass sie voneinander unabhängig sind. Es gibt in unserem Erleben keine Funktion, die jeweils unabhängig von anderen Funktionen sein könnte, und dies gilt auch auf die verschiedenen Formen des Wissens. Die drei Wissenssysteme bilden ein gemeinsames Wirkungsgefüge, in dem jeweils nur unterschiedliche Orientierungen unseres Wissens deutlicher betont werden. Diese Orientierungen spiegeln sich beispielsweise in der „Ich-Nähe“ oder der „Ich-Ferne“ des Wissens. Explizites Wissen, das Wissen über Sachverhalte, das uns in den Stand setzt, andere zu informieren, ist Ich-fernes Wissen; es bezieht sich auf Informationen, die allen bekannt sind oder bekannt sein können. Wenn wir von „Wissensmanagement“ sprechen, dann wird auf dieses Ich-ferne Wissen Bezug genommen, das verbal und durch Anweisungen vermittelt wird, und das zwischen Trägern von Information ausgetauscht werden kann. Nur weil explizites Wissen Ich-fern ist, und man sich davon distanzieren kann, nur deshalb kann es auch unverfälscht ausgetauscht werden. (Das bedeutet natürlich auch, dass beim „Wissensmanagement“ die größte Menge menschlichen Wissens unberücksichtigt bleibt).

Das implizite Wissen ist hingegen eine Ich-nahe Wissensform. Wenn wir uns intuitiv zu einer Handlung entscheiden, wenn wir in eingeübte Abläufe und Rituale des Alltags eingebunden sind, wenn wir automatisierte Bewegungen durchführen, dann geschieht dies gleichsam ohne eigenes Zutun aus uns heraus. Es geschieht mit uns, doch wird das Handeln nicht als Ich-fremd, sondern als ein Geschehen erlebt, das ein Teil von uns ist. Wiederum können hier Beobachtungen an Patienten ein besseres Verständnis erschließen; es kann geschehen, dass nach bestimmten Störungen im Gehirn einzelne Körperteile, beispielsweise eine Hand, als Fremdkörper empfunden werden, als würden sie nicht mehr zum Selbst gehören. Solche Beobachtungen zum „neglect“ zeigen, dass die Ich-Nähe des impliziten Wissens keine Selbstverständlichkeit ist, sondern durch spezifische neuronale Prozesse bereit gestellt wird. Die Ich-Nähe des impliziten Wissens zeigt sich auch in unseren Ausdrucksbewegungen, die sich uns manchmal instinktiv aufdrängen. Der Ausdruck unserer Gefühle beruht auf jenen uns mitgegebenen Bewegungsformen, die mühelos hergestellt und mühelos verstanden werden können, und die als unmittelbar zu uns gehörend empfunden werden.

Wie das implizite Wissen so ist auch das bildliche Wissen durch Unmittelbarkeit und Ich-Nähe gekennzeichnet sein. Dies trifft insbesondere für das Erinnerungswissen zu. Das Wissen um unsere Identität, das auf Bildern unserer Lebensgeschichte beruht, begründet sich in jenen Bildern, die wir mit niemandem teilen können; sie sind in höchstem Maße subjektiv. Auch das gegenwärtige Anschauungswissen ist durch Ich-Nähe gekennzeichnet, denn sinnliche Wahrnehmung ist unmittelbar, und es ist immer nur meine Wahrnehmung. Die Ich-Nähe des Anschauungswissens können wir aber aufheben, wenn wir uns in den Zustand des Inspizierens versetzen. Man kann in den sternenklaren Himmel schauen und sich als Teil des Universums

empfinden, oder sich darin verloren finden, und andererseits kann man die Sterne zählen.

Bildliches Wissen ist aber auch abstrahierendes Wissen, und als solches ist es distanziertes Wissen über Sachverhalte, das Wissen in das Bild setzt, damit ich im Bilde bin; („man kann die Sterne zählen“). Doch erweist sich das abstrahierende Wissen auch als Ich-nahes Wissen, wenn ein weiterer Gesichtspunkt berücksichtigt wird, der für alle visuelle Erfahrung gilt. Bildlich vermittelte Zusammenhänge müssen eine ästhetische Qualität haben, und nur wenn sie diese haben, dann wird auch Information durch das Bild vermittelt, und es entsteht Wissen. Eine S-förmige Kurve in einem Koordinatensystem gibt auf ästhetische Weise der funktionellen Beziehung, die zum Ausdruck gebracht werden sollen, Bedeutung und Gewicht. Offenbar nur dann, wenn ein Sachverhalt in einfacher Weise zum Ausdruck gebracht wird, kann dieser im abstrahierenden Wissen des Betrachters nachhaltig verankert werden. Hier verbirgt sich ein fundamentales Prinzip unseres Wahrnehmens und Erkennens, nämlich nur dann etwas als richtig oder wahr begreifen zu können, wenn es in einfacher Form, sei es in einem einfachen Bild oder in einer einfachen Formel, dargestellt werden kann. Die Schönheit einer Lösung gilt in den Naturwissenschaften geradezu als ein Kriterium für deren Richtigkeit.

Hier stellt sich die Frage, ob das ästhetische Prinzip nicht allgemein als ein Rahmen für Wissen gilt, also nicht nur für das Vorstellungswissen, um graphisch Sachverhalte in möglichst klarer Weise zu veranschaulichen, sondern dass es für alle Wissensformen gilt, und dass damit eine innere Verbindung zwischen den verschiedenen Formen des Wissens gegeben ist. Dies ist in der Tat meine These, dass also das „ästhetische Wissen“ ein Grundprinzip ist, und daß Inhalte der drei Wissensformen nur dann in uns verankert sind, wenn sie dem ästhetischen Prinzip gehorchen.

Explizites Wissen strebt immer nach Ordnung, um in die Unübersichtlichkeit der uns umgebenden Information ein System zu bringen. Explizites Wissen muss klassifiziert sein, und wenn Ordnung nicht gegeben ist, entwickelt man eigene Schemata. Das Kriterium für eine gelungene Klassifikation ist deren Stimmigkeit, Einfachheit und Klarheit; dies sind ästhetische Kriterien. Für das implizite Wissen gilt ebenfalls Stimmigkeit als Kriterium; wenn Handlungs- oder Bewegungsabläufe harmonisch sind, wenn sie eine Gestalt bilden, wenn ein wissenschaftliches Experiment erfolgreich durchgeführt wird, dann wird dies alles als richtig, als befriedigend und auch als schön empfunden.

Für das Anschauungswissen und das Erinnerungswissen gilt das ästhetische Prinzip in seiner ursprünglichen Weise. Der Begriff der Ästhetik leitet sich aus dem griechischen „aisthesis“ ab und meint ursprünglich Wahrnehmung, Gefühl und auch Erkenntnis. An diese Bedeutung des Begriffs lehnt sich Immanuel Kant an, wenn er

die Ästhetik als die „Wissenschaft von den Regeln der Sinnlichkeit überhaupt“ bezeichnet. Wenn Ästhetik in diesem ursprünglichen Sinne verstanden wird, also nicht eingegrenzt wird auf die Philosophie der Kunst oder die Theorie des Schönen, dann gilt in der Tat das ästhetische Prinzip auch für das Anschauungs- und das Erinnerungswissen. Ein Kennzeichen des inneren Theaters (Anschauungswissen) oder des inneren Museums (Erinnerungswissen) ist der Rahmen. Bildliches Wissen ist immer begrenzt, und in dem Rahmen, der durch die Begrenzung vorgegeben ist, repräsentiert sich eine vergangene oder gegenwärtige Wirklichkeit. Für diese Repräsentation gilt nicht das Kriterium der Schönheit, sondern das Kriterium der Bedeutung. Wie in einem Bild eines Künstlers muss diese Bedeutung erkennbar sein, und sei sie durch Verzerrungen oder Symmetriebrüche noch so verfremdet. In unserem inneren Museum wird nichts aufbewahrt, was langweilig ist, in unserem inneren Theater wird nichts aufgeführt, was nicht unsere Aufmerksamkeit auf sich zieht. Für die Bilder, die wir in uns tragen, und die wir uns von Augenblick zu Augenblick schaffen, gilt das ästhetische Prinzip also in seinem eigentlichen Sinn, nämlich als Ausdruck der Weise unserer Erfahrung, wie sie vom Gehirn und den Sinnesorganen vorgegeben wird.

Wenn das ästhetische Prinzip für jede der Wissensformen gilt, dann gilt es auch für das Zusammen des expliziten, impliziten und bildlichen Wissens. Jede der drei Formen des Wissens ist wesentlich, und keine Form des Wissens kann für sich alleine stehen, (auch wenn wir uns in unserer kulturellen Tradition vorzugsweise auf das explizite Wissen konzentrieren). Würden wir nur explizites Wissen kultivieren, dann würden wir uns genauso zu Karikaturen unserer selbst machen, wie wenn es für uns nur implizites oder bildliches Wissen gäbe. Nur explizites oder begriffliches Wortwissen ist *unfruchtbar*. Nur implizites oder intuitives Wissen ist *ziellos*. Nur individuelles Bildwissen ist *unverbindlich*. Auf keine Form des Wissens können wir als einzelner oder auch als Gemeinschaft, insbesondere auch als Gemeinschaft der Wissenschaftler, verzichten; alle drei Koordinaten des Wissens müssen bestimmt sein.

Wenn sich in den verschiedenen Formen des Wissens explizite Sachverhalte, implizite Abläufe oder bildliche Repräsentationen zeigen, und dieses Sich-Zeigen nach dem ästhetischen Prinzip erfolgt, gibt es neben diesem formalen Prinzip der Abbildung, wie also etwas abgebildet wird, ein weiteres Prinzip, das die drei Wissensformen miteinander verbindet, das sich auf das bezieht, was abgebildet wird? Dies ist der Fall. Das ästhetische Prinzip wird ergänzt durch das mimetische Prinzip. Wissen ist Nachahmen, Spiegeln, Abbilden, Verdoppeln. „Mimesis“ ist das Grundprinzip des Wissens.

Wie steht es mit dem bildlichen Wissen hinsichtlich des Nachahmens? In einer unmittelbaren Weise entspricht der Seheindruck der Wirklichkeit. Was ich sehe, ist die Kopie der Welt, die sich in meinen Augen spiegelt, und die in meinem

Wahrnehmungssystem als Seheindruck festgehalten wird. Anschauungen sind der unmittelbare Ausdruck des mimetischen Prinzips. Aber dennoch: was ich sehe, ist nicht eine photographische Kopie dessen, was auf dem Augenhintergrund, auf der Retina, abgebildet wird, denn Verarbeitungsprozesse des Gehirns verwandeln die optischen Informationen in einen Seheindruck mit Bedeutung; ohne die optischen Informationen jedoch gäbe es keine Abbildung und somit auch keine Anschauung. Das bildliche Anschauungswissen entspricht somit dem mimetischen Prinzip, doch dieses nicht als ein passives Spiegeln oder gedankenloses Verdoppeln der äußeren Wirklichkeit, sondern als ein aktives Gestalten von Wirklichkeit, indem eine verbindende Brücke zwischen der optischen Information, die über die Augen aufgenommen wird, und der subjektiven Bedeutung dieser Information, hergestellt wird. Diese verbindende Brücke wird von Aufmerksamkeitsprozessen gesteuert, die auswählen, welche Informationen zum gegebenen Zeitpunkt „sehenswert“ gemacht und als inneres Bild erzeugt werden. Sehen als Grundlage für das Anschauungswissen ist ein aktiver Vorgang, der ins Leere laufen würde, gäbe es nicht das Abzubildende. Bei allen Lebewesen mit Augen findet ein aktives Verdoppeln statt, damit das Gesehene, was in einem gegebenen Augenblick an einem bestimmten Ort wichtig ist.

Diese aktive Verdoppelung der Welt gilt auch für das Erinnerungswissen. Die Bilder, die sich in unser episodisches Gedächtnis seit der frühesten Kindheit eingeprägt haben, halten für immer jene Ereignisse fest, die im Augenblick des Erlebens für den Einzelnen bedeutsam waren. Es ist wiederum nicht das photographische Bild, das gespeichert wird, sondern die inneren Bilder sind verändert und auch inszeniert, indem sie nur Wichtiges festhalten, und damit in die persönliche Lebensgeschichte hineinpassen. Diese Bilder gäbe es jedoch nicht, wenn nicht Information mit den Sinnesorganen aufgenommen worden wäre. Die Komplementarität der Informationen von außen und der Gestaltung von innen ist Grundlage des Erinnerungswissens.

Das Festhalten von Bildern im Erinnerungswissen steht einerseits im Bezug zu unserer personalen Identität; wir versichern uns unserer Selbst durch das Erinnerungswissen, indem wir in diesen Bildern im Bewusstsein spiegeln, was vergangen ist, und dieses Spiegeln erfolgt in der Gegenwart. Wenn ich jetzt ein Bild aus meinem Erinnerungswissen aufrufe, dann weiß ich aber auch, dass ich dies morgen wieder machen kann; somit ermöglicht mir das Erinnerungswissen, mich in die Zukunft zu projizieren. Das Erinnerungswissen ist damit ein zeitliches „Brückenwissen“, das mich aus der Vergangenheit in die Zukunft trägt, indem ich mich immer wieder selbst verdoppeln kann. Diese Selbstspiegelung verlangt aber auch, dass ich zu mir selbst eine Außenperspektive einnehmen kann, dass ich also weiß, (oder zumindest davon ausgehe), dass ich morgen noch derselbe bin.

Auch für das implizite Wissen gilt das mimetische Prinzip, doch in einer anderen Ausprägung: Eine Bewegung, ein Ritual, das heuristische Erfahrungswissen des

Experimentators, dies sind Wiederholungen von im Gehirn gespeicherten Programmen, die genetisch oder gelernt zur Verfügung stehen. Intuitionen als ein weiterer Ausdruck des impliziten Wissens repräsentieren Lösungen; sie sind Antworten auf Fragen. Doch was sind Lösungen? Sie sind Abbildungen zwischen Fragen und Antworten, die sich in der Realität bewähren müssen, und als Intuition mit seinem impliziten Probehandeln schon bewährt haben. Während explizites Wissen eine abstrahierte Wirklichkeit repräsentiert, und durch Ich-Ferne gekennzeichnet ist, sind das implizite und das bildliche Wissen sind durch eine Verschränkung der Erfahrungswelt mit dem „Wissenden“ charakterisiert; diese Ich-Nähe des Wissens gibt Vertrauen und Sicherheit.

Für das topologische Wissen gilt das mimetische Prinzip in einer ganz anderen Weise. Diese Form des bildlichen Wissens entspricht dem expliziten Wissen auf der begrifflichen Ebene, indem ein Abbild einer reduzierten Realität auf abstrakter Ebene vorgenommen wird. Eine Teilmenge dessen, das in einem anschaulichen Rahmen repräsentiert ist, wird herausgegriffen, und es wird explizit gemacht, indem es benannt oder topologisch (etwa durch eine einfache Graphik) dargestellt wird. Manchmal wird das explizite Wissen formelhaft festgehalten, wie es typisch für die Naturwissenschaften ist. Etwas sagen und etwas formalisieren ist dasselbe. Was als explizites Wissen festgehalten wird, ist ein Abzug, eine Kopie aus der Realität in einem neuen Rahmen, den man Wissensrepräsentation nennen kann, und dieses erfolgt in einer anderen Sprache, in einem neuen Medium. Üblicherweise wird eine abstrakte Ebene angestrebt, auf der ein verallgemeinertes Abbild von Sachverhalten vorgenommen wird. Mimesis ist in diesem Fall abhängig von der vorausgehenden Abstraktion.

Doch wie kann man dies alles kommunizieren? Es ist offenkundig, dass die drei Formen des Wissens in jeweils unterschiedlicher Weise vermittelt werden und vermittelt werden müssen. Doch wissenschaftliche Kompetenz und wissenschaftlicher Erfolg beruhen auf allen Formen des Wissens. Hier stellt sich eine völlig neue Herausforderung der Wissensvermittlung, die über die Repräsentation des expliziten oder des topologischen bildlichen Wissens in den üblichen Formen weit hinausgeht. Bibliotheken müssen auch Orte sein, in denen alle Formen des Wissens zur Geltung kommen.

(Dieser Text ist eine gekürzte und wesentlich überarbeitete Version des Kapitels „Drei Formen des Wissens“ aus: Ernst Pöppel: „Der Rahmen. Ein Blick des Gehirns auf unser Ich“, Hanser-Verlag München, 2006).

E-Science: Neue Formen wissenschaftlicher Zusammenarbeit

Von Cyberscience zu e-Science

Anne-Katharina Weilenmann

Einleitung

Die Fülle an Daten und Informationen hat bis heute fast unvorstellbare Ausmaße angenommen: „Die Datenmenge im Internet wird heute von Experten auf 100 Millionen Exabyte geschätzt (1 Exabyte = 10¹⁸). Dies entspricht ungefähr dem Volumen, das 100 000 Bibliotheken in der Größe der Library of Congress aufbrächten, wenn ihre Bestände digitalisiert wären (zum Vergleich: die reale Library of Congress besitzt 19 Millionen Bücher). Als „Durchschnittsleser“ kann man sich davon wohl nur einen Nano-Bruchteil erschließen“ (Börsenblatt, 2007). Der Mensch in der Informationsgesellschaft ist jedoch nicht mehr nur Leser bzw. Konsument dieser vielen Informationen, er nimmt immer häufiger auch selbst die Rolle des Informationsproduzenten ein und kann so direkt am Publikationsprozess teilhaben (Stichwort „Web 2.0“), was zu einer kaum mehr bewältigbaren Informationsflut führt. „Wir schreiben mit einer Verdopplungsrate von 20 Jahren das jeweils vorhandene Wissen immer wieder neu, und erkennen an der Halbwertszeit des Zitierten von 5 Jahren, dass wir mit der Literatur der letzten 35 Jahre bereits 99 Prozent des vorhandenen Wissens abdecken.“ (Umstätter, 2005).

Diese Zahlen zeigen eindrücklich, in welchen Dimensionen wir uns heute bewegen.

Doch was heißt „Wissen“ überhaupt? „Unter Wissen versteht man im Bewusstsein verfügbare Kenntnisse über Gegenstände, Sachverhalte, Personen, Ereignisse, Methoden, Regeln etc. einschließlich des zugehörigen lebensweltlichen (historischen) Begründungszusammenhangs – im Unterschied zu bloßen Meinungen, Vermutungen – verknüpft mit der Einsicht in ihre Gewissheit, die (objektiv) auf Tatsachenfeststellung, logischen, wissenschaftlichen Kriterien bzw. (subjektiv) auf Intuition oder eigener Anschauung beruht.“ (Kuhlen et al., 2004).

Aus den verfügbaren (Er-)Kenntnissen erzeugen wir ein Bild, welches wir immer wieder aus anderen Blickwinkeln sehen und betrachten können (Wissen stammt etymologisch von lat. videre = sehen). Wir können Teile dieses Bildes weglassen, neue Formen hinzufügen. Auf diese Weise entsteht ein Werk, das sich stetig verändert und sich dabei selbst wieder neu erschafft, wie dies Owusu-Ansah (2005) treffend ausdrückt: „Wenn aber Wissen als Teil eines sich im Kreis drehenden intellektuellen Universums wieder zum Ausgangspunkt zurückkehren könnte und auf der nächst höheren Ebene erneutes Wachstum anstoßen würde, also quasi wieder zur Information würde, dann wäre die Chance, neues Wissen zu kreieren ungleich größer.“

Das Hinzufügen neuer Formen und damit das kontinuierliche Verändern derselben bezieht sich nicht mehr allein auf mentale Prozesse, sondern auch auf technische Vorgänge im weltweiten Netz, dies begründet durch die Entwicklung der

Informations- und Kommunikationstechnologien. Die Struktur des Internets verdeutlicht exemplarisch, wie „Wachstum anstoßen“ und damit die Generierung neuen Wissens verstanden werden kann. Durch die Verlinkung der einzelnen Webseiten kann sich der Internetbenutzer von Punkt zu Punkt bewegen, dies in beliebiger Reihenfolge. Im Internet wird die Linearität des Geschriebenen aufgehoben. Der Leser kann in einem Text willkürlich von Abschnitt zu Abschnitt springen, einzelne Passagen herausfiltern, um diese wiederum in einen ganz anderen Zusammenhang zu setzen und daraus ein neues Gebilde (Gewebe) zu formen: „Virtuelle Bücher beruhen auf der Verknüpfung einzelner Wissensfragmente, die nicht aus einem einzigen linearen Ausgangstext stammen müssen, zu größeren (hoffentlich kohärenten) Einheiten.“ (Kuhlen, 1993). Spinnt man den Faden weiter, kann man sich zukünftig eine Informationswelt vorstellen, in der die verschiedenen Einheiten nicht mehr nur zu eindimensionalen Texten zusammengesetzt werden, sondern Mehrdimensionalität einnehmen können, wie dies Nentwich (2000) in seiner Idee einer Cyberwissenschaft formuliert: „Ich nenne diese Vision ‚subdisziplinäre‘, also in jedem Fach speziell organisierte, ‚vernetzte‘ ‚Hyperbasen‘, also Text- und Multimedia-„Räume“, die keine linearen Texte, sondern Hypertextstrukturen enthalten,... und damit eine Art dynamisches Abbild des Status Quo einer Disziplin abbilden.“

Solche Visionen mögen vielleicht gewagt sein; sie werfen vor allem auch Fragen auf: Wie gestaltet sich der Informations- bzw. Wissensraum in naher Zukunft, können wir eine „Wissensvision“ formulieren? Welche Arbeits- und Wissensformen werden sich entwickeln, in welche Richtung bewegt sich die wissenschaftliche Kommunikation? Prognosen abzugeben ist immer schwierig; doch es ist unerlässlich, sich kritisch mit aktuellen Trends und Tendenzen auseinander zu setzen. Anhand eines kurzen Rückblicks auf die noch junge Entwicklungsgeschichte von e-Science wird deren Entstehen nachgezeichnet und die wichtigsten Eckpunkte hervorgehoben. Bedeutende internationale Projekte werden vorgestellt, abschließend innovative Suchtechnologien aufgezeigt.

Die Aufzeichnung der Welten

Die massive Zunahme der Informationsproduktion, die in immer kürzeren Abschnitten geschieht, bedeutet eine große Herausforderung für die heutige, vor allem aber auch für zukünftige Generationen. Neben der inhaltlichen Qualität der Informationen sind es besonders Aspekte der Langzeitarchivierung sowie der (Wieder)Auffindbarkeit mittels ausgeklügelter Suchtechnologien, was allmählich zu einer nicht zu unterschätzenden Problematik wird und Lösungsansätze erfordert.

Auf nationalen wie internationalen Fachkongressen sind diese Themen schon seit einiger Zeit immer wieder Gegenstand von Expertendiskussionen, so auch auf der Europäischen Konferenz „Herausforderung: Digitale Langzeitarchivierung. Strategien

und Praxis europäischer Kooperation“¹, welche vom 20. bis 21. April 2007 in Frankfurt am Main stattfand. Einer der herausragenden Beiträge dieser Veranstaltung war der Vortrag von Kim Veltman vom Virtual Maastricht McLuhan Institute mit dem sinnigen Titel „Von der aufgezeichneten Welt zur Aufzeichnung der Welten“ (englischer Originaltitel: From Recorded World to Recording Worlds) (Veltman 2007). Darin entwirft er die Vision einer „allumfassenden Welt-Bibliothek“, welche aus den vier folgenden Elementen bestehen könnte: erstens die Schaffung einer „Beobachtungsstation“, mit dem Auftrag, die europäischen Gedächtnisinstitutionen (Bibliotheken, Archive, Museen) stets über neueste technologische Entwicklungen zu informieren. Zweitens soll mit dem Aufbau einer europäischen Digitalbibliothek zugleich die Errichtung eines verteilten Repositoriums, eines digitalen Referenzraums sowie einer virtuellen Agora ermöglicht werden, um die Gedächtnisinstitutionen mit aktuellen Forschungsergebnissen vernetzen zu können. Drittens regt er die Gründung eines Instituts zur Wissensorganisation an (in Anlehnung an die Vision von Otlet), viertens schliesst er seine Ideen einer „Welt-Bibliothek“ mit der Forderung, die Voraussetzungen für eine „Universal Digital Library“ zu erforschen.

Blendet man zurück, erfährt man, dass in der einschlägigen Literatur erste Denkansätze eines solchen übergeordneten Gesamt-Bildes bereits in früheren Jahren erkennbar sind. Neben Nentwich (2000), der schon im Jahre 2000 den Ausdruck „Cyberwissenschaft“ definierte, findet man diese Formulierung in anderer Prägung und Ausgestaltung bei Patkar (2006) wieder. Patkar sieht als Grundlage für eine Cyberwissenschaft zuerst den Aufbau einer „Cyberinfrastructure“, die er sich als ein Zusammenspiel der drei folgenden Faktoren vorstellt: Cyberenvironments, Cyber-Resources und Cyber-Education. Der Wissenschaftler arbeitet und bewegt sich gedanklich in einem Informationsraum, der zunehmend von Virtualität geprägt ist, er hat zeit- und ortsunabhängig Zugriff auf digitale sowie gedruckte Ressourcen und Medien, seine Forschungsergebnisse lässt er wiederum in seine Lern- und Experimentierumgebung einfließen, die mit dem Arbeitsplatz anderer Wissenschaftler vernetzt ist.

Damit die Informations- und Dokumentationseinrichtungen den Anforderungen an eine Cyberwissenschaft nachkommen können, bedarf es einer leistungsfähigen Infrastruktur, welche dem vernetzten Forschen förderlich ist. Diese Infrastruktur wird gemäß John Taylor, dem ehemaligen Direktor von Hewlett-Packard (HP), jedoch nicht mehr wie üblich eingekauft und institutionsweit zur Verfügung gestellt, sondern der Endanwender benützt und bezahlt solche IT-Dienstleistungen nur dann, wenn er diese auch wirklich benötigt (pay per use). Taylor hegt damit die Vorstellung einer Wissenschaft, die auf kollaborativer und vernetzter Basis arbeitet und führt den

¹ Herausforderung: Digitale Langzeitarchivierung. Strategien und Praxis europäischer Kooperation:
<http://www.langzeitarchivierung.de/downloads/EU/KONFERENZPUBLIKATION.pdf>
(Stand: August 2007)

Ausdruck „e-Science“ ein: „e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.“ (Taylor, 2001). Mit „e-Science“ wird jedoch keine neue Wissenschaftsrichtung begründet, „e-Science“ besteht vielmehr aus einer Sammlung von Tools und Technologien zur Unterstützung des vernetzten Forschens.

Die Terminologie ist keineswegs einheitlich; während sich in Europa der Begriff „e-Science“ etabliert hat, benützt man in den USA das Wort „Cyberscience“.

Das Bewusstsein für diese veränderte Wissenschaftskommunikation und somit für e-Science wird durch Analysen und Empfehlungen diverser nationaler Gremien gefördert. Wegweisend dabei ist der Bericht „Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure“ von Atkins (2003), der auf 84 Seiten detailliert die Vision einer „Cyberinfrastruktur“ entwickelt. Dabei wird auch den Informations- und Dokumentationseinrichtungen große Bedeutung beigemessen. In Großbritannien wird zeitgleich mit der Einführung der neuen Terminologie im Jahre 2001 ein sogenanntes e-Science-Programm ins Leben gerufen, mit dem Ziel, innovative technische Lösungen voranzutreiben. Eine e-Science-Initiative gibt es ebenfalls in Deutschland, gefördert vom Bundesministerium für Bildung und Forschung (BMBF); im Mittelpunkt steht dabei der Aufbau der digitalen Infrastruktur für das wissenschaftliche Arbeiten der Zukunft. Diese Thematik wird erstmals am e-Science-Forum vom Oktober 2005² zum Leitgedanken gemacht; die erste Konferenz zu e-Science in Deutschland, die den aktuellen Forschungsstand widerspiegelt, lief unter dem Namen "German E-Science Conference 2007", GES2007³, organisiert von der Helmholtz-Gemeinschaft, der Max-Planck-Gesellschaft und der Hochschul-Rektorenkonferenz (HRK).

Projekte

Neues, noch Unbekanntes wird immer auch begleitet durch das Ausloten bestehender Möglichkeiten innerhalb eines Testfeldes. So sind verschiedene Projekte initiiert worden, um die Nachhaltigkeit und Praxistauglichkeit von e-Science zu untersuchen.

Besondere Erwähnung verdienen das EU-Projekt „DILIGENT“ sowie als Beispiel aus den Geisteswissenschaften das Parzival-Projekt.

Bei DILIGENT, Digital Library Infrastructure on Grid ENabled Technology⁴, handelt es sich um eine dreijährige Initiative im 6. Rahmenprogramm für Forschung und technologische Entwicklung (RP6) der Europäischen Kommission, welches von 2004-2007 läuft (Castelli, 2006). Die Herausforderung besteht darin, eine

² E-Science Forum: <http://www.e-science-forum.de/> (Stand: August 2007)

³ German E-Science Conference 2007, GES2007: <http://www.ges2007.de> (Stand: August 2007)

⁴ DILIGENT: <http://www.diligentproject.org/> (Stand: August 2007)

Infrastruktur für eine Digitale Bibliothek aufzubauen, die es den Teilnehmern dynamischer virtueller Forschungsorganisationen ermöglicht, bei Bedarf eigene sogenannte „flüchtige“ digitale Bibliotheken zu kreieren, um auf verteilte Ressourcen zuzugreifen. Diese Ressourcen können das gesamte Angebot an verfügbaren elektronischen, multimedialen Inhalten und Applikationen umfassen.

Der Einsatz von e-Science-Technologien eignet sich auch vorzüglich für das Gebiet der Geisteswissenschaften, geht es dort doch oft um umfangreiche Texteditionen und Überlieferungen, welche nicht nur große Anforderungen an den vorhandenen Speicherbedarf stellen, sondern auch an die Arbeitsinstrumente zur Auswertung und Präsentation der Texte.

Als Beispiel sei hier Wolfram von Eschenbachs Parzival angeführt, einer der bedeutendsten Erzähltexte des Mittelalters, welcher aus 25 000 Versen besteht; zudem sind neben 16 fast vollständig erhaltenen Handschriften 68 Fragmente sowie ein Druck bekannt. Im Parzival-Projekt⁵ soll nun die Überlieferung untersucht und die Neuedition verfasst werden. Solche Initiativen illustrieren eindrücklich, wie hier dezentrale und kollaborative Verfahren zum Einsatz kommen, und wie so groß angelegte Projekte erst durch die Anwendung solcher Verfahren realisiert werden können. Weitere interessante und vielversprechende Anwendungen, die in den Geisteswissenschaften entwickelt werden, findet man in Aschenbrenner (2007).

Suchtechnologien

Wie werden die Suchtechnologien von morgen aussehen, wie werden die Instrumente beschaffen sein, die uns nicht nur das Suchen, sondern vor allem auch das Finden erlauben? Es würde zu weit führen, dieses Thema hier ausführlich zu behandeln, dennoch sollen einige Gedanken dazu präsentiert und innovative Ansätze vorgestellt werden.

„Search is the new interface to knowledge“ (Battelle, 2007). Mit diesem Ausspruch signalisiert Battelle die neue Art des Suchens, die sich nicht mehr länger nur auf das Formulieren einer Anfrage und die Ausgabe einer Treffermenge reduziert.

Der bedarfsgerechte Zugriff auf Expertenmeinungen und –wissen wird immer wichtiger. Wie kann der gegenwärtige Informationsraum permanent beobachtet werden, wie Expertenwissen, das erst in den Köpfen der Wissenschaftler existiert, sofort einbezogen und dauerhaft zugänglich gemacht werden, damit der Prozess der „Aufzeichnung der Welt“ für jeden einzelnen Forscher zum richtigen Zeitpunkt und nachhaltig geschehen kann? Zimmer (2001) formuliert dies sehr bezeichnend: „In ihrer fluiden Phase fließt die wissenschaftliche Information auf jede nur erdenkliche Art: bei Arbeitskonferenzen und ihren Nachsitzungen, auf Fachtagungen und in ihren Kaffeepausen, per Brief, per Telefon, per E-Mail, per Newsgroup, per Preprint, ob auf Servern oder aus dem Faxgerät, vielleicht sogar per Telepathie.... Aus der fortwährenden Diskussion gehen Ergebnisse hervor, und diese müssen jenen

⁵ Parzival-Projekt: <http://www.parzival.unibas.ch> (Stand: August 2007)

kristallinen Zustand annehmen, der „Zitierfähigkeit“ heißt und auf den dann andere aufbauen können.“ Je eher die Forschungsergebnisse also in die kristalline Phase überführt werden, desto eher können diese unter anderen Umständen, durch zusätzliche Erkenntnisse erweitert, wieder in fluiden Zustand versetzt werden.

Um sich einen, wenn möglich, vollständigen Überblick über diesen Informationsraum zu verschaffen, kann es nützlich sein zu erfahren, welche Personen gleiche oder ähnliche Forschungsthemen bearbeiten. Um solche Aufgaben erledigen zu können, bedarf es geeigneter Tools, die hier Unterstützung bieten.

Ein Anwendungsgebiet ist z.B. das „Scholarly Semantic Web“ (Tho, 2007). Auf der Grundlage des semantischen Webs wird eine Technologie ausgearbeitet, die dem Forscher bei der Suche Hilfestellung bieten soll. Dieses „Scholarly Semantic Web“ setzt sich aus drei Komponenten zusammen:

- scholarly service provider: dient dazu, aus den von den einzelnen wissenschaftlichen Institutionen und Universitäten erstellten Ontologien das „richtige Wissen“ (appropriate knowledge) herauszusuchen
- scholarly service requester: Web-Service, welcher für die Interaktion mit dem Endanwender zuständig ist
- matchmaking agent: dient der Suche eines geeigneten „scholarly service providers“, welcher die an ihn gerichtete Anfrage erledigen kann

Die Suche im „Scholarly Semantic Web“ gestaltet sich sehr ausgeklügelt und besteht aus folgenden Funktionen: simple document search (Stichwortsuche), simple author search (Autorensuche), document-clustering search (Trefferliste zu ähnlichen Dokumenten), author-clustering search (findet Autoren, die an gleichem Forschungsthema arbeiten), expert search (Suche nach Experten), trend search (identifiziert Trends auf dem Forschungsgebiet), schließlich fuzzydocument search (Ähnlichkeitssuche). Besonders interessant scheinen mir die „expert search“ und die „trend search“ zu sein, was einerseits erlaubt, das Forschungsgeschehen im eigenen Interessensgebiet zu beobachten, andererseits die immer wichtiger werdende Voraussicht für Trends gewährleistet.

Zusammenfassung

Ausgehend von der heutigen Informationsflut wurde aufgezeigt, was „Wissen“ denn eigentlich bedeutet, wie dieses Wissen generiert und verändert werden kann. Der Faden wurde aufgenommen bei einzelnen Wissensfragmenten, die (theoretisch) beliebig zusammengefügt und wieder auseinander genommen werden können, wurde dann weiter gesponnen bis zum globalen Informationsraum, der sich in ständiger Bewegung zwischen fluidem und kristallinem Zustand befindet.

Literatur

Aschenbrenner, A. et al. (2007). Von e-Science zu e-Humanities – Digital vernetzte Wissenschaft als neuer Arbeits- und Kreativbereich für Kunst und Kultur. In: Bibliothek. Forschung und Praxis, 31(1), 11-21

Atkins, D. E. et al. (2003). Revolutionizing Science and Engineering Through Cyberinfrastructure.
Im Internet: <http://www.nsf.gov/cise/sci/reports/atkins.pdf> (Stand: August 2007)

Battelle, J. (2007). Search for Tomorrow. In: Nextspace 6, 7
Im Internet: <http://www.oclc.org/nextspace/006/1.htm> (Stand: August 2007)
Börsenblatt (2007). Börsenblatt : Wochenmagazin für den Deutschen Buchhandel, 174(5), 5

Castelli, D. (2006). Digital libraries of the future – and the role of libraries. In: Library Hi Tech, 24(4), 496-503
Im Internet: http://conference.ub.uni-bielefeld.de/2006/proceedings/castelli_final_web.pdf (Stand: August 2007)

Kuhlen, R. et al. (2004). Grundlagen der praktischen Information und Dokumentation: Band 2: Glossar. 5., völlig neu gefasste Ausg. München : K.G. Saur-Verlag, 130

Kuhlen, R. (1993). Wie real sind virtuelle Bibliotheken und virtuelle Bücher? In: Herget, J. et al. Neue Dimensionen in der Informationsverarbeitung. Konstanz : Universitätsverlag, 52

Nentwich, M. (2000). Die Zukunft des wissenschaftlichen Publikationswesens im Zeitalter der Cyber-Wissenschaft : IBLC-Symposium, 21-22 Oktober. Frankfurt a. Main, 15.
Im Internet: <http://www.oeaw.ac.at/ita/ebene5/frankfurt.pdf> (Stand: August 2007)

Owusu-Ansah, E. K. (2005). Umgang mit Information und Wissen : Bibliothek, Hochschule und studentisches Lernen. In: ABI-Technik, 25(1), 27

Patkar, V., & Chandra, S. (2006). e-Research and the Ubiquitous Open Grid Digital Libraries of the Future : World Library and Information Congress : 72nd IFLA General Conference and Council. 20-24 August, Seoul, Korea
Im Internet: http://www.ifla.org/IV/ifla72/papers/140-Patkar_Chandra-en.pdf (Stand: August 2007)

Taylor, J.M. (2001). Talk given at UK e-Science Town Meeting

Tho, Q.T., Fong, A.C.M., & Hui, S.C. (2007). A scholarly semantic web system for advanced search functions. In: Online Information Review 31(3), 353-364

Umstätter, W. (2005). Wissenschaftskommunikation. In: BibSpider (2005). Wissenschaftskommunikation im Netzwerk der Bibliotheken = Scholarly communication in libraries networking. Berlin : BibSpider, 11-12

Veltman, K. H. (2007). From Recorded World to Recording Worlds. Im Internet: http://www.langzeitarchivierung.de/downloads/EU/PPT/Veltman_TEXT.pdf (Stand: August 2007)

Zimmer, D. E. (2001). Analoge und digitale Informationsmedien. In: Wissenschaftspublikation im digitalen Zeitalter : Verlage, Buchhandlungen und Bibliotheken in der Informationsgesellschaft : ein Symposium des Börsenvereins des Deutschen Buchhandels e.V., der Bundesvereinigung Deutscher Bibliotheksverbände und der Deutschen Bibliothek am 8. und 9. Februar 2001 in Berlin. Wiesbaden : Harrassowitz, 131

Semantisches und vernetztes Wissensmanagement für Forschung und Wissenschaft

Katrin Weller, Indra Mainz, Ingo Paulsen, Dominic Mainz

Zusammenfassung

Der Text befasst sich mit den verschiedenen Vernetzungsebenen einer kommenden e-Science-Landschaft. Diese werden unterschieden und zueinander in Beziehung gesetzt. Den Schwerpunkt bildet dabei die Betrachtung der Möglichkeiten zu Wissensaustausch und wissenschaftlicher Kommunikation, zur gemeinschaftlichen Wissensaufbereitung und zur semantischen Vernetzung. Bezüge zu den aktuellen Themen Web 2.0 und Ontologien werden aufgezeigt.

Abstract

This text discusses the different levels of networking approaches for a future e-science scenario. The levels will be distinguished and also interrelated. A focus is laid on the possibilities for exchanging knowledge, collaborative knowledge management and semantic integration. References to the current topics Web 2.0 and ontologies are identified.

Einleitung – Ebenen in e-Science

In der Wissenschaftsförderung wird derzeit nach Lösungen gesucht, welche Wissenschaftler in Organisations- und technischen Routineaufgaben entlasten, wissenschaftliche Arbeitsabläufe vereinfachen und eine umfassende Vernetzung von wissenschaftlichen Einrichtungen sowie effizientere Infrastrukturen ermöglichen. Auf nationaler Ebene wurden daher vielfach sogenannte e-Science Initiativen begonnen. Der Begriff e-Science steht dabei für „enhanced science“ und wurde von John Taylor, ehemals Hewlett-Packard-Manager und darauf Generaldirektor der Forschungsbeiräte beim britischen Office of Science and Technology, geprägt und 2001 in Großbritannien erstmals für die Bezeichnung eines umfassendes Forschungsförderungsprogramms eingesetzt⁶.

Von großer Bedeutung innerhalb von e-Science-Programmen ist die Erarbeitung technischer Möglichkeiten für verteilte Rechenleistung (Grid-Computing) (Foster, Kesselmann 2004). Insbesondere in den USA spricht man in dem Zusammenhang auch von Cyberinfrastructures (Hey, Trefethen 2005). Große Datenzentren und physikalische Netzwerke bilden die Basis für Bearbeitung, Auswertung und Austausch grundlegender Primär-Forschungsdaten.

Wenngleich die technische Vernetzung wissenschaftlicher Einrichtungen und verteiltes Bearbeiten rechenintensiver Aufgaben eine wichtige Basis für effizientere Forschung darstellen, so müssen für eine umfassende Neugestaltung der Wissenschaftslandschaft und eine vollständige neue Infrastruktur weitere Aspekte

⁶ UK e-Science Program: <http://www.rcuk.ac.uk/escience/default.htm>.

berücksichtigt werden. Neben den Rechnernetzen ist eine Wissensvernetzung (Knowledge Networking) unbedingt erforderlich. Wissensnetze können ihrerseits auf zwei Arten ausgelegt werden: es stehen entweder die Menschen im Vordergrund, die ihr gemeinsames Wissen zusammenbringen und kollaborativ anwenden sollen (Human Infrastructures oder Social Networking), oder es geht um die Vernetzung von Daten bzw. (digitalisierten) Wissenseinheiten (Resource Networks, Data Grid) (Hey, Trefethen 2005; Lee et al. 2006; Paulsen et al. 2007). Goble et al. (2006, S. 2) bringen die die Gesamtheit der aktuellen Bedürfnisse auf den Punkt:

„Scientific progress increasingly depends on pooling resources, know-how and results; making connections between ideas, people, and data; and finding and interpreting knowledge generated by others, in ways that may have not been anticipated when it was created. [...] It has as much to do with intelligent information management as with sharing scarce resources like large scale computing power or expensive instrumentation.“

Somit sind für eine übergreifende Wissensorganisation und neue Forschungsk Kooperationen verschiedene konkrete Aspekte von Belang (Abbildung 1). Der offene Zugang zu Publikationen und anderen wissenschaftlichen Ergebnissen (Open Access) spielt eine wichtige Rolle und wird bereits ausgiebig diskutiert (beispielsweise Cram 2005; Deutsche Forschungsgemeinschaft 2005; Guédon 2004; Harnad et al. 2004), auf die damit zusammenhängenden Fragen soll daher in diesem Rahmen nicht näher eingegangen werden. Ein weiterer Schritt zu neuer Wissenschaftskommunikation ist die Entwicklung von Plattformen für die wissenschaftliche Zusammenarbeit. Auf nicht-wissenschaftlicher Ebene gibt es bereits diverse Web-Plattformen, wo Benutzergemeinschaften zusammen Inhalte erschaffen oder aufbereiten (was im Allgemeinen als Merkmal des so genannten Web 2.0 gilt). Einige dieser aktuellen Ansätze lassen sich komplett oder zumindest teilweise in den Wissenschaftsbetrieb übertragen, z. B. gemeinschaftliches Verwalten von Literaturlisten oder Internet-Lesezeichen (Social Bookmarking), sowie verlinkte Wissensaufbereitung in Form von Wikis (Weller et al. 2007). Systeme, die sich speziell an die Bedürfnisse wissenschaftlicher Arbeit anpassen, sind für die Zukunft zu erwarten.

Ebene	Ziel	Schlagwort
Rechenleistung	Verteilte Durchführung rechenintensiver Aufgaben.	Grid, Cyber-infrastructure
Menschen	Kontakte knüpfen, Experten finden, Diskussion.	Social Networking
Daten, Dokumente	Zentraler Zugang zu Ressourcen.	Data Networks, Open Access
Wissen	Informationseinheiten plus Kontext.	Semantic Integration

Abbildung 1: Ebenen der Forschungsvernetzung.

Zur kooperativen Erstellung und Verwaltung von Inhalten kommt dann die verbesserte Wissensvernetzung auf Datenebene hinzu. Hier konzentriert sich die Forschung vor allem auf den Aufbau eines Semantic Web (Antoniou, van Harmelen 2004, Breitman et al. 2007). Dabei geht es um Informationsintegration, und zwar in Form von Semantic Integration: Wissen soll nicht nur besser auffindbar gemacht werden, sondern auch unmittelbar in einen globalen Kontext eingegliedert werden; neue Informationen werden zum bestehenden Informationsgeflecht in Beziehung gesetzt (Alexiev et al. 2005; Stuckenschmidt, van Harmelen 2005; Leser, Naumann 2007). Dabei sollen semantische Annotationen (Handschuh, Staab 2003; Uren et al. 2006) und die strukturierte Umsetzung von Wissen in Form von Ontologien helfen (Davies et al. 2003, Davies et al. 2006; Staab, Studer 2004). Ontologien dienen zudem dazu, das Wissen einer Forschungsgemeinschaft in eine eindeutige Form zu bringen, und dabei möglichst einen aktuellen Konsens zum Wissenstand abzubilden (Neumann, Prusak 2007).

Soziale Netzwerke und Communities

In jüngster Zeit ist im World Wide Web die Entwicklung zu beobachten, dass es für Internetnutzer immer einfacher wird, eigene Texte zu veröffentlichen und (Multimedia-) Inhalte einzustellen – was von einer wachsenden Zahl an Nutzern auch intensiv wahrgenommen wird. Dies gehört zu den wesentlichen Merkmalen des neuen „Mitmach-Web“ oder Web 2.0 (Alby 2007). Neue Technologien und Internetdienste schaffen dabei auch neue Kommunikationsformen und neue Kommunikationskulturen (Möller 2005). Mit der wachsenden Nutzerbeteiligung im WWW bilden sich vielfach auch Benutzer- oder virtuelle Interessengemeinschaften (*Communities*). Wenngleich aktuelle Web 2.0 Entwicklungen bisher eher selten speziell auf wissenschaftliche Zielgruppen ausgerichtet sind, so steckt in den aktuellen Trends doch Potential für den Einsatz in Forschungsumgebungen. *Social Software* (Gordon-Munrane 2006; Szugat et al. 2006), insbesondere spezielle Social-Networking-Plattformen (wie beispielsweise XING⁷ oder Facebook⁸) bieten genau das, was es braucht, um wissenschaftliche Personennetzwerke abzubilden, zu pflegen und weiterzuentwickeln: jeder Nutzer kann ein persönliches Profil anlegen und dabei auch bestehende Kontakte zu anderen Plattformnutzern mit angeben. Auf diese Weise werden die Beziehungen zwischen der Nutzergemeinschaft erfasst, und beispielsweise als Pfade zwischen zwei Personen angezeigt.

Ein solches System hilft, bestehende Kontakte trotz Arbeitsplatz- oder Ortswechsel nicht aus den Augen zu verlieren, da jeder seine Kontaktdaten stets aktuell halten kann, und ermöglicht gleichzeitig die Suche nach neuen Bekanntschaften und Kooperationspartnern. Mit Hilfe von Eingabefeldern für Interessen, Kenntnisse oder Spezialisierungen im persönlichen Profil können so Experten für besondere Fragestellungen oder potentielle Projektpartner für Forschungskooperationen gefunden werden. Für den Einsatz in Wissenschaft-Communities müssten die

⁷ XING: <http://www.xing.com>.

⁸ Facebook: <http://www.facebook.com>.

Optionen für die Profilgestaltung und damit die Suchfelder entsprechend spezifisch angelegt werden. Denkbar wäre auch, bereits bestehende Kontaktketten bei einer Suche zu berücksichtigen und Suchergebnisse nach Bekanntschaftsgraden zu sortieren, um so eine erste Kontaktaufnahme zu erleichtern (Braun et al. 2007). Erweiterungsmöglichkeiten für die Vernetzung von Wissenschaftlern könnten darin bestehen, dass die Art einer Bekanntschaft spezifiziert werden kann, und so z. B. Co-Autoren oder Partner innerhalb von Forschungsprojekten besonders gekennzeichnet werden.

Aktuell können sich innerhalb von Networking-Plattformen in der Regel Interessengruppen formieren, die als Rahmen für thematische Diskussionen genutzt werden. Wenn sich hieraus tatsächlich aktive *Communities of Practice* (Timbrell et al. 2007, Wenger 2002) zusammenfinden, die durch gemeinsame Interessen oder Fragestellung verbunden sind, ihre Erfahrungen und Sichtweisen austauschen und einander bei Problemen helfen, so ist ein wichtiger Baustein für vernetzte und ortsungebundene Forschung und Wissensaustausch getan.

Die Kommunikation innerhalb solcher Communities findet nahezu ausschließlich schriftlich statt (Timbrell et al. 2007). Das bedeutet, dass hier zahlreiche neue Texte produziert, gespeichert und ggf. referenziert werden – die internen Gruppen-Diskussionen sind eine von zahlreichen neuen Formen wissenschaftlichen Kommunikationsverhaltens.

Kategorie	Gruppen-Anzahl	Kategorie	Gruppen-Anzahl
Alumni	1.384	Kunst und Kultur	166
Branchen	463	Politik und Gesellschaft	71
Community und Soziales	189	Regionales	581
Dies und Das	39	Themen	296
Events	66	Verbände und Organisationen	450
Freizeit und Sport	296	Wirtschaft und Märkte	311
Geographie und Umwelt	19	Wissenschaft	129
Internet und Technologie	361	XING	21
Jobs und Karriere	209		
Aktivste Gruppen (nach Artikeln)		Größte Gruppen (nach Mitgliedern)	
Das WortWerk (Spiel mit Worten und Begriff)	30.780	Absolventen – Gesuche und Angebote - Praj	67.282
Akquisition / Kundengewinnung	7.176	Freiberufler Projektmarkt / freelance projects	41.542
		Neue Gruppen	
		Offshore-Standort VAE - Null Steuer, uneingetragene Zweigniederlassung	
		XING Greater Washington DC Region	

Abbildung 2: Thematische Gruppen bei XING, sortiert nach Kategorien, Aktivität, Größe. Quelle: <http://xing.com>, 14.08.2007.

Die Business-Networking-Plattform XING beherbergt derzeit rund 5000 thematische Gruppen, 129 davon sind der Kategorie „Wissenschaft“ zugeordnet⁹ (Abbildung 2). Ihre Ausrichtung reicht von sehr eng gesetzten Zielgruppen und Thematiken (z. B. „Metallpulverspritzguss: Hart wie Stahl – Formbar wie Kunststoff“, 10 Mitglieder, „Schlafmedizin“, 30 Mitglieder), bis hin zu Themen von aktuellem oder allgemeinem Interesse (derzeitige Spitzenreiter: „Geistreich – Interdisziplinärer Dialog“ mit 2.958, „Arbeits- und Organisationspsychologie“ mit 2456 und „Life Sciences“ mit 2340 Mitgliedern). Hinzu kommen forschungsrelevante Gruppen, die anderen Kategorien zugeordnet sind. Ein Beispiel für eine Social-Networking-Plattform mit gezielter Ausrichtung auf (Natur-) Wissenschaftler ist Nature Network¹⁰ der Nature Publishing Group, diese befindet sich allerdings noch im Aufbau und verzeichnet bislang eher geringe Nutzerzahlen. Weitere Beispiele für wissenschaftliche Communities werden ebenfalls besonders im Bereich der Life Sciences diskutiert und beobachtet (Clark, Kinoshita 2007; Neumann, Prusak 2007).

Ein Problem aktueller sozialer Netzwerke und anderer Community-Portale besteht darin, dass die zugehörigen Systeme auf die aktive Teilnahme (angefangen mit der Registrierung) der Mitglieder angewiesen sind. Vertrauenswürdigkeit und Sorgfalt im Umgang mit Personendaten sind die minimalen Anforderungen, damit diese grundlegende Hürde überwunden wird. Weitere Anreizsysteme, beispielsweise verfeinerte und explizite Möglichkeiten für den Aufbau von Reputation, werden benötigt um insbesondere Wissenschaftler zur aktiven Beteiligung zu bewegen. Damit es zum Wissensaustausch innerhalb von Communities kommen kann, muss eine Vertrauensbasis geschaffen werden, dass Beiträge nicht zweckentfremdet oder missbraucht werden (Neumann, Prusak 2007).

Social Networking Tools können als eigenständige Portale aufgebaut werden, sind aber oft auch Bestandteil anderer Social Software Angebote oder werden mit solchen kombiniert. So können beispielsweise gemeinsame Terminkalender¹¹ über Konferenzen und wissenschaftliche Veranstaltungen geführt werden. Nutzer können angeben, welche Veranstaltungen sie besucht haben, so dass darüber die soziale Vernetzung erweitert wird und ein neuer Einstieg für die Suche nach Personen mit ähnlichen Interessen ermöglicht wird. Besuchte Veranstaltungen können kommentiert und mit Fotos oder Dokumenten ergänzt werden.

Gemeinschaftliche Wissenssammlung und -aufbereitung

Besonderes Potential ergibt sich aus der Kombination von Community-Netzen mit Systemen zur gemeinsamen Verwaltung von Dokument- oder Datenbeständen. Mit steigenden Publikationszahlen und zahlreichen Datenarten werden mehr als jemals

⁹ Diese und die folgenden Daten beziehen sich auf den Stand vom 15.08.2007 und den deutschsprachigen Zugang zur Plattform XING.

¹⁰ Nature Network: <http://network.nature.com>.

¹¹ Beispielsweise in einer Form wie bei Yahoos Portal Upcoming: <http://upcoming.yahoo.com>.

zuvor einfache aber effiziente Wege benötigt, um gezielt passende Informationen zu finden und Relevantes von Irrelevantem zu trennen. Eine Errungenschaft des Web 2.0 ist, dass Nutzer sich nicht nur aktiv an der Bereitstellung von Inhalten beteiligen, sondern auch damit beschäftigt sind, diese zu organisieren und zu indexieren (bzw. zu *taggen*). Ein früher Web 2.0 Ansatz für gemeinsame Aufbereitung von Informationen ist Wikipedia¹², der Vorreiter und bekannteste Vertreter einer Gruppe enzyklopädischer Wissensaufbereitung in Wiki-Systemen.

Das aktuelle Erfolgsmodell der kollaborativen Wissens- und Datenorganisation ist eine Kombination aus gemeinschaftlicher Datenverwaltung und einer Indexierung über *Social Tagging*, das den Nutzern erlaubt, frei wählbare Schlagworte zu einzelnen Einträgen zu vergeben; es entstehen sog. Folksonomies als neue Erschließungsmethoden und Suchhilfen (Fichter 2006; Gordon-Murnane 2006; Macgregor, McCulloch 2006; Stock 2007). Solche *Filesharing* oder *Social Bookmarking* Systeme existieren derzeit im Privatanutzerbereich für das Verwalten von Fotos und Videos¹³, für die wissenschaftliche Nutzung ist das gemeinsame Verwalten von Weblinks und bibliographischen Angaben¹⁴ interessant (Hammond, Hannay 2005). Portale für die nutzerzentrierte Archivierung von wissenschaftlichen Bild- und Multimediadateien oder gar experimentelle und primäre Forschungsdaten gibt es bislang nicht, ließen sich jedoch auf ähnliche Weise anbieten.

In den meisten Fällen legen Nutzer die Tags (oder alternative Strukturierungen) in erster Linie für den eigenen Gebrauch an; gleichzeitig besteht die Möglichkeit, die Herangehensweise anderer Nutzer einzusehen und alternative Perspektiven auszunutzen (Hammond, Hannay 2005). Der Ansatz von Wu und Gordon vereint hierarchische Ordernavigation mit nutzergenerierten Links zwischen Dokumenten, wobei die Nutzer ihre Hierarchien mit anderen teilen können und die Strukturierungsansätze anderer Nutzer einsehen und nutzen können (Wu, Gordon 2007). Besonders interessant ist das, wenn man gleichzeitig über ein Soziales Netz Informationen darüber erhält, wessen Ansichten man dort vor sich hat (Braun et al. 2007). Über gemeinschaftliche Wissensorganisation können zudem wiederum Experten gefunden werden (Diederich, Iofciu 2006; Wu, Gordon 2007).

Neben Folksonomies stehen weitere Möglichkeiten für den Zugang zu Inhalten zur Verfügung, man spricht insgesamt von *Social Navigation* oder *Collaborative Filtering*. Diese beruhen auf direkt oder indirekt gewonnenen Nutzerurteilen; indirekt über durch Logfile-Untersuchungen ermittelte Daten zur Navigation, direkt über vom

¹² Wikipedia: <http://www.wikipedia.org>.

¹³ Die bekanntesten Beispiele sind Flickr (<http://www.flickr.com>) und YouTube (<http://www.youtube.com>), ähnliche Angebote bieten beispielsweise Zoomr (<http://www.zoomr.com>), Buzznet (<http://www.buzznet.com>) oder Bubbleshare (<http://www.bubbleshare.com>).

¹⁴ Beispielsweise Connotea, Bibster (<http://bibster.semanticweb.org>), CiteULike (<http://www.citeulike.org>) und Bibsonomy (<http://www.bibsonomy.org>), Delicious (<http://del.icio.us>), Furl (<http://www.furl.net>), Spurl (<http://www.spurl.net>), LibraryThing (<http://www.librarything.com>).

Nutzer selbst abgegebene Bewertungen oder Kommentare (Forsberg 1998). Die bekannteste Ausprägung sozialer Navigationsmechanismen sind Vorschlagsysteme (*Recommender Systems* bzw. *Recommendation Systems*) (Resnick, Varian 1997), die vor allem im E-Commerce eine wichtige Rolle spielen¹⁵. Mit ihrer Hilfe lassen sich aber auch beispielsweise die meistgelesenen oder am besten bewerteten Artikel einer Nachrichtenplattform ermitteln und empfehlen oder innerhalb eines Bookmarking-Systems die Webseiten, die Nutzer mit ähnlichem Profil wie dem eigenen, zusätzlich in ihren Favoritenlisten vermerkt haben.

Ontologien und semantische Informationsintegration

Gemeinschaftliches Tagging von Dokumenten und anderen Informationseinheiten ist ein wesentlicher Aspekt für neue Formen der Wissensaufbereitung, reicht aber für die erhöhten Anforderungen an Wissenschaftsinformationen im Forschungskontext langfristig alleine nicht aus, wenngleich sie einige Vorteile bieten (Peters, Stock 2007). Als weiterführenden Schritt konzentriert sich die Forschung auf den Aufbau von semantischen Integrationsansätzen im Sinne des Semantic Web. Damit erreichen wir die oberste und schwierigste Ebene einer vernetzten Wissenschaftslandschaft. Kontrollierte und strukturierte Metadaten sollen eine Kontexteinbettung und semantische Suchen ermöglichen (Bügel 2005; Koivunen, Swick 2004).

Kernstück der Semantic Web Forschung sind Ontologien. Ontologien (im engen Sinne¹⁶) sind formale Wissensrepräsentationen. Sie erlauben eine strukturierte, maschinenlesbare Umsetzung von Wissen. Die Grundstruktur wird in Form von Konzepten bzw. Klassen und den Beziehungen zwischen diesen angelegt. Zusätzlich werden Instanzen als Referenz auf reale Entitäten eingefügt¹⁷. Mit Hilfe von Regeln können Merkmale und Kriterien für die Klassenzugehörigkeit definiert werden, die Verwendung formaler Logiken lässt Schlussfolgerungen und Ableitungen impliziter Zusammenhänge zu. Unter Ausnutzung komplexer Ontologiesprachen wie OWL-DL¹⁸ lassen sich Ontologien nicht nur wie herkömmliche kontrollierte Vokabulare zur Annotation verwenden, sie eignen sich auch dazu, Wissensausschnitte in eindeutiger Form festzuhalten. Diese können somit als gemeinsame Wissensbasis innerhalb von Communities aber auch zwischen Mensch und Maschinen oder Maschinen untereinander dienen.

¹⁵ Bekanntestes Beispiel: Der Online-Versandhändler Amazon ermittelt Produktvorschläge anhand von Verkaufsdaten sowie mit Daten zu Nutzerverhalten und Navigation.

¹⁶ In der Literatur wird „Ontology“ vielfach auch als Sammelbegriff für unterschiedlichste Arten der Wissensrepräsentation verwendet.

¹⁷ Ähnlich den Individualbegriffen in traditionellen Formen der Wissensrepräsentation. Dem Konzept „Auto“ können so einzelne PKW-Modelle, etwa ein Toyota Corolla, als Instanz zugewiesen werden.

¹⁸ OWL: Web Ontology Language, siehe <http://www.w3.org/2004/OWL/>. OWL ist in den Dialekten OWL lite, OWL-DL (description logics) und OWL full verfügbar, wobei derzeit OWL-DL die gängigste Variante ist.

Ontologien ermöglichen es, wissenschaftliche Erkenntnisse eines Themenbereichs hierarchisch zu klassifizieren und mit verschiedenen Eigenschaften weiter zu definieren, um so einen gemeinsamen Wissenshintergrund festzuhalten. Neue Erkenntnisse müssen dann nach und nach in bestehende Schemata eingefügt werden, was unter Umständen eine Umstrukturierung bereits vorhandenen Wissens erfordert. Das so zusammengestellte formalisierte Wissen kann als Referenzwerk dienen, Wissenschaftler können darin Informationen nachschlagen oder Definitionen überprüfen. Über Reasoning Mechanismen lassen sich darüber hinaus die Zusammenhänge erfassen, die nicht explizit in die Ontologie eingegeben wurden. Nun kommt es in erster Linie darauf an, semantische Technologien bzw. Ontologien mit den oben beschriebenen Netzwerken zu verbinden. Hierzu gibt es bereits zahlreiche Ansätze mit ganz unterschiedlichen Ausgangspunkten. Zum Spektrum gehören beispielsweise:

Semantisches Grid: Semantic Grid Projekte befassen sich mit einer Verbesserung der Interoperabilität aktueller Grid-Applikationen (Corcho et al. 2006). Semantische Strukturen helfen hier in erster Linie bei der Maschine-Maschine Kommunikationen und sollen einen verbesserten Workflow zwischen einzelnen Anwendungen gewährleisten.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

  <foaf:Person rdf:nodeID="harry">
    <foaf:name>Harry Osborn</foaf:name>
    <rdfs:seeAlso rdf:resource="http://www.osborn.com/harry.rdf"/>
  </foaf:Person>

  <foaf:Person>
    <foaf:name>Peter Parker</foaf:name>

    <foaf:knows rdf:nodeID="harry"/>

    <foaf:knows>
      <foaf:Person>
        <foaf:name>Aunt May</foaf:name>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>

</rdf:RDF>
```

Abbildung 3: Zwei Personen und die Beziehung zwischen ihnen im FOAF Format. Quelle: <http://www.foaf-project.org/2004/us/about.html>

Semantisch angereicherte Wikis, Repositories und Community-Portale: Einige Projekte konzentrieren sich auf Ontologien zur Erfassung von Projekten, Aktivitäten, Personen und Veranstaltungen (Bachler et al. 2004). Das FOAF¹⁹ Projekt setzt soziale Netzwerke in einem speziellen Formular maschinenlesbar um (Abbildung 3).

¹⁹ FOAF, Friend of a Friend Project: <http://www.foaf-project.org/>.

Eine Vielzahl an Projekten erarbeitet derzeit Modelle für die Verbindung von semantischen Technologien und gemeinschaftlicher Wissensaufbereitung. Das Projekt DBin²⁰ entwickelt beispielsweise ein System für eine durch Ontologien strukturierte Wissenssammlung für Communities. Verschiedene Interessengemeinschaften können sich hier formieren und ihr Wissen (z. B. Kommentare und Bilder) direkt an eine thematische Ontologie knüpfen (Abbildung 4). Mit Hilfe der Ontologie kann dann in den Einträgen auch semantisch gesucht werden. Auch bei OntoWiki²¹ wird die Eingabe von neuen Informationen in strukturierter Form durch Ontologien unterstützt, das WIKINGER Projekt erprobt die semantische Annotation von Dokumenten (Bröcker, Paal 2007). Auch für Wikipedia werden Ansätze zur semantischen Erweiterung erarbeitet (Völkel et al. 2006).

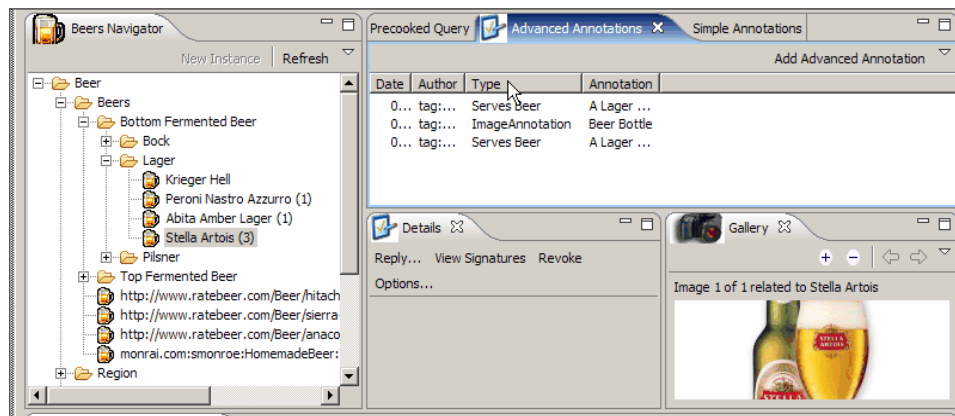


Abbildung 4: DBin. Neue Informationen werden direkt mit Instanzen der Ontologie verknüpft. Quelle: <http://www.dbin.org/screendemo/dbintease.html>, 14.08.2007.

Kollaborativ erstellte Ontologien: Wenngleich es Folksonomies an terminologischer Kontrolle mangelt, so haben sie doch den entscheidenden Vorteil, Sichtweisen und Vokabular der Nutzercommunity direkt abzubilden. Diesen Effekt kann man in Ontologien übertragen, wenn diese kollaborativ erarbeitet werden. Es gibt bereits einige Ansätze für gemeinschaftlich nutzbare Ontologieeditoren²² (Bao, Hanovar 2004, Hepp et al. 2005). Im Ontoverse Projekt²³ soll zudem eine vernetzte Community – bestehend aus Fachexperten verschiedener Bereiche und Experten für Wissensformalisierung – an den kollaborativen Ontologieeditor angeschlossen werden (Paulsen et al. 2007). Auf diese Weise können direkt gemeinschaftliche Wissensbasen aufgebaut und gepflegt und der wachsende Bedarf an angepassten

²⁰ DBin: <http://www.dbin.org>, Plattform-Demonstration <http://www.dbin.org/screendemo/dbintease.html>.

²¹ OntoWiki: <http://ontowiki.net/Projects/OntoWiki>

²² Aktuelle Ansätze gibt es auch z. B. bei Soboleo (<http://soboleo.fzi.de:8080/webPortal/>), Collaborative Protégé (<http://protege.stanford.edu/doc/collab-protege/>).

²³ Ontoverse: <http://www.ontoverse.org>

domänenspezifischen Ontologien für andere Applikationen aufgegriffen werden. Gleichzeitig werden die Vorteile einer wissenschaftlichen Community für die fachliche Diskussion genutzt.

Mit diesen vielfältigen Ansätzen zeichnet sich für die Zukunft die nächste Herausforderung ab: Formalisierungen verschiedener Wissensbereiche müssen auf einer weiter übergeordneten Ebene wiederum zueinander in Beziehung gesetzt werden, Konkordanzen und Querverweise müssen aufgebaut werden. Aktuelle Forschungen im Bereich *Ontology Mapping* bereiten diese Aufgaben bereits vor. Erschwert werden diese Bemühungen vor allem noch dadurch, dass sich Ontologien durch fortdauernde Aktualisierungen und Ergänzungen im ständigen Wandel befinden können.

Fazit

Aktuelle Trendbegriffe wie e-Science, Web 2.0 und Social Software oder Ontologien und Semantic Web bzw. die jeweils dahinterstehenden Techniken und Entwicklungen sind nicht isoliert zu betrachten, sondern entwickeln zusammengenommen ein Potential für grundlegende Veränderungen. Es ergeben sich daraus für eine vernetzte Wissenschaftslandschaft insbesondere die folgenden Merkmale:

- Ressourcenverknüpfung und freier Zugang zu Wissensbeständen
- Soziale Netzwerke und Wissensaustausch
- Kollaborative Wissensaufbereitung
- Semantische Informationsintegration

Im Idealfall hätten wir es also mit Technologien zu tun, die mit vernetzter Rechenleistung wissenschaftliche Fragestellungen bearbeiten, Daten gewinnen und umfassende und allgemein zugängliche Archive betreiben. Darin liegen sowohl wissenschaftliche Primärdaten (z. B. Ergebnisse aus Experimenten) wie auch wissenschaftliche Publikationen vor. Diese werden kollaborativ gepflegt und von Communities mit neuen Ergebnissen angereichert. Interessensgruppen kommentieren und diskutieren vorhandene Daten und nutzen bereitgestellte Portale für die strukturierte Aufbereitung der enthaltenen Informationen, wobei bestehende Ontologien und semantische Technologien für die Informationsintegration genutzt werden.

Bis so ein System jedoch technisch realisiert und praktisch implementiert werden könnte, sind noch zahlreiche technische und organisatorische Probleme zu lösen, und selbst dann wird es letztlich davon abhängig sein, inwieweit die Forschungsgemeinschaft zur Teilnahme bereit ist.

Danksagung

Dank gilt den Kolleginnen und Kollegen im Projekt Ontoverse, sowie dem Bundesministerium für Bildung und Forschung (BMBF) für die Förderung des Projektes.

Literatur

Alby T (2007): Web 2.0. Konzepte, Anwendungen, Technologien. München, Wien: Carl Hanser Verlag.

Alexiev V, Breu M et al. (2005): Information Integration with Ontologies. Experiences from an Industrial Showcase, Chichester: Wiley & Sons.

Antoniou G, van Harmelen F (2004): A Semantic Web Primer, Cambridge, Mass.: MIT Press.

Bachler M, Chen-Burger J, et al. (2004): Collaborative Tools in the Semantic Grid. In: GGF11 – The Eleventh Global Grid Forum 2004, Honolulu, Hawaii.

Bao J, Honavar V (2004): Collaborative Ontology Building with Wiki@nt. A Multiagent Based Ontology Building Environment. In: Proceedings of the 3rd International Workshop on Evaluation of Ontology-based Tools (EON), Hiroshima 2004, S. 1-10.

Braun S, Hefke M, Schmidt A, Sevilmis N (2007): Im Wissensnetz. Linked Information Processes in Research Networks. In: Proceedings of the German eScience Conference 2007, Max Planck Digital Library, ID 316584.0.

Breitman KK, Casanova MA, Truszkowski W (2007): Semantic Web. Concepts, Technologies and Applications, London: Springer.

Bröcker L, Paal S (2007): WIKINGER. Wiki Next Generation Enhanced Repositories. In: Proceedings of the German eScience Conference 2007, Max Planck Digital Library, ID 316543.0.

Bügel U (2005): Technologische Trends beim Einsatz semantischer Netzwerke. In: Knowledge Extended. Die Kooperation von Wissenschaftlern, Bibliothekaren und IT-Spezialisten. 3. Konferenz der Zentralbibliothek. Schriften des Forschungszentrums Jülich, 2005, S. 225-236.

Clark T, Kinoshita J (2007): Alzforum and SWAN. The Present and Future of Scientific Web Communities. In: Briefings in Bioinformatics 8 (2007), S. 163-171.

Corcho O, Pinar A, et al. (2006): An Overview of S-OGSA. A Reference Semantic Grid Architecture. In: Journal of Web Semantics 2006 4(2).

Cram HR (2005): Die Auswirkungen der Open-Access-Initiative auf die Wertschöpfungskette. In: Knowledge Extended. Die Kooperation von Wissenschaftlern, Bibliothekaren und IT-Spezialisten. 3. Konferenz der Zentralbibliothek. Schriften des Forschungszentrums Jülich, 2005, S. 57-71.

Davies J, Fensel D, van Harmelen F (Hrsg., 2003): Towards the Semantic Web. Ontology-Driven Knowledge Management, Chichester: Wiley & Sons.

Davies J, Studer R, Warren P (2006): Semantic Web Technologies. Trends and Research in Ontology-Based Systems, Chichester: Wiley & Sons.

Deutsche Forschungsgemeinschaft (2005): Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access, Wiley VCH.

Diederich J, Iofciu T (2006): Finding Communities of Practice From User Profiles Based on Folksonomies. In: Proceedings of the 1st International Workshop on Building Technology Enhanced Learning Solutions for Communities of Practice (TEL-CoPs'06).

Fichter D (2006): Intranet applications for tagging and folksonomies. In: Online, 30(3), S. 43-45.

Forsberg J (1998): Social Navigation: An Extended Definition. Verfügbar über: www.nada.kth.se/~forsberg/Documents/ (letzter Aufruf am 12.08.2007).

Foster I, Kesselmann C (2004): The Grid. Blueprint for a New Computing Infrastructure, San Francisco: Morgan Kaufmann, Elsevier.

Goble C, Corcho, O, et al. (2006): E-Science and the Semantic Web: A Symbiotic Relationship. In: Proceedings of Discovery Science, 9th International Conference, DS 2006, Barcelona, Spain, S. 1-12.

Gordon-Murnane L (2006). Social Bookmarking, Folksonomies, and Web 2.0 Tools. In: Searcher. The Magazine for Database Professionals, 14(6), S. 26-38.

Guédon J (2004): The Green and Gold Roads to Open Access: The Case for Mixing and Matching. In: Serials Review, 30 (4), S. 315-328.

Hammond T, Hannay T, et al. (2005): Social Bookmarking Tools (I): A General Review. In: D-Lib Magazine, 11 (4).

Handschuh S, Staab S (Hrsg.) (2003): Annotation for the Semantic Web, Amsterdam: IOS Press.

Harnad S, Brody T, Vallieres F, Carr L, Hitchcock S, Yves G, Charles O, Stamerjohanns H, Hilf E (2004): The Access/Impact Problem and the Green and Gold Roads to Open Access. In: Serials Review 30(4), S. 310-314.

- Hey T, Trefethen A (2005): Cyberinfrastructures for e-Science. In: Science 308, S. 817-821.
- Hepp M, Bachlehner D, Siorpaes K (2005): OntoWiki. Community-driven Ontology Engineering and Ontology Usage based on Wikis. In: Proceedings of the International Symposium on Wikis (WikySym), San Diego 2005.
- Koivunen MR, Swick RR (2004): Collaboration through Annotations in the Semantic Web. In: Handschuh, S.; Staab, S. (Hrsg.): Annotation for the Semantic Web. Frontiers in Artificial Intelligence and Applications Vol. 96, Amsterdam: IOS Press.
- Lee CP, Dourish P, Mark G. (2006): The Human Infrastructure of Cyberinfrastructure. In: Computer Supported Cooperative Work. Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, Banff Canada, New York: ACM Press, S. 483-492.
- Leser U, Naumann F (2007): Informationsintegration. Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen, Dpunkt.Verlag: Heidelberg.
- Macgregor G, McCulloch E (2006): Collaborative tagging as a knowledge organisation and resource discovery tool. In: Library Review, 55(5), 291-300.
- Neumann E, Prusak L (2007): Knowledge Networks in the Age of the Semantic Web. In: Briefings in Bioinformatics 2007 8 (3), S. 141-149.
- Paulsen I, Mainz D, Weller K, Mainz I, Kohl J, von Haeseler A (2007): Ontoverse. Collaborative Knowledge Management in the Life Science Network. In: Proceedings of the German eScience Conference 2007, Max Planck Digital Library, ID 316588.0.
- Peters I, Stock WG (2007): Folksonomy and Information Retrieval. In: Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology (Vol. 45) (CD-ROM).
- Resnick P, Varian HR (Hrsg.) (1997): Special Section on Recommender Systems. In: Communications of the ACM 40, 3 (März 1997), S. 56-89.
- Schuster M, Rappold D (2006): Social Semantic Software. Was soziale Dynamik im Semantic Web auslöst. In: Pellegrini T, Blumauer A (Hrsg): Semantic Web. Wege zur vernetzten Wissensgesellschaft, Berlin, Heidelberg: Springer, S. 189-199.
- Staab S, Studer R (Hrsg) (2004): Handbook on Ontologies, Berlin, Heidelberg, New York: Springer.

Stock WG (2007): Folksonomies and Science Communication. A Mash-up of Professional Science Databases and Web 2.0 Services, In: Information Services & Use 27(3).

Stuckenschmidt H, van Harmelen F (2005): Information Sharing on the Semantic Web, Berlin, Heidelberg: Springer.

Szugat M, Gewehr JE, Lochmann C (2006): Social Software. Blogs, Wikis & Co., Frankfurt (Main): Entwickler.press.

Timbrell G, Lambe P, Taule O (2007): Throwing Pebbles into a Dark Cave. A Study of Participation and Behaviour in the {act-km} Online Community. In: Journal of Information & Knowledge Management, 6 (2) 2007, S. 139-152.

Uren V, Cimiano P, et al. (2006): Semantic Annotation for Knowledge Management. Requirements and a Survey of the State of the Art. In: Journal of Web Semantics 2006 4(1).

Völkel M, Krötzsch M, Vrandečić D, Haller H, Studer R (2006): Semantic Wikipedia. In: Proceedings of the 15th International Conference on World Wide Web, Edinburgh, Scotland, May 2006 (WWW06), New York: ACM Press, S. 585-594.

Weller K, Mainz D, Mainz I, Paulsen I (2007): Wissenschaft 2.0? Social Software im Einsatz für die Wissenschaft. Erscheint in: 29. Online-Tagung der DGI, 59. Jahrestagung der DGI, Proceedings, Frankfurt (Main).

Wenger E, McDermott R, Snyder WM (2002): Cultivating Communities of Practice. A Guide to Managing Knowledge, Boston: HBS Press.

Wu H, Gordon M (2007): Collaborative Structuring. Organizing Document Repositories Effectively and Efficiently. In: Communications of the ACM 50, 7 (Juli 2007), S. 86-91.

Wissensvernetzung im Forschungsprozess

Ute Rusnak, Matthias Razum, Leni Helmes

Abstract

Der gesamte Prozess der wissenschaftlichen Kommunikation, Kollaboration und Publikation ist einer grundlegenden Änderung unterworfen. Wissenschaftler arbeiten zunehmend vernetzt und interdisziplinär an wechselnden Orten in ortsungebundenen virtuellen Organisationen und auf der Grundlage weltweit verteilter Ressourcen. Im Wissenschaftsbereich entstehen derzeit neue Plattformen für netzbasierte kooperative Forschung und Zusammenarbeit, die unter dem Begriff „E-Science“ zusammengefasst werden.

Das Informationsmanagement in Wissenschaft und Forschung war bisher darauf konzentriert, Ergebnisse aufzubereiten, diese über leistungsfähige Informationssysteme bereitzustellen und die Nutzung langfristig zu sichern. Für ein zukünftiges vernetztes Wissensmanagement, das sich mit Entstehung, Erwerb, Bewahrung, Bewertung, Ordnung, Verknüpfung, Identifikation, Abruf, Austausch und Nutzung von Wissen beschäftigt, ist jedoch die innovative und nachhaltige politische, organisatorische und technische Unterstützung des gesamten wissenschaftlichen Wertschöpfungsprozesses essentiell. Innovative Dienstleistungen und Infrastrukturen in diesem Umfeld sind notwendig, um Wissenschaft und Forschung bei der Informationsbeschaffung, Kommunikation, Kollaboration und Wissensgenerierung bis hin zur Publikation optimal und nachhaltig zu unterstützen.

1. Einführung

E-Science („Enhanced Science“) steht für die erweiterten Möglichkeiten der netzbasierten wissenschaftlichen Arbeit der Zukunft. Neue Informations- und Wissenstechnologien helfen, die etablierten organisatorischen Strukturen, das Informationsmanagement und die verfügbaren Werkzeuge weiter zu entwickeln und zu ergänzen und somit Forschungsprozesse zu vereinfachen, zu intensivieren und zu beschleunigen. E-Science beschreibt eine neue wissenschaftliche Arbeitsumgebung, die die Zusammenarbeit in verteilten Teams, den einfachen Austausch von Ressourcen und optimierte Publikations- und Kommunikationsprozesse umfasst. John Taylor (Director General of Research Council, UK) hat dies folgendermaßen zusammengefasst: „eScience is about global collaboration in key areas of science and the next generation of infrastructure that will enable it.“

Oftmals wird der Begriff E-Science im Zusammenhang mit rechenintensiven Anwendungen in verteilten Systemen verwendet, z.B. im Umfeld von Grid-Infrastrukturen. Gerade das Grid assoziiert man mit der Einrichtung großer Rechenzentren und leistungsfähiger Netze, um die heute anfallenden riesigen Mengen an Experimental- und Primärdaten verarbeiten zu können. Das Wachstum dieser Datenmengen geht ungebremsst weiter (Becla et al. 2006; Kovac 2003). Allerdings muss E-Science in

einem weiteren Sinn verstanden werden. Der direkte Zugang zu Primärdaten, Publikationen und anderen wissenschaftlichen Materialien erfordert nicht notwendigerweise den Einsatz großer Rechen- oder Speicherressourcen. Immer mehr Institutionen bauen Digitale Bibliotheken und Institutional Repositories auf, um ihren Forschern und Studenten den einfachen, schnellen und effektiven Zugang zu relevanten Informationen zu ermöglichen. Allerdings werfen diese vielfältigen neuen Informationsquellen auch neue Probleme auf: Persistenz, Zitierbarkeit, Auffindbarkeit, disziplinübergreifende Suchen, usw. Ein umfassendes Knowledge Management ist eine der Grundvoraussetzungen für die Umsetzung der „Vision“ von E-Science. Weiterhin sollten sich diese neuen Repositorien nicht auf die traditionell von Bibliotheken verwalteten Daten beschränken, sondern alle Artefakte speichern, die im Prozess der Wissensentstehung anfallen (Henry 2003). Nur so ist der gesamte Prozess nachvollziehbar dokumentiert und gleichzeitig ein Arbeiten in verteilten Teams (*collaboration*) überhaupt erst sinnvoll möglich.

2. Das eSciDoc-Projekt

Das im Rahmen der nationalen E-Science-Initiative geförderte Projekt eSciDoc (Dreyer et al. 2007) erarbeitet Lösungen für diese Herausforderungen. Dazu haben FIZ Karlsruhe und die Max-Planck-Gesellschaft (MPG) eine strategische Partnerschaft gebildet, um gemeinsam eine integrierte Informations-, Kommunikations- und Publikationsplattform für netzbasiertes wissenschaftliches Arbeiten zu entwickeln und in multidisziplinären Anwendungen zu erproben. Dabei wird ein produktiv einsetzbares System entstehen.

Das Projekt zielt nicht auf eine bestimmte Disziplin oder einen Anwendungsfall ab, sondern bietet eine generische Infrastruktur und Anwendungsumgebung für alle wissenschaftlichen Sektionen der MPG. Dazu erhebt das Projekt Anforderungen direkt in den Instituten, untersucht dort entstandene Anwendungen (GAVO 2007; ECHO 2007) und generalisiert sie teilweise, um von den vorhandenen Erfahrungen und Ergebnissen zu profitieren und mit den späteren Nutzern gemeinsam praxismgerechte Lösungen zu erarbeiten. Damit wird eSciDoc allerdings auch mit einer Vielzahl von Anforderungen konfrontiert, die eine hohe Flexibilität hinsichtlich Metadatenprofilen, Datentypen und Wiederverwendbarkeit von Diensten einfordern. Dies führt zur Verwendung von generischen Objektstrukturen (content models), die auch den Umgang mit bisher noch nicht vorgesehenen Datentypen und –strukturen erlauben. Durch diesen „agnostischen“ Ansatz – das System trifft keine Annahmen hinsichtlich späterer Anwendungsfälle – erreicht eSciDoc ein ungewöhnlich breites Einsatzspektrum. Gleichzeitig erlauben die einheitlichen Objektstrukturen die einfache Nachnutzung von Daten, auch über unterschiedliche Anwendungen hinweg. Entsprechend teilt sich das System in zwei große Bereiche auf: einen allgemeinen Infrastrukturtail („eSciDoc Framework“) und in die Gruppe der Anwendungen („eSciDoc Solutions“), die auf dieser Infrastruktur aufsetzen und deren Dienste nutzen. Das Framework stellt grundlegende Funktionen wie die Speicherung und Verwaltung der Objekte, die Authentifizierung und Autorisierung, die Validierung und

Konvertierung von Objekten und Metadaten, die Indexierung und Suche über die Objekte sowie die semantische Verknüpfung von Objekten in Form von Ontologien in Form unabhängiger Dienste bereit (im Sinne einer service-orientierten Architektur, SOA). Die Solutions führen diese Dienste zu disziplinspezifischen oder übergreifenden Anwendungen zusammen. Im Projekt konzentriert sich FIZ Karlsruhe auf die Konzeption und den Aufbau des Frameworks, während die MPG mit der größeren Nähe zu den späteren Nutzern – den Wissenschaftlern und Bibliothekaren in der MPG – für die darauf aufsetzenden Anwendungen verantwortlich ist. Durch die Kooperation mit anderen E-Science-Projekten und -Initiativen (z.B. D-Grid und kopal) kann die Infrastruktur um dort entwickelte und angebotene Dienste ergänzt werden (DGrid 2007; kopal 2007). Gleichzeitig können aber auch Dritte die Dienste von eSciDoc nachnutzen.

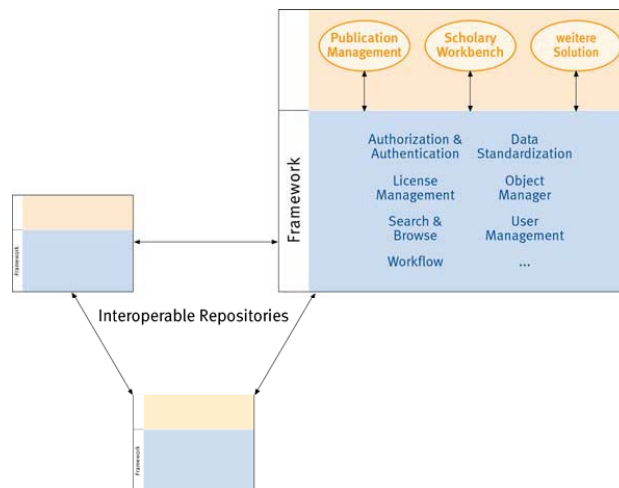


Abb. 1: Unterschiedliche Solutions können auf gemeinsam genutzte Dienste des Frameworks zugreifen; das Framework wiederum stellt nur einen Knoten in einem Netzwerk interoperabler Repositories dar.

3. Die eSciDoc Solutions

Das eSciDoc Projektteam erstellt einige Anwendungen bzw. Solutions auf Basis des Frameworks und unterstützt die Wissenschaftler in der MPG dabei, eigene Solutions zu realisieren. Der Fokus liegt momentan auf dem Publication Management, über das die Wissenschaftler und Bibliothekare der MPG Publikationsdaten aller Art (Artikel, Konferenzbeiträge, Poster, Reports, Bücher, Supplementary Material usw.) langfristig verfügbar machen können. Mit den Daten werden notwendige Metadaten für das Retrieval und die Langfristarchivierung gespeichert. Dazu gehört auch die Vergabe von persistenten Identifiern, die die verlässliche Zitierung und Referenzierung der Publikationen ermöglichen. Die Solution unterstützt semantische Beziehungen zwischen den Publikationen und damit verbundenen Personen (z.B. Autoren) und Organisationen. Die Suche in Metadaten und Volltexten ist über

entsprechende Masken und über eine standardisierte Schnittstelle (SRW/SRU) möglich. Das Einpflegen neuer Inhalte erfolgt über weitgehend konfigurierbare Arbeitsabläufe (*workflows*), die auch mehrstufige Qualitätskontrollen und Versionierung (Razum et al. 2007) beherrschen. Schließlich ermöglichen vielfältige Exportformate die Integration der Daten mit bestehenden Werkzeugen und Systemen. Die Mächtigkeit des Publication Managements bringt eine gewisse Komplexität mit sich. Durch rollenbezogene Sichten (z.B. für Wissenschaftler, Bibliothekare, Administratoren), die nur die jeweils benötigten Funktionen anbieten, bleibt diese Komplexität dem Anwender verborgen.

Weiterhin entstehen eine Reihe von Solutions, die sich mit dem wissenschaftlichen Arbeitsplatz der Zukunft beschäftigen, der so genannten Scholarly Workbench. Sie soll den Wissenschaftlern auf einfache Weise das Erstellen, Verarbeiten und Publizieren von digitalem Quellenmaterial ermöglichen. Dabei sollen auch geografisch verteilte Arbeitsgruppen gemeinsam an solchen digitalen Kollektionen arbeiten können. Durch die Einbindung lokaler und im Netz verfügbarer disziplinspezifischer Werkzeuge lassen sich die Daten analysieren, bewerten, bearbeiten, anreichern, verknüpfen und annotieren. Die Realisierung erfolgt in Schritten, ausgehend von relativ einfachen Szenarien (etwa dem Umgang mit Bildern). Mit jeder weiteren Solution kommen komplexere Funktionen hinzu. Die Solutions entstehen jeweils in enger Zusammenarbeit mit einem oder mehreren Instituten der MPG, die einen konkreten Bedarf haben. Deren Mitarbeiter arbeiten direkt oder indirekt an der Entwicklung mit und stellen so die Praxisnähe der entstehenden Solution sicher.

4. Das eSciDoc Framework

Übersicht

Die eSciDoc Solutions haben, wie auch viele andere E-Science-Anwendungen, trotz ihrer unterschiedlichen Einsatzzwecke und Anwendungsbereiche, eine Reihe von Gemeinsamkeiten:

- ein Verwaltungssystem für vielfältige digitale Objekte, einschließlich ihrer Metadaten (Object Manager)
- die Unterstützung einer erweiterbaren Menge von Objekttypen (Artikel, Bücher, Bilder, Datensätze, Lexika, Videos, Präsentationen, usw.)
- eine verteilte Authentifizierung (Standard Shibboleth)
- eine ausgefeilte Rechteverwaltung (Zugriffsrechte und Lizenzen)
- eine verlässliche und langfristige Identifizierung der digitalen Objekte über Persistent Identifier (PID)
- den Import und Export der digitalen Objekte in das oder aus dem Framework
- Such- und Ausgabemethoden für die digitalen Objekte.

Diese gemeinsam genutzten Funktionen bilden die Grundlage für den weniger sichtbaren Teil des eSciDoc Projekts: das eSciDoc Framework. Es besteht aus

einem Kern mit dem Object Manager und der Rechteverwaltung sowie weiteren Diensten, die typische Aufgabenstellungen im E-Science-Umfeld abdecken und damit auch außerhalb von eSciDoc in vielen Anwendungen sinnvoll einsetzbar sind. Das Framework bietet keine Dienste an, die Endnutzer direkt ansprechen, sondern stellt seine Funktionen über einfach und intuitiv zu bedienende Programmierschnittstellen Software-Entwicklern zur Verfügung.

Aus technischer Sicht besteht das Framework aus einer Reihe unabhängiger Dienste, die über eine service-orientierte Architektur (SOA) lose miteinander verbunden sind (Krafzig et al. 2004). Anwendungen können auch nur einzelne Dienste aus dem Angebot des Frameworks nutzen und/oder sie mit anderen, von Dritten bereitgestellten Diensten kombinieren. Durch die Verknüpfung solcher Dienste erstellen Software-Entwickler einfach und in kurzer Zeit neue leistungsfähige Anwendungen.

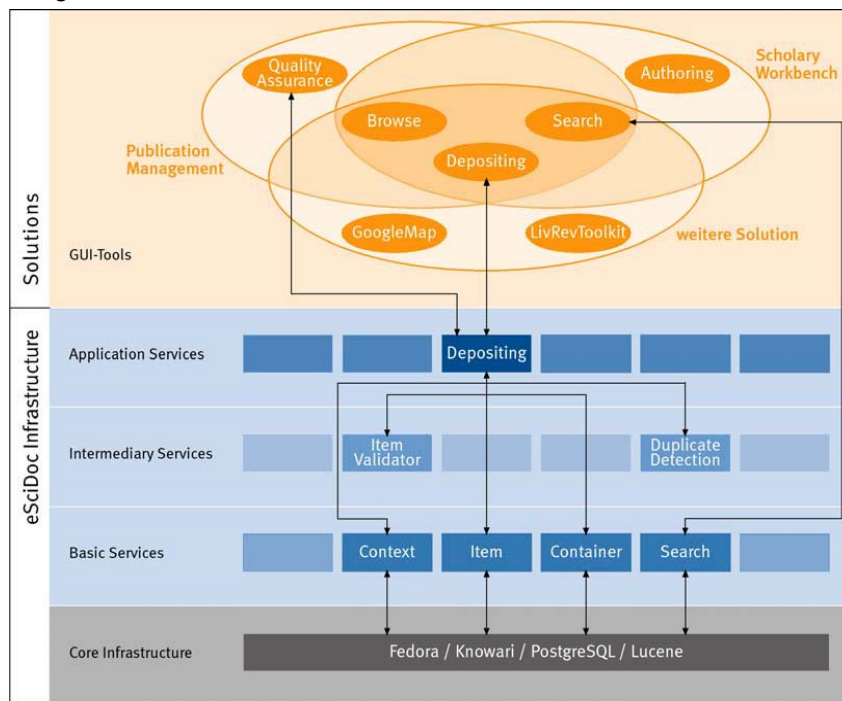


Abb. 2: Systemarchitektur mit symbolischer Darstellung der Schichten und Dienste, beispielhaft wird hier der Anwendungsfall „Depositing“ gezeigt.

Das Framework stellt alle notwendigen Funktionen zur Verfügung, um den gesamten Lebenszyklus eines Objekts (z.B. ein wissenschaftliches Dokument) sowie den zu seinem Verständnis notwendigen Kontext abzubilden. Dabei gibt es nahezu keine Beschränkungen für den Aufbau und die digitalen Bestandteile eines solchen Objekts. Aufgrund ausgefeilter Berechtigungsmechanismen können Autoren bereits den ersten Entwurf ihrer Arbeit ablegen und anfangs nur einem begrenzten Auditorium zugänglich machen, z.B. nur ihrer Arbeitsgruppe. Im Laufe der Zeit

können sie Erweiterungen und Korrekturen an ihren Objekten vornehmen und weitere Teile wie Bilder, Videos oder Grafiken hinzufügen. Auch die beschreibenden Informationen in den Metadaten können sie schrittweise ergänzen und vervollständigen. Dabei sind die Nutzer nicht auf eine spezielle eSciDoc Anwendung angewiesen. Über die vom Framework zur Verfügung gestellten REST- und SOAP-Schnittstellen (Fielding 2000) sind bestehende Werkzeuge einfach zu integrieren. Damit müssen sich die Anwender nicht an das Framework anpassen, sondern können ihre gewohnte Arbeitsumgebung beibehalten. Mit dem Fortschreiten der Arbeit können die Zugriffsrechte immer wieder bedarfsgerecht angepasst werden – von vertraulichen Daten bis hin zu Open Access Publikationen. Spezielle Rechte für die Mitarbeiter der eigenen Organisation sind möglich, ebenso die Berücksichtigung von Embargofristen kommerzieller Verleger.

Für die Darstellung oder Auswertung der Bestandteile eines Objekts können Web-Services eingebunden werden, die nicht notwendigerweise Bestandteile des Frameworks sind. So lassen sich auch sehr spezifische Funktionen, die z.B. nur für ein Institut oder eine Projektgruppe relevant sind, über das Netz einbinden. Dieses Vorgehen ermöglicht, sehr spezielle Objekte im Framework abzulegen, ohne dass hierzu eine Weiterentwicklung der Framework-Software notwendig wäre.

Das Framework versteht sich auch als „Enabling Technology“. Neben dem eSciDoc Team sollen auch andere ermutigt werden, auf Basis des Frameworks eigene Solutions zu entwickeln. Durch die Bereitstellung vieler immer wieder für solche E-Science-Anwendungen benötigter Funktionen vereinfacht sich die Programmierung erheblich. Damit versetzt das Framework z.B. Wissenschaftler in die Lage, sich auf die für ihre Arbeit wesentlichen Teile – nämlich die eigentliche „Geschäftslogik“ – zu konzentrieren. Das Framework garantiert die einheitliche und sichere Implementierung der zugrunde liegenden, immer wiederkehrenden Aufgaben. Der Aufbau der (fachspezifischen) Solution erfolgt damit auf einer verlässlichen und nachhaltigen Infrastruktur, die auch den Betrieb in einer dedizierten Produktionsumgebung erlaubt.

Die zentralen Komponenten des eSciDoc Frameworks

Object Manager

Elementare Bausteine von E-Science-Infrastrukturen sind Repositories. Im eSciDoc Framework heißt diese Komponente Object Manager und speichert die digitalen Objekte, legt bei Änderungen Versionen an und verknüpft die Objekte mit Verwaltungsdaten (z. B. beschreibenden, technischen und administrativen Metadaten, Konfigurationen, Änderungshistorie). Die Grundlage des Object Managers bildet das bewährte Open-Source-Projekt Fedora (Flexible Extensible Digital Object Repository Architecture) (Lagoze et al. 2006). Hierdurch kann diese kritische Komponente auf einer ausgereiften Software aufsetzen und von der sehr aktiven weltweiten Community profitieren.

Viele heute verfügbare E-Science-Anwendungen sind auf einen speziellen Einsatzzweck ausgelegt. Die Daten, oftmals aufwändig erstellt, stehen nur für diesen Zweck zur Verfügung: sie liegen in einem „Silo“. Wissenschaftliches Arbeiten jedoch ist ein sehr dynamischer Prozess, bei dem sich die Sicht auf die Daten und ihre Verwendungsmöglichkeiten schnell ändern können. Gleichzeitig erfordern E-Science-Anwendungen Objekte unterschiedlichster Typen – vom in sich abgeschlossenen wissenschaftlichen Artikel bis hin zu komplexen Objekten mit mehreren Bestandteilen wie Bildern, Experimentaldaten und Beschreibungen. Der Object Manager verwaltet deshalb die Daten in generischen Objektstrukturen, die sich nicht auf die zurzeit bekannten Anwendungsszenarien und Objekttypen festlegen. Durch die Trennung von Framework und Solutions stehen von Anfang an klar definierte Schnittstellen zur Verfügung, um auf die Daten zuzugreifen. Damit können die gespeicherten Objekte später auch in neuen, bisher noch nicht angedachten Solutions wieder verwendet werden. Der Object Manager dient so als Quelle für eine Vielzahl von Anwendungen.

Darüber hinaus ermöglicht das Framework die Verlinkung zwischen Objekten innerhalb wie außerhalb des eSciDoc Frameworks – hilfreich für Referenzen, Annotationen und Zitierungen. Auf diese Weise entsteht eine vernetzte Wissensbasis, die die Information in den Objekten selbst um eine zusätzliche Schicht erweitert, die die Objekte untereinander in Beziehung setzt und so Mehrwert schafft. Beziehungen der Objekte untereinander lassen sich durch den Standard RDF (Resource Description Framework) beschreiben (Manola 2004). Dabei beschreibt jeder RDF-Ausdruck eine Beziehung mit Quelle, Ziel und Bedeutung der Beziehung¹. Folgt diese Bedeutung gewissen Regeln, baut man damit eine Ontologie auf. Eine Suche innerhalb einer solchermaßen definierten Ontologie ist ebenfalls möglich – sehr ähnlich den Konzepten des „Semantic Web“.

Alle Objekte können in verschiedenen Repräsentationsformen (z. B. Dateiformaten) und mit mehreren Metadatensätzen unterschiedlicher Schemata angelegt werden. Erst durch eine solche Speicherung entfalten die Objekte ihren vollen Wert.

Die Objekte lassen sich in verschiedenen logischen Strukturen (z. B. Collections oder Bundles) organisieren. Auch hierfür stellt das Framework entsprechende Dienste zur Verfügung. Solche „logischen Strukturen“ – digitale Editionen, Sammlungen, komplexe Wissensobjekte u. a. – können auch als Ganzes publiziert und zitierfähig gemacht werden.

Authentifizierung und Rechteverwaltung

Ein wichtiger Aspekt bei allen E-Science-Anwendungen ist die Authentifizierung (wer ist der Benutzer) und Autorisierung (was darf der Benutzer). Voraussetzung für die Authentifizierung ist eine Benutzerverwaltung. Viele Anwendungen bringen ihre eigene Benutzerverwaltung mit, was dazu führt, daß die gleiche Person in diversen

¹ Genauer ausgedrückt umfasst ein RDF-Ausdruck ein Subjekt (Ressource), ein Prädikat (Eigenschaft) und ein Objekt (Wert). Die drei Elemente eines RDF-Statements bilden zusammen ein Triple.

Systemen als Benutzer verwaltet werden muss. eSciDoc geht hier einen anderen Weg und nutzt die bereits in den meisten Institutionen vorhandenen Benutzerverwaltungssysteme (*Identity Management Systems, IDM*). Soll eSciDoc institutsübergreifend eingesetzt werden, erlaubt eine verteilte Authentifizierung auch die Integration mit mehreren IDM-Systemen. Das hat neben verringertem Verwaltungsaufwand und aktuelleren Daten noch den Vorteil, dass Benutzernamen und Passworte nie die „Heimatorganisation“ des Benutzers verlassen. eSciDoc setzt dabei auf den im E-Science-Umfeld weit verbreiteten Standard Shibboleth (Shibboleth 2007), der z.B. auch im D-Grid häufig eingesetzt wird.

Gerade für ein System, das auch sensible Daten wie noch nicht publizierte Zwischenergebnisse verwaltet, ist eine zuverlässige und fein abgestufte Rechteverwaltung wichtig. Benutzer haben hier sehr detaillierte Rechte zum Erstellen, Bearbeiten und Lesen von digitalen Objekten oder auch nur von Teilen davon. Beispiel hierfür sei ein Artikel, auf dessen Metadaten frei zugegriffen werden kann, dessen Volltext aber aufgrund einer Vereinbarung mit dem Verlag noch gesperrt ist. Daher unterstützt das eSciDoc Framework eine breite Palette von Regeln: ausgehend von sehr generischen Regeln auf der Ebene des Object Managers oder von Benutzerrollen bis hin zur Regelung des Zugriffsrechtes auf einzelne Bestandteile eines Objekts. Das Framework übernimmt die komplette Zugriffsverwaltung und nimmt damit diese komplexe Aufgabe den Anwendungen ab. So garantiert das eSciDoc Framework jederzeit eine konsistente Rechteverwaltung und deren korrekte Anwendung für alle gespeicherten digitalen Objekte – unabhängig von den darüber liegenden Anwendungen. Die Entwickler der Anwendungen können sich also ganz auf die eigentliche Geschäftslogik konzentrieren und müssen sich keine Sorgen machen, dass sie durch Fehler die Integrität des Systems gefährden.

Metadatenverwaltung

Das eSciDoc Framework unterstützt neben einem internen noch weitere, auch selbst definierte Metadaten-Schemata. Unterschiedliche Objekttypen benötigen jeweils eigene Metadatenprofile, teilweise gibt es auch mehr als einen Metadatenatz pro Objekt. Die Metadatenverwaltung übernimmt hierbei mehrere Aufgaben: die Registrierung von Schemata, die Validierung von Metadatenätzen gemäß eines Schemas sowie die automatische Transformation von einem Schema in ein anderes. Dazu können Transformationsregeln hinterlegt werden.

Suche

Ein wesentlicher Bestandteil der meisten E-Science-Anwendungen ist die Suche nach digitalen Objekten anhand ihrer Eigenschaften. Typischerweise sind dies die Metadaten, bei Textobjekten auch der Volltext. Je nach Domäne und gespeicherten Objekten sind aber viele weitere Suchen denkbar, z.B. nach chemischen Substanzen oder Strukturen, nach ähnlichen Bildern oder nach charakteristischen Ton- bzw. Melodiefolgen.

Schon bei der Textsuche kommen weitere Aspekte hinzu: die Einbindung von Normdateien (*authority files*), Thesauri und Wörterbüchern, die sprachabhängige Analyse und Rückführung auf Normalformen und Lemmata, semantische Suchen oder der Umgang mit mehrsprachigen Texten. Für viele diese Probleme gibt es bereits Lösungen, die in diesem Zusammenhang nachgenutzt werden können. Spezielle Suchen können über unabhängige Dienste eingebunden werden und so die spezifischen fachlichen Probleme der Wissenschaftler lösen helfen. Solutions aus dem Bereich der Scholarly Workbench zeigen beispielhaft ausgefeilte Suchen in altsprachlichen Textbeständen und deren mehrstufige linguistische Analyse. Die dazu notwendigen Funktionalitäten sind als eigenständige Webservices umgesetzt, so dass sie in vielfältigen E-Science-Umgebungen einzusetzen sind.

Verwaltung von Organisationen und Personen

Digitale Objekte haben Beziehungen zu Personen, z.B. zu Autoren oder Herausgebern. Diese wiederum haben eine Zugehörigkeit zu einer Institution und ihren Organisationseinheiten wie Abteilungen oder Arbeitsgruppen. Das eSciDoc Framework stellt Dienste zur Verwaltung dieser Zusammenhänge bereit, um darüber z. B. Eigentums- und Urheberverhältnisse darzustellen und indirekt auch die Zugriffsrechte auf die Objekte zu steuern. Weiterhin ist über diese Dienste eine Zuordnung von Personen und Organisationseinheiten zu den in der Benutzerverwaltung bereitgestellten Daten möglich.

Interoperabilität

eSciDoc existiert nicht auf einer Insel. Viele Institutionen wie auch die MPG haben bereits Vorläufersysteme. Viele Wissenschaftler arbeiten mit Programmen zur Verwaltung von Referenzen wie Endnote oder BibTex. Über Standards wie OAI-PMH will man Metadaten aus anderen Repositories harvesten, also einmalig oder regelmäßig in das eigene System übernehmen. All dies erfordert flexible Möglichkeiten der Datenübernahme - Flexibilität nicht nur hinsichtlich der Formate, sondern auch hinsichtlich der Automatisierung solcher Prozesse. Dies umfasst die Beschreibung von Quellsystemen, die zu verwendenden Verfahren (Protokolle, Häufigkeit, notwendige Umwandlungsschritte) sowie eventuelle Qualitätssicherungsschritte.

Auch der Austausch mit anderen Systemen oder mit Kollegen sowie die Erstellung von z.B. Publikationslisten erfordern umfangreiche und flexible Exportfunktionen. Hier spielt die Umwandlung von Metadaten und Dateiformaten eine wichtige Rolle. Über entsprechende Mechanismen kann der Exportdienst auch externe Harvester nach OAI-PMH bedienen und speziell für Suchmaschinen wie Google optimierte Darstellungen ausliefern.

5. Nachnutzung der Projektergebnisse: eSciDoc Solutions und Framework

Beide eSciDoc Partner unterstützen den Bereich E-Science nachhaltig, auch über die Projektlaufzeit hinaus. FIZ Karlsruhe hat als langjähriger Servicepartner der Wissenschaft für Informationsdienstleistungen die Entwicklung von E-Science-Lösungen als eigenständiges Geschäftsfeld eingerichtet. Die Max Planck Digital Library (MPDL) bündelt als neue zentrale Serviceeinheit die Aktivitäten zur Versorgung der MPG-Institute mit wissenschaftlicher Information und zur Unterstützung von webbasierter wissenschaftlicher Kommunikation.

Das Projekt eSciDoc wird alle Open-Source-Softwarekomponenten des Gesamtsystems (eSciDoc Solutions und Framework) veröffentlichen, um sie interessierten Informationseinrichtungen in Wissenschaft und Forschung zur Verfügung zu stellen und die Grundlage dafür zu schaffen, die Software gemeinsam mit anderen Beteiligten weiter zu entwickeln. Hierdurch ergibt sich die Chance, die im Rahmen des Projekts eSciDoc entwickelten Softwarekomponenten als Infrastruktur für E-Science-Anwendungen in einem nationalen Rahmen zu etablieren.

FIZ Karlsruhe wird die Zusammenarbeit mit weiteren Forschungseinrichtungen und Universitäten aktiv vorantreiben, indem es die neuen Anwendungen und darauf aufbauende Mehrwertdienste durch geeignete Kommunikationsmaßnahmen einem breiten Nutzerkreis vorstellt. Der Aufbau einer offenen Kommunikationsplattform unterstützt die Bildung einer aktiven E-Science-Community, die ähnliche Fragestellungen aus der täglichen Arbeit mit wissenschaftlichen Informationen über Organisations- und Disziplingrenzen hinweg diskutiert und somit weitere Entwicklungen von E-Science-Anwendungen vorantreibt.

Literatur

(Becla et al. 2006) Becla, J., Hanushevsky, A., Nikolaev, S. et al. 2006. Designing a Multi-petabyte Database for LSST. SLAC Publication 12292.

<http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-pub-12292.pdf>

(DGrid 2007) D-Grid Initiative, Homepage. <http://www.d-grid.de/>

(Dreyer et al. 2007) Dreyer, M., Bulatovic, N., Tschida, U., Razum, M. 2007. eSciDoc – a Scholarly Information and Communication Platform for the Max Planck Society.

Konferenzbeitrag. German e-Science Conference, Baden-Baden.

<http://edoc.mpg.de/get.epl?fid=36036&did=315471&ver=0>

(Fielding 2000) Fielding, R.T. 2000. Architectural Styles and the Design of Network-based Software Architectures. Dissertation. University of California, Irvine.

http://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf

(GAVO 2007) German Astrophysical Virtual Observatory (GAVO) Projekt, Homepage. <http://www.g-vo.org/>

(ECHO 2007) European Cultural Heritage Online (ECHO), Homepage
(momentan betrieben vom Max-Planck-Institut für Wissenschaftsgeschichte).
<http://echo.mpiwg-berlin.mpg.de/>

(Henry 2003) Henry, G. 2003. On-line publishing in the 21-st Century. Challenges and Opportunities. D-Lib Magazine, Volume 9, Issue 10.
<http://dx.doi.org/10.1045/october2003-henry>

(kopal 2007) kopal: Ein Service für die Langzeitarchivierung digitaler Informationen, Homepage. <http://kopal.langzeitarchivierung.de/>

(Kovac 2003) Kovac, C. 2003. Computing in the Age of the Genome. The Computer Journal. Volume 46, Issue 6, 593-597.
<http://dx.doi.org/10.1093/comjnl/46.6.593>

(Krafzig et al. 2004) Krafzig, D., Banke K., Slama D. 2004. Enterprise SOA: Service-Oriented Architecture Best Practices. Prentice Hall.
<http://proquest.safaribooksonline.com/0131465759>

(Lagoze et al. 2006) Lagoze, C., Payette, S., Shin, E., Wilper, C. 2006. Fedora: an architecture for complex objects and their relationships. International Journal on Digital Libraries. Volume 6, Issue 2. Springer Berlin / Heidelberg.
<http://dx.doi.org/10.1007/s00799-005-0130-3>

(Manola et al. 2004) Manola, F., Miller, E. 2004. RDF Primer. W3C Recommendation.
<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>

(Shibboleth 2007) Shibboleth - Web Single Sign-On and Federating Software, Homepage. <http://shibboleth.internet2.edu/>

(Razum et al. 2007) Razum, M., Schwichtenberg, F., Fridman, R. 2007. Versioning of Digital Objects in a Fedora-based Repository. Konferenzbeitrag. German e-Science Conference, Baden-Baden.
<<http://edoc.mpg.de/get.epl?fid=36061&did=315520&ver=0>>

Wissenschaftskommunikation in Lehre und Forschung

Eine Ontologie der Wissenschaftsdisziplinen. Entwicklung eines Instrumentariums für die Wissenskommunikation

Elena Semenova, Martin Stricker

Abstract

Interdisziplinarität als Kennzeichen des modernen Wissenschaftslebens setzt in Forschung und Lehre eine effiziente Wissenschaftskommunikation voraus, bei der sich die Partner über eine gemeinsame Sprache verständigen können. Klassifikationen und Thesauri übernehmen dabei eine wichtige Rolle. Zu beobachten ist, dass vorhandene Instrumente in ihrem Gefüge zu inflexibel sind, um die komplex ineinander verwobenen Felder der Wissenschaft in ihrer dynamischen Entwicklung adäquat abzubilden, zur (Selbst-)Verständigung über das Wesen und Struktur der Wissenschaftslandschaft sowie zum erfolgreichen Wissensaustausch beizutragen. Ontologien erschließen neue Wege zur Lösung dieser Aufgaben. In einigen Einzelwissenschaften und Teilgebieten ist diesbezüglich eine rege Tätigkeit zu beobachten, es fehlt allerdings noch ein fachübergreifendes Instrumentarium.

Im Vortrag wird das von der DFG geförderte Projekt "Entwicklung einer Ontologie der Wissenschaftsdisziplinen" vorgestellt. Es gilt, die oben beschriebene Lücke zu schließen und eine umfassende Ontologie für Erschließung, Recherche und Datenaustausch bei der Wissenschaftskommunikation zu erstellen. Diese Ontologie soll dazu beitragen, eine effiziente Wissenskommunikation, besonders bei interdisziplinären Projekten, zu unterstützen, verfügbare Ressourcen auffindbar zu machen und mögliche Knotenstellen künftiger Kooperationen zu verdeutlichen. Ausgehend von der Kritik an vorhandenen Instrumenten wird derzeit ein Begriffsmodell für die Beschreibung von Wissenschaftsdisziplinen, ihrer zentralen Facetten sowie ihrer interdisziplinären Beziehungen untereinander entwickelt. Das Modell, inspiriert vom Topic Maps Paradigma, basiert auf einer überschaubaren Menge zentraler Konzepte sowie prinzipiell inverser Beziehungen. Eine entsprechende Ontologie wird in unterschiedlichen (technischen) Beschreibungsformaten formuliert werden können. Dies bildet den Grundstein für den Fokus des Projekts, flexible, verteilte, benutzer- wie pflegefreundliche technische Umsetzungen zu entwickeln und mit Kooperationspartnern zu implementieren.

Ontology for Academic Disciplines. The development of a tool for knowledge communication.

Interdisciplinarity, as a mark of the modern science life, presumes an efficient knowledge communication, in which partners can interact sharing a common language. Classifications and Thesauri play hereto an important role. It is noticeable, that available instruments are too inflexible in their diversity to display the complex interwoven science fields appropriately in their dynamic development to contribute to the understanding about the nature and the structure of the knowledge landscape, as

well as to a successful knowledge exchange. Ontologies provide new ways for the solution of this task. Some action can be observed in a few solitary science disciplines and sub-areas as to that, however most lack interdisciplinary tools to tackle the task.

The presentation introduces the project *Developing an Ontology for Academic Disciplines* funded by the German Research Foundation (Deutsche Forschungsgemeinschaft). The aim is to close the above-mentioned gap and to create a concise ontology for set-up, research and data exchange in science communication. This ontology should take over to sustain a useful science communication, especially in regard to interdisciplinary projects, to retrieve available resources and to highlight possible knot spots for future cooperation. Due to criticism of existing tools, a term framework for the description of science disciplines, their central layers as well as their interdisciplinary relationships, is currently in development. The framework, inspired by Topic Maps Paradigm, is based on a controlled amount of central concepts and regularly inverse relationships. An according ontology can be formulated in different (technical) description formats. This lays the groundwork for the focus of the project, to develop flexible, distributed, user-friendly technical implementation for optimal partner cooperation.

Die drastische Veränderung der Wissenschaftslandschaft ist zu einem markanten Zeichen unserer Zeit geworden. Die Wissenschaftsentwicklung der letzten Jahrzehnte zeichnet sich neben der Etablierung neuer wissenschaftlicher Disziplinen sowie ständiger Erneuerung und Erweiterung der Inhalte von bereits bestehenden Einzelwissenschaften, insbesondere durch eine ausdrückliche Tendenz zur Interdisziplinarität, aus. Die Logik dieser Entwicklung macht Wissenschaftskommunikation zum unentbehrlichen Teil des wissenschaftlichen Lebens. Dementsprechend wächst die Rolle der Kommunikationsinstrumente. Eine gelungene Kommunikation kann nur dann erfolgen, wenn alle Kommunikationspartner eine gemeinsame Sprache beherrschen oder zumindest mit ihren Grundstrukturen vertraut sind. Die Begriffssysteme, die auch als Dokumentationssprachen bezeichnet werden, wie Klassifikationen, Thesauri und zunehmend Ontologien, erhalten dabei neue Funktionen und gewinnen dadurch an Bedeutung. Längst werden diese Begriffssysteme nicht nur für die Erschließung und Recherche der Wissensbestände verwendet. Im Bereich des Datenaustausches, einem wichtigen Teil der Wissenschaftskommunikation, ist die Anwendung systematisch erfasster Begriffssysteme heute unabdingbar geworden. Mit zunehmender Rolle, die Begriffssysteme für den Wissensaustausch spielen, vermehren sich auch die Anforderungen an sie. Die modernen Instrumentarien sollten möglichst adäquat die gewählte Domäne in ihrer Komplexität abbilden, zeitnah, flexibel und gleichzeitig eindeutig, konsistent, skalierbar und effizient in der Benutzung sein.

Die wachsende Zahl der interdisziplinären Vorhaben verlangt besonders nach einer solchen Struktur. Bei der Zusammenarbeit von Wissenschaftlern aus unterschiedlichen Bereichen erschwert die semantische Pluralität die Arbeit enorm. Die interdisziplinäre Kommunikation stößt auf den verschiedensten Ebenen auf zahlreiche Hindernisse, die größtenteils mit Hilfe eines normierten Begriffssystems überwunden werden können. Die Notwendigkeit eines für alle Beteiligten verständlichen und klar strukturierten Vokabulars, welches die Fachsprache kanalisiert, das Wissensaustausch – vor allem im Rahmen technischer Informationssysteme - ermöglicht, die Kommunikation auf allen Ebenen erleichtert und in einer transparenten und nachvollziehbaren Struktur repräsentiert wird, ist kaum zu überschätzen. Die heute vorhandenen Strukturen entsprechen jedoch diesen Anforderungen nicht: Keines der verfügbaren Vokabularien kann die zunehmende Komplexität der wissenschaftlichen Erkenntnisse und Strukturen zufriedenstellend repräsentieren.

Traditionell wird das System der Wissenschaftsdisziplinen in Form von Klassifikationen dargestellt, d.h. dass als Ordnungssystem eine Baumstruktur zugrunde gelegt wird. Diese Jahrtausende alte Tradition ist jedoch seit geraumer Zeit an ihre Grenzen gestoßen. In der Wissenschaftswelt, wo fast jede Disziplin mit einer anderen verknüpft ist, lassen sich Wissenschaften nur bedingt zu großen Gruppen zusammenfassen und die Hierarchien auf den zweiten, dritten und weiteren Ebenen können oft nur unter zahlreichen Vorbehalten gebildet werden.

Vor einigen Jahren war eine Tendenz zu beobachten, die Nachteile der starren Baumform mit Hilfe der Bildung von polyhierarchischen Vokabularien zu überwinden.

Die Entwicklung von Ontologien als Begriffssysteme war die logische Folge dieser Tendenz, ging aber darüber hinaus und führte zu einem gewissen Paradigmenwechsel in diesem Bereich: Seit der Entwicklung von Ontologien verdrängt die Vorstellung einer Netzstruktur die alten Muster. Die Ontologien entwickeln sich rasch. In der Wirtschaft längst Norm geworden, gewinnen sie seit einiger Zeit sichere Positionen im Wissenschaftsalltag. In vielen Teilbereichen der Wissenschaft werden fachspezifische Ontologien entwickelt. Wenn jedoch die Struktur der Wissenschaftslandschaft selbst repräsentiert werden soll, wird nach wie vor zumeist auf die konventionelle Systematik in Form der Klassifikation zurückgegriffen. Die Gründe dafür sind ohne Zweifel in der Geschichte zu suchen. Durch die historische Entwicklung hat es sich so ergeben, dass die Bibliotheken im Bereich der informationswissenschaftlichen Technologien eine Vorreiterrolle übernommen haben. Dementsprechend wurden alle bekannte Systematiken in erster Linie für die Erschließung und Recherche der Bibliotheksbestände erstellt. Als Melvil Dewey in den 70er Jahren des 19. Jahrhunderts seine Dezimalklassifikation entwickelte, entsprach diese einerseits den Anforderungen, eine ortsgebundene, klare und einheitliche Ordnung in die Bibliothek zu bringen, andererseits ist in ihr die zeittypische Ordnung der Wissenschaftsdisziplinen ablesbar.

Durch die Wandlung der Inhalte einzelner Wissenschaftsdisziplinen sowie durch die drastische Veränderung der gesamten Wissenschaftslandschaft erscheinen heute - mehr als hundert Jahre nach Deweys Entwicklung – trotz beständiger Überarbeitung einige Positionen der DDC hinsichtlich der Wissenschaft nicht zeitgemäß. Außerdem ist das Ziel der DDC nicht, explizit wissenschaftliche Disziplinen zu erfassen, sondern eher Wissens- und Wirkungsbereiche. Schon in der ersten Übersichtsebene der letzten Fassung der DDC¹ wird dieses deutlich:

- 000 Informatik, Informationswissenschaft, allgemeine Werke
- 100 Philosophie und Psychologie
- 200 Religion
- 300 Sozialwissenschaften
- 400 Sprache
- 500 Naturwissenschaften und Mathematik
- 600 Technik, Medizin, angewandte Wissenschaften
- 700 Künste und Unterhaltung
- 800 Literatur

Wissenschaftliche Disziplinen (*Informationswissenschaft, Philosophie und Psychologie, Religion* u.a.) befinden sich auf einer Ebene mit Phänomenen (*Sprache*) und Erzeugnissen der menschlichen Kreativität (*Künste und Unterhaltung, Literatur*). Von der zweiten Ebene an wird diese Art der uneinheitlichen Klassenbildung in Hinsicht der Wissenschaftsdisziplinen noch deutlicher: In der Klasse *Sprache* steht *Linguistik* (eine Disziplin) neben der Auflistung verschiedener Sprachen (einem Phänomen) wie z.B. *Englisch, Altenglisch* oder *Deutsch*, germanische Sprachen allgemein etc. Die hierarchische Gliederung in der DDC wird manchmal sogar regelrecht verunstaltet: in der Klasse *Naturwissenschaften* stehen *Biowissenschaften/Biologie* auf der gleichen Ebene wie *Pflanzen (Botanik)* und *Tiere (Zoologie)*. Ein solches Ordnungssystem orientiert sich an der Besonderheit des Bibliothekssystems und ist den Erfordernissen der Erschließung von Bibliotheksbeständen angepasst. Dieses ist jedoch nur bedingt auf andere Bereiche übertragbar. Nach dem gleichen Prinzip wurden die Internationale Universale Dezimalklassifikation (UDK) und die Regensburger Verbundklassifikation (RVK)² aufgebaut. Obwohl beide Klassifikationen in der Auswahl der Bezeichnungen viel konsequenter vorgehen als die DDC, weisen aus Sicht der wissenschaftlichen Fachgebiete diese Klassifikationssysteme einen Mangel an Homogenität der Beschreibung, an ausgewogenen Strukturen und an präziser Abbildung der Relationen zwischen den Disziplinen auf³. Mit anderen Worten: Diese

¹ <http://www.ddc-deutsch.de/produkte/uebersichten/summaries1.htm> [02.08.2007]

² http://www.bibliothek.uni-regensburg.de/rvko_neu/ [02.08.2007]

³ Vgl. Konrad Umlauf, Einführung in die bibliothekarische Klassifikationstheorie und –praxis: <http://www.ib.hu-berlin.de/~kumlau/handreichungen/h67/> (02.08.07)

Klassifikationssysteme sind streng zweckbezogen, vertreten die Sichtweise einer bestimmten Zielgruppe und erfüllen auf keine Weise den Anspruch auf Universalität. Die primäre Position, die Bibliotheken im Informationsbereich innehatten, wird zunehmend durch die Verbreitung des Internet in Frage gestellt. Die Verbreitung der neuen Technologien und das ständig anwachsende Wissensgut führt zu einem neuen Selbstverständnis der Gesellschaft, die heute als Wissensgesellschaft bezeichnet wird. Wissensportale, die der Bündelung der Informations- und Wissensströme dienen, besetzen heute in der Wissenschaftskommunikation eine wichtige Stellung. Virtuelle Bibliotheken, Subject Trees bzw. Gateways und Suchmaschinen für Wissenschaftsserver bieten neben den Suchmöglichkeiten auch eine Browsing-Funktion, die hauptsächlich auf einer Wissenschaftssystematik basiert. Die überwiegende Zahl dieser Portale sind Fachportale, die jeweils nur ein Fachgebiet präsentieren und dementsprechend eine fachspezifische Systematik verwenden.⁴ Viele dieser Portale entwickeln eine eigene Systematik. Bei aller Differenz haben sie ein wesentliches Merkmal gemeinsam: Die Bezeichnungen der Wissenschaftsdisziplinen werden hier in erster Linie für den thematischen Einstieg in die Recherche genutzt. Die Systematiken gehen also nicht in die Tiefe, sondern sind relativ grob gegliedert und bilden die Fachgebiete nicht in vollem Maße ab. Die Fächerauswahl von ‚vascoda‘ z. B. bietet nur 30 Fächer an, die in 4 großen Klassen zusammengefasst sind. Obwohl die Suchfunktion von ‚vascoda‘ Informationen für eine erheblich größere Anzahl der Fächer recherchieren lässt, mangelt es an Navigationsmöglichkeiten. Aus den ausgegebenen Treffern kann man nur bedingt die nötigen Zusammenhänge herauslesen. Als Folge dessen kann diese Systematik nur beschränkt für eine erfolgreiche wissenschaftliche Kommunikation verwendet werden. Offensichtlich wurde bei der Erstellung der Systematik ein Bottom-Up Prinzip gewählt. Die Vorteile dieses Prinzips bestehen darin, dass das Erschließungs- und Rechercheinstrument in Umfang und fachspezifischem Standpunkt an die zu erschließende Datensammlung angepasst wird. Über die Datensammlung hinaus verliert ein solches Instrument seine Relevanz.

Gleiches gilt für andere Portale. Die vorgenommene Systematisierung erfüllt durchaus das eigentliche Ziel des einzelnen Portals, ist dennoch für andere Datensammlungen nicht zu verwenden. Als Folge dessen stellen sie keineswegs ein komplettes Bild der Wissenschaftsdisziplinen dar und können nur begrenzt in anderen Kontexten genutzt werden.

Für die Erstellung einer neuen Struktur sprechen viele Gründe, die im folgenden zu skizzieren sind:

⁴ Vgl. Clio-online, Ein Fachportal für die deutschsprachigen historischen Wissenschaften; EVIFA, die Virtuelle Fachbibliothek Ethnologie; Anglistik Guide: Virtual Library of Anglo-American Culture - Literature at SUB Göttingen (GoeAng); DAINet - German Agricultural information Network, Germany (DAINet); Geo-Guide: Virtual Library of Earth Sciences, Geography, Thematic Maps, and Mining at SUB Göttingen (GoeGeo); History Guide: Virtual Library of Anglo-American Culture - History at SUB Göttingen (GoeHist); MathGuide: Part of the Virtual Library of Mathematics at SUB Göttingen (GoeMath) u.a.

1. Als erste sind quantitative Gründe zu nennen: Es gibt wesentlich mehr Fachbereiche, als in den oben genannten Systematiken aufgelistet sind. Allein nach Angaben des Statistischen Bundesamtes⁵ werden an deutschen Hochschulen 619 Fächergruppen, Lehr- und Forschungsbereiche sowie Fachgebiete gezählt, wobei einige Disziplinen in dieser Aufstellung nicht ausdifferenziert werden. So wird z.B. die Sprachwissenschaft nur im Hinblick auf verschiedene Sprachen wie *Anglistik*, *Niederlandistik*, *Skandinavistik*, *Slavistik* präsentiert und nicht in Teilgebiete wie *Phonetik/Phonologie*, *Morphologie*, *Syntax/Grammatik*, *Semantik/Lexikographie* gegliedert. Unter Berücksichtigung der fehlenden Gliederung könnte die Anzahl der Fachbereiche etwa auf eintausend geschätzt werden. Um die Wissenschaftsdisziplinen adäquat darstellen zu können, ist die Entwicklung eines möglichst vollständigen Vokabulars unabdingbar.

2. Die qualitativen Gründe, die für die Entwicklung einer neuen Beschreibungssystematik sprechen, lassen sich in zwei Bereiche untergliedern: Einerseits gilt es Probleme zu bearbeiten, die für jedes kontrollierte Vokabular typisch sind, andererseits gilt es Ansätze zu entwickeln, die den spezifischen Anforderungen bei der Beschreibung von Wissenschaftsdisziplinen genügen:

- Synonymie, Polysemie und Homonymie. Die für jedes kontrollierte Vokabular typischen Probleme sind auch für den Bereich der wissenschaftlichen Disziplinen relevant. Das heißt: Alle Bezeichnungen der Fachgebiete müssen berücksichtigt werden. So sollte im Vokabular nicht nur der Begriff *Phylogenie*, sondern auch *Phylogenese*, neben *Stammesentwicklung* auch *Stammesgeschichte* enthalten sein und damit das Synonymproblem bewältigt werden.
- Das Polysemieproblem wird in vielen vorhandenen Systematiken mit Hilfe von polyhierarchischen Strukturen gelöst, was durchaus berechtigt wäre, wenn alle relevanten Bedeutungen vollständig abgebildet würden. Allerdings findet man in keiner der Systematiken eine komplette Darstellung beispielsweise des klassischen Polysems *Morphologie*. Die *Morphologie* ist ein Teilgebiet, das nicht nur in der *Sprachwissenschaft* und der *Biologie* eine Rolle spielt, sondern auch in der *Geomorphologie* oder in der *Stadt morphologie*.
- Ein grundlegendes Problem der Präsentation von Wissenschaftsdisziplinen ist die Vielfältigkeit der Forschungsaspekte. Die Anthropologie beispielsweise unterscheidet *Philosophische Anthropologie*, *Theologische Anthropologie*, *Biologische Anthropologie*, *Forensische Anthropologie*, *Historische Anthropologie*, *Kultur- und Sozialanthropologie*, *Kybernetische Anthropologie* und *Visuelle Anthropologie*. Der Forschungsgegenstand ist jedes mal der Gleiche - der Mensch - aber je nach Forschungsbereich wird er unter verschiedenen Aspekten betrachtet. Bis jetzt gibt es keine Systematik, die diese Relationen anschaulich und umfassend abbildet.

⁵<http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Klassifikationen/BildungKultur/StudentenPruefungsstatistik.property=file.pdf> (02.08.07)

- Manche Grundwissenschaften sind Bestandteil unterschiedlicher Disziplinen. Ein Beispiel ist die *Papyrologie*, die sowohl für die *Altphilologie* als auch für die Archäologie relevant ist. Auch diese Relationen müssen zum Ausdruck gebracht werden.
- Viele moderne Fachgebiete haben sich aus unterschiedlichen Disziplinen heraus entwickelt. So ist etwa die Computerlinguistik ein Teilbereich der *Künstlichen Intelligenz* und gleichzeitig Schnittstelle zwischen *Sprachwissenschaft*, *Kognitionswissenschaft* und *Informatik*. Solche Schnittstellen müssen genau dokumentiert werden.
- Es gibt kaum Wissenschaftsdisziplinen, die keine Kooperationen mit anderen Fachgebieten eingehen. In der *Telekommunikation* beispielsweise setzen sich die Forschungsteams aus den Bereichen *Ingenieurwissenschaften*, *Nachrichten- und Elektrotechnik*, *Mathematik* und *Physik* sowie *Informatik* zusammen. Heutzutage wird zwar oft über die Vernetzung der Wissenschaft gesprochen, es fehlt jedoch eine entsprechende Abbildung solcher Netze. Mit dem traditionellen Instrumentarium der Hierarchie- und Äquivalenzbeziehungen lässt sich dieses Problem nicht lösen.
- Auch die Beziehungen zwischen Studiengängen und Wissenschaftsdisziplinen sind zu berücksichtigen. Der Studiengang *Religionswissenschaft* z.B. beinhaltet u.a. auch das Studium der *Philosophie*, *Soziologie* und *Psychologie*. Eine Systematik der universitären Disziplinen hat auch solche Beziehungen repräsentativ darzustellen.

Diese Aufgaben können nur mit Hilfe eines Begriffssystems bewältigt werden, das die Möglichkeiten der traditionellen Erschließungs- und Rechercheinstrumente – Klassifikationen und Thesauri – erweitert und Relationen zwischen Konzepten, sowie pragmatische Aspekte wie unterschiedliche Perspektive in die Struktur mit einbezieht. Ein solches Instrumentarium sollte nicht nur möglichst unmissverständlich, sondern auch anschaulich und intuitiv nachvollziehbar sein, um größere Benutzerkreise mit der Struktur der Domäne vertraut zu machen und mögliche Schnittstellen für potenzielle Kooperationspartner zu verdeutlichen. Die aktuelle Entwicklung im Bereich Ontologieaufbau bildet das Fundament für die Erarbeitung eines solchen Instrumentariums. Die Ontologien erlauben, unterschiedlichste Formen von Beziehungen abzubilden und die Wissenschaftslandschaft, die sich selbst als eine vernetzte Welt versteht, in einer geeigneten Form zu repräsentieren.

Die Erarbeitung der Ontologie der wissenschaftlichen Disziplinen wird, wie bei solchen Projekten üblich, in zwei Phasen ablaufen: Modellerstellung mit Bildung von Klassen für Konzepte, Relationen, Namen und Rollen sowie anschließend die Erstellung der Ontologie selbst. Zugleich erfolgen parallel die Entwicklung spezifischer Software und Software-Schnittstellen für die Implementierung.

Unter einer „Ontologie“ verstehen wir für unser Projekt ein verzweigtes, verflochtenes Begriffssystem, das die Elemente eines kontrollierten Vokabulars in sich einschließt

und das auf einem standardisierten Referenzmodell für einen ausgewählten Bereich⁶ (in diesem Fall für die wissenschaftlichen Disziplinen) beruht. Dabei sollen die Vorteile der Potentiale der maschinellen Begriffsermittlung (*Reasoning*) genutzt werden, es ist jedoch davon auszugehen, dass der Anteil intellektueller Arbeit sehr hoch sein wird.

Für die Bestimmung der Anwendungsdomäne wird ein Verständnis des Begriffs „Wissenschaft“ als differenzierter und institutionalisierter Teilbereich der Gesellschaft zugrunde gelegt. Zusätzlich beschränkt sich derzeit das Projekt auf Forschung und Lehre in Deutschland (beim Referenzmodell soll auf eine weitgehende internationale Anwendbarkeit geachtet werden). Für eine kohärente Definition des Gegenstandsbereiches sollen Wissenschaftsdisziplinen von Wissensfeldern, Wissensgebieten und Wissensbereichen deutlich abgegrenzt werden. Die Anerkennung eines Wissensgebietes als wissenschaftliche Disziplin ist mitunter sehr umstritten. Da unser Projekt keinen normativ-setzenden, sondern einen deskriptiven Charakter hat, ist hier eine praktische Lösung gefragt, die durch einen Rückgriff auf die Verfahren gesellschaftlicher Legitimierung (zum Beispiel Institutionalisierung an Universitäten) Konsens zu erzielen sucht.

Die Festlegung der Vorgehensweise für die Erstellung der Ontologie wurde aus diesen zwei Grundverständnissen abgeleitet. Für unser Begriffssystem bevorzugen wir als Arbeitsweise einen Top-Down Ansatz. Dieses Verfahren hat eine Reihe von spezifischen Merkmalen, die bei der Erstellung dieser Art von Ontologien von Vorteil sind: Unter anderem ermöglicht die Unabhängigkeit von einem festgelegten Datenbestand eine breite Perspektive; auf diese Weise können unterschiedliche Auffassungen nebeneinander zum Ausdruck gebracht werden.

Die projektierte Ontologie wird mit dem standardisierten Datenmodell des Semantic Web (OWL/RDF) kompatibel sein bzw. auf einem Subset davon beruhen. Eine Kompatibilität mit weiteren Ontologien oder entsprechenden Anwendungen ist somit gewährleistet. In der derzeitigen Anfangsphase der Modellierung, die, da von standardisierten Verfahren bei der Ontologieerstellung noch nicht die Rede sein kann, sehr stark experimentell und iterativ geprägt ist, kommt der ISO-Standard Topic Maps zum Einsatz. Dessen Paradigma sowie entsprechende Softwareinstrumente ermöglichen aufgrund des – im Vergleich zu OWL/RDF – Reichtums an konkreten Konzepten einen flexiblen Entwicklungsprozess. Beabsichtigt ist, Referenzmodell und Ontologie in beiden Standards zu publizieren. Zu den weiteren Projektzielen zählt die Entwicklung und Bereitstellung einer Web-basierten Anwendung, die hypertextualisiert und grafisch-visuell zur Recherche der Ontologie sowie der im Laufe des Projektes erarbeiteten damit verknüpften Informationsquellen genutzt werden kann. Ein dritter Weg, auf die Ontologie

⁶ Vergl. z.B. Wolfgang Hesse, Barbara Krzensk: Ontologien in der Softwaretechnik <http://www.mathematik.uni-marburg.de/~krzensk/mod04/WorkshopBand.pdf> (31.07.07).
Definition der Ontologie von Wolfgang Hesse in Glossar der Gesellschaft für Informatik;
Diskussion: Katrin Weller : „Kooperativer Ontologieaufbau“ http://www.alt.phil-fak.uni-duesseldorf.de/inforwiss/admin/public_dateien/files/35/1166536740katrinwell.pdf (31.07.07)

zuzugreifen oder sie in kollaborativen Verfahren zu bearbeiten oder zu ergänzen (von Interesse vor allem für Kooperationen und die Integration in verteilte Grid-Systeme), werden (voraussichtlich Web-basierte) API-Schnittstellen sein, die gegebenenfalls gemeinsam mit konkreten Kooperationspartnern entwickelt werden sollen.

Historische Fachkommunikation im Wandel. Analysen und Trends

Peter Haber, Jan Hodel

Abstract

Der digitale Wandel hat in den letzten Jahren auch die Kommunikation in den Geistes- und Kulturwissenschaften stark verändert. In mehreren Phasen veränderten sich Recherchepraktiken, Publikationsverhalten und die fachinterne Kommunikation. Mit den neuen Möglichkeiten von Web 2.0 stehen weitere, vielleicht sogar tiefgreifendere Veränderungen an. Dieser Beitrag skizziert die bisherige Entwicklung im Feld der Geschichtswissenschaften und diskutiert die Auswirkungen der aktuellen Trends im Kontext von Web 2.0.

In recent years, the digital change in media environment had a significant impact on professional communication in the humanities as well. The practices of searching, the publication behaviour and communication comportment all have altered step by step. The newly evolved web 2.0 now gives way to further changes in this field. The authors outline the development in the history sciences to date and discuss the effects, that current trends of web 2.0 might have on this topic.

Die neuen Informations- und Kommunikationstechnologien haben in den rund zehn letzten Jahren auch in den Geschichtswissenschaften die meisten wissenschaftlichen Arbeitsschritte tiefgreifend verändert. So sind E-Mail, Newsletters oder bibliographische Datenbanken heute aus dem Historiker-Alltag kaum mehr wegzudenken und kaum ein Historiker mag heute noch auf die Dienste von Google, auf die retrodigitalisierten Zeitschriften bei JSTOR oder auch nur auf den Online-Katalog der nächstgelegenen Universitätsbibliothek verzichten. Doch die meisten Historikerinnen und Historiker reagierten zu Beginn eher zurückhaltend auf die Veränderungen und waren kaum bereit, sich aktiv mit den neuen Möglichkeiten der digitalen Medien auseinanderzusetzen. Diese Unlust mag mit dem rasanten Tempo zu tun haben, mit der das Internet und die damit verbundenen Möglichkeiten viele liebgewonnene Gepflogenheiten über den Haufen zu werfen drohte.

Entwicklungsphasen im Rückblick

Als sich Mitte der neunziger Jahre der Durchbruch des Internet abzuzeichnen begann, hatte der Personal Computer für die meisten Historiker allenfalls den Status einer komfortablen Schreibmaschine inne. Ansonsten nutzten sie weiterhin vor allem die vertrauten Hilfsmittel: den eigenen Zettelkasten, die nächstgelegene wissenschaftliche Bibliothek und den für das eigene Forschungsfeld in Frage kommenden Archive.

Das Internet etablierte sich nur langsam in der historischen Forschungspraxis. Einen ersten Einschnitt markierten die übers Netz zugänglichen Bibliothekskataloge. Diese begannen sich in der Schweiz zwar bereits zu Beginn der neunziger Jahre durchzusetzen. Sie wurden aber anfänglich kaum genutzt, da die Bibliotheken ihre Kataloge über das damalige Datennetz Telepac zugänglich machten, was aufwendige Loginprozeduren bedingte. Zudem hatte fast jede Bibliothek ein anderes Katalogsystem mit einer je eigenen Abfragesyntax. So blieben die meisten Historiker lieber beim Zettelkasten und den Microfichen und nutzten den elektronischen Bibliothekskatalog allenfalls vor Ort in der eigenen Bibliothek. Trotzdem eröffnete der Online Public Access Catalog (OPAC) eine grundsätzlich neue Arbeitsweise: Relevante Literatur ließ sich nicht mehr nur mit Hilfe der lokal greifbaren Bibliotheksbestände und einschlägiger (meist umständlich zu nutzender) Fachbibliographien erschließen. Mit ein wenig Einarbeitungszeit waren plötzlich die Bestände wichtiger Bibliotheken zugänglich und erlaubten die gezielte Literatur- und zum Teil auch Quellensuche am Bildschirm. Entscheidend dabei war, dass große US-amerikanische Bibliotheken mit historisch relevanten Altbeständen wie zum Beispiel die New York Public Library oder die Bibliothek der Harvard University ihre Bestände sehr früh annähernd vollständig und sehr gut erschlossen zugänglich gemacht haben.

Das Internet als wissenschaftliches Arbeitsinstrument setzte sich – sieht man von den technischen Wissenschaften ab, wo es bereits früher breite Verwendung fand – gegen Mitte der neunziger Jahre durch. Die eigentliche „Killerapplikation“ war – neben Email – das World Wide Web. Das WWW wurde einige Jahre zuvor von Tim Berners-Lee in Genf entwickelt. Er wollte damit den Wissenschaftlern ein Arbeitsinstrument zur Verfügung stellen, um Forschungsergebnisse auf einfache Art und Weise auszutauschen. Berner-Lees Kerngedanke war, Texte miteinander zu verknüpfen, mit Bildern und anderen nicht-textlichen Medienformaten anzureichern und einfach zu bearbeiten. Diese drei Merkmale – Hypertextualität, Multimedialität und kollaborative Arbeitsweise – sind bis heute die drei zentralen Punkte bei der wissenschaftlichen Nutzung des Internet.

Die Rezeption des Internets in den Geschichtswissenschaften folgte indes einer anderen Logik. Zunächst wurde das Internet fast ausschließlich als Instrument zur Recherche wahrgenommen und genutzt. Ende der neunziger Jahre lernten die Geschichtsstudierenden in Einführungskursen, wie sie mit den neuen Hilfsmitteln arbeiten sollen und was Quellenkritik – eine der zentralen Punkte der historischen Methode – im digitalen Zeitalter bedeuten könnte. Grundsätzlich blieb die Einstellung gegenüber historischen Quellen aus dem Netz von misstrauischer Vorsicht geprägt, denn das herkömmliche Instrumentarium der Quellenkritik schien sich nicht oder nur mit einigen Modifikationen auf Online-Ressourcen anwenden zu lassen.

So übten sich die Historikerinnen und Historiker im Umgang mit dem WWW als einer komfortablen, aber auch fragwürdigen Recherchierhilfe. Zentrale Fragen waren die sinnvolle Erschließung und korrekte Zitierweise von Online-Ressourcen sowie der Aufbau von geschichtswissenschaftlichen Internet-Portalen. Zu den einflussreichsten Pionierprojekten dieser Phase zählte der HistoryGuide, der an der Niedersächsischen Staats- und Universitätsbibliothek Göttingen konzipiert wurde und der nach einer zweijährigen Entwicklungsphase 1997 ans Netz ging. Das Konzept war einfach und zugleich überzeugend: Das bibliothekarische Fachwissen sollte nicht nur zur Erschließung und Bereitstellung von gedruckter Literatur, sondern auch von Online-Ressourcen genutzt werden. Zu diesem Zwecke wurde ein Set von Metadaten entwickelt, das den medienspezifischen Anforderungen bei der Beschreibung von Web-Angeboten entsprach. Dabei mussten einige Fragen, die bei Büchern relativ klar waren, neu gestellt werden: Wie definiert sich zum Beispiel das Publikationsdatum einer Website? Wann kann man von einer Neuauflage sprechen und wie stellt man diese fest? Schon bald zeigte sich, dass das noch relativ junge Medium des WWW nicht diejenigen Paratexte entwickelt hatte, die das Buchwesen schon seit langer Zeit kannte und die den Bibliothekaren ihre Arbeit erleichterten: Klappentexte, Inhaltsverzeichnisse, Register, aber auch Danksagungen und Rezensionen trifft man bei Websites nur vereinzelt und in einer kaum standardisierten Form an.

Auf diese erste Phase, in der die Historiker das WWW eigentlich nur als Recherchierinstrument nutzten, folgte eine zweite Phase, in der sie das Netz als mögliche Plattform zur Selbstdarstellung und zur Publikation eigener Forschungsergebnisse entdeckten. Im deutschen Sprachraum entwickelte sich dabei die Diskussionsliste H-Soz-u-Kult zur wichtigsten Plattform der Scientific Community. H-Soz-u-Kult wurde bereits 1996 gegründet, zählt heute über 14.000 Abonnenten und ist Teil des amerikanischen H-Net, das an der Michigan State University angesiedelt ist. Auf H-Soz-u-Kult werden jährlich rund 1.000 Fachrezensionen veröffentlicht und es gibt kaum eine geschichtswissenschaftliche Stellenausschreibung im deutschen Sprachraum, die nicht auch über die Liste verschickt würde.¹ Der Erfolg von H-Soz-u-Kult liegt vermutlich daran, dass das Angebot gekonnt die Vorteile der verschiedenen Internet-Dienste kombiniert und zugleich Elemente der bisherigen Kommunikationsgewohnheiten innerhalb des Faches berücksichtigt. Der Erfolg von H-Soz-u-Kult wirft aber auch etliche Fragen auf: Welche Auswirkung hat der Beschleunigungseffekt der digitalen Medien auf die Qualität der fachinternen Debatten? Wie verschieben sich dabei die Grenzen zwischen dem wissenschaftlichen Fachdiskurs und einer „Feuilleton-Kultur“, die

¹ Hohls, Rüdiger: H-Soz-u-Kult: Kommunikation und Fachinformation für die Geschichtswissenschaften, in: Mruck, Katja / Gersmann, Gudrun (Hrsg.): Neue Medien in den Sozial-, Geistes- und Kulturwissenschaften. Elektronisches Publizieren und Open Access: Stand und Perspektiven, Köln 2004 (= Historical Social Research; 29 (2004)/1), S. 212-232.

gerade in den letzten Jahren immer wieder auch historische Themen aufgegriffen hat? Obwohl H-Soz-u-Kult die Merkmale eines interaktiven Kommunikationsmediums aufweist, steht in der momentanen Nutzung des Dienstes die nicht-interaktive Publikation insbesondere von Fachrezensionen im Vordergrund.

Dieses und weitere Beispiele zeigen, dass sich das WWW als neuer Publikationskanal für geschichtswissenschaftliche Arbeiten nur sehr zögerlich durchsetzt. Die neuen Dienste orientieren sich fast durchgängig an den Strukturen und Mechanismen des bisherigen wissenschaftlichen Publikationssystems. Die Möglichkeiten des Internet werden nur sehr partiell akzeptiert, am herrschenden Machtgefüge des Faches, der sich natürlich auch in den Strukturen des Publikationsprozesses widerspiegelt, wird nicht gerüttelt.

Medialitäten der Historiographie

Nach dieser kurzen Übersicht über die Entwicklung der fachinternen geschichtswissenschaftlichen Kommunikation im digitalen Zeitalter soll nun der Blick in die Zukunft gerichtet werden. Dabei wollen wir weniger darüber spekulieren, welche technischen Entwicklungen zu erwarten sind und mit welchen Auswirkungen auf die Geschichtswissenschaften zu rechnen sein könnten. Vielmehr wollen wir als Ausgangspunkt unserer Überlegungen die Bedürfnisse der geschichtswissenschaftlichen *Scientific community* nehmen: Von welchen digitalen Hilfsmitteln erwarten wir uns eine Erleichterung und Bereicherung für den wissenschaftlichen Alltag?

Um technophilen Machbarkeitsphantasien vorzubeugen und um zu verhindern, dass wir im Folgenden die Eierlegendewollmilchsau der digitalen Geschichtsschreibung entwerfen, sollen die Anforderungen systematisiert werden. Zur wissenschaftlichen Kommunikation zählen wir dabei nicht nur die Vermittlung wissenschaftlicher Ergebnisse an eine interessierte Öffentlichkeit, sondern auch – und an dieser Stelle primär – die Kommunikation innerhalb der Wissenschaft.

Die wissenschaftliche Kommunikation im Bereich der Geschichtswissenschaften besteht seit rund 200 Jahren, im Schreiben von Darstellungen, in denen vergangene gesellschaftliche Ereignisse, Prozesse und Konstellationen erzählt, gedeutet und erklärt werden. Dabei stellen sich bei allen Arbeitsschritten Fragen, welche die wissenschaftliche Kommunikation betreffen: so etwa bei der Formulierung der Fragestellung, bei der Suche nach entsprechenden Materialien und bei der Auswertung und Interpretation desselben. Die durch die digitalen Medien hervorgerufenen Veränderungen stellt die klassische historische Methode vor teilweise völlig neue Herausforderungen, vor allem bezüglich des Wissensmanagements.

Das hat wohl unbestreitbar mit dem Gegenstand der historischen Disziplin zu tun: Der Untersuchungsgegenstand der Geschichtsschreibung, die Vergangenheit, lässt sich schließlich niemals unmittelbar beobachten, sondern kann nur mittels medialer Aufzeichnungen erschlossen werden. Die Medialität der Aufzeichnung prägt also den Prozess der historischen Interpretation tiefgreifend.

Computer Aided Historiography und virtual H-Desk

Im digitalen Zeitalter muss der geschichtswissenschaftliche Arbeitsplatz eine Reihe von Voraussetzungen erfüllen: Es muss möglich sein, mit digitalen Materialien zu arbeiten, egal ob es sich um Texte, Bilder, Töne oder Videos handelt. Der Arbeitsplatz muss die Möglichkeit bieten, im Austausch mit Fachkollegen sämtliche zur Verfügung stehenden Trägerformate zu verwenden. Und schließlich sollte der *digital historian* in der Lage sein, sein Material in adäquater Form in seinen eigenen Darstellungen zu präsentieren. Diese Art von *Computer Aided Historiography* (CAH)² wird weitgehend netzbasiert funktionieren und es stellt sich deshalb die Frage, wie der klassische Arbeitsplatz des Historikers – der Schreibtisch – in Zukunft aussehen wird.

Wir schlagen vor, die vertraute Metapher des Schreibtisches weiterhin zu verwenden und in diesem Zusammenhang von einem virtuellen Geschichtsschreibtisch, dem *virtual H-Desk* zu sprechen. Der *virtual H-Desk* soll dem Historiker des 21. Jahrhunderts helfen, den wissenschaftlichen Anforderungen im digitalen Zeitalter zu genügen. Die hier präsentierte erste Skizze eines *virtual H-Desk* folgt der Logik des kommunikativen Alltagsgeschehen in den Geisteswissenschaften: Informationen beschaffen, Informationen auswerten, Informationen mit Kollegen austauschen, Ergebnisse formulieren, Publizieren.

In „vordigitaler“ Zeit ließen sich diese einzelnen Prozessschritte noch einigermaßen einfach auseinander halten: Zuerst galt es, in der Bibliothek den Forschungsstand zu überprüfen und einschlägige Bücher und Aufsätze zu suchen und zu exzerpieren. Anschließend konnten relevante Quellenbestände eruiert und erfasst werden – was meistens aufwändige Archivreisen bedingte. In einem nächsten Schritt folgten Diskussionen der vorläufigen Ergebnisse an Tagungen, Seminaren oder allenfalls in den *Miszellen* spezialisierter Fachorgane. Schließlich galt es, die Ergebnisse in Form von Monographien oder Fachaufsätzen niederzuschreiben und nach einem nicht selten mehrstufigen Prozess der Qualitätssicherung mit einer Verzögerung von einigen Jahren der Fachöffentlichkeit zu präsentieren.

Im digitalen Zeitalter, vor allem unter den Voraussetzungen des web 2.0, sind diese Arbeitsschritte immer enger miteinander verzahnt und oftmals gar nicht mehr voneinander zu trennen.

² Haber, Peter: CAH - Computer Aided Historiography, in: weblog.histnet.ch vom 30. März 2007 [<http://weblog.histnet.ch/archives/328>].

Information Retrieval und die Zugänglichkeit von Information

Den Ausgangspunkt bildet die Frage, wie wir an die benötigten Informationen kommen respektive wie wir die gewünschten Informationen zu uns holen. Die Möglichkeit, digitalisierte Daten einfach und in grossen Mengen über grosse Distanzen zu übertragen, hat schon in den Anfangsjahren des Internet zur Vorstellung geführt, das Wissen der Welt werde nun in seiner Gesamtheit allen und jederzeit und womöglich jeweils im gewünschten Aufbereitungsmodus zugänglich.³ Die Realität der Informationsflut einer kaum noch strukturierten Daten-Mélange hat schließlich zu einem Perspektivenwechsel geführt und die Frage dringend werden lassen, wie wir aus dieser ungeordneten Informationsmenge die gewünschte Information herausfischen können.⁴

Die Möglichkeiten „intelligenter“ Informationsdienste wie zum Beispiel RSS-Feeds und Alert-Services, die uns die gewünschte oder vermeintlich gewünschte Information direkt zustellen, definieren lediglich neue Handlungsfelder: Wir müssen die uns zugestellten Informationen nun filtern, abwehren und nicht selten einfach nur noch vernichten. Die Menge an Informationen, die wir dabei zu verarbeiten haben, täuscht leicht darüber hinweg, dass der Zugang zu bestimmten Informationen – insbesondere auch zu Fachinformationen – unter der zunehmend rigiden Auslegung des Urheberrechts in den letzten Jahren zunehmend erschwert worden ist. Allen OpenAccess-Initiativen zum Trotz: Information ist eine Ware, deren Herstellung, Veredelung und Distribution nicht umsonst ist. Die Produzenten und mehr noch die Distributoren wollen für ihre Arbeit abgegolten werden. So wird die Zugänglichkeit von (Fach-)Information für den Einzelnen immer mehr abhängig von Lizenz-Vereinbarungen derjenigen Institution, die den Zugang vermittelt.

Systeme der Ordnung

Ob man die Informationen, die man für die wissenschaftliche Arbeit braucht, nun aktiv zusammensucht, von „intelligenten“ Informationsagenten zugestellt erhält, ob sie umsonst sind, bestellt oder bezahlt werden müssen: sie kommen in den verschiedensten Formaten daher. Dies erschwert ihre Weiterverarbeitung im Kontext eines individuellen Wissensmanagements. Die Ordnung und Strukturierung von

³ Haber, Peter: „Google-Syndrom“. Phantasmagorien des historischen Allwissens im World Wide Web, in: Eppler, Angelika (Hg.), Vom Nutzen und Nachteil des Internet für die historische Erkenntnis. Version 1.0, Zürich: Chronos 2005, S. 73-89; Haber, Peter: Weltbibliothek oder Diderots Erben? Traditionslinien von Wikipedia (Preprint), in: Tagungsband INFORMATIK 2007. 37. Jahrestagung der Gesellschaft für Informatik e.V. (GI), Bremen 2007; Haber, Peter: Alles Wissen dieser Welt. Historische Referenzpunkte von Wikipedia (Manuskript), Basel und Bern 2007.

⁴ Degele, Nina: „Neue Kompetenzen im Internet. Kommunikation abwehren, Informationen vermeiden“, in: Lehmann, Kai (Hg.), Die Google-Gesellschaft. Vom digitalen Wandel des Wissens, Bielefeld: transcript 2005, S. 63-74; Hodel, Jan: „Heidegger in der Strassenbahn oder Suchen in den Zeiten des Internet“, in: Haber, Peter (Hg.), Geschichte und Internet: „Raumlose Orte - geschichtslose Zeit“, Zürich: Chronos 2002, S. 35-48.

Informationen in Ordnungssystemen sind der Mehrwert, den Individuen aus der unendlichen Menge jedermann zugänglicher Informationen generieren können.

In einem von Google und anderen Suchmaschinen dominierten Arbeitsumfeld, in dem man mit wenigen Mausklicks zu fast jedem Suchbegriff Tausende von Treffern generieren kann, werden Fähigkeiten der Auswahl und der Strukturierung von Suchvorgängen und Suchergebnissen zu Distinktionsmerkmalen bei der erfolgreichen wissenschaftlichen Kommunikation. Umso bemühter ist es daher, wenn man die ausgewählten Informationen einfach in einem unstrukturierten Format ablegen muss. Das Problem liegt bei der fehlenden Einheitlichkeit anwendbarer Metadaten und der mangelnden Konsequenz ihres Einsatzes. Viel wäre schon gewonnen, wenn alle Informationen, die im WWW zugänglich sind, zumindest mit Metadaten zu Autor, Titel, Datum und URL enthielten. Wünschenswert wäre eine flexible Lösung eines Informationssystems für die Ablage von beliebig strukturierten Daten, die einfach ein- und ausgelesen werden können.

Schon seit Jahren ist vom semantischen Web die Rede, in dem Dokumente nach klaren Regeln so strukturiert werden, dass Maschinen den Inhalt nicht nur lesen sondern auch „verstehen“ können. Im semantischen Web erkennt beispielsweise eine Suchmaschine dank dieser inhaltlichen Strukturen, dass ein Dokument biografische Informationen zu „Picasso“ enthält und Teil des Katalogs einer großen Sonderausstellung ist. Im semantischen Web kann dank der Struktur nun maschinell eine Verbindung von „Picasso“ zu „Maler, Künstler“ und „20. Jahrhundert“ erstellt werden, da in anderen Texten Picasso als Künstler des 20. Jahrhundert definiert wurde. Eine Suchanfrage zu „Ausstellungen Künstler 20. Jahrhundert“ führt folglich besagtes Dokument auf. Dass eine solche Informationsstrukturierung den Umgang mit Wissensressourcen nachhaltig verändern wird, ist offensichtlich: Es wird möglich, Informationsbestände ad hoc zu generieren und nach immer neuen Kriterien zu ordnen. Die Trefferliste wird kombiniert aus eigenen Informationsbeständen ebenso wie aus Fremddaten, die zum Beispiel von Bibliotheken oder digitalen Repositorien stammen. Die Herkunft der Informationen und ihre frühere Verwendung bleibt zwar transparent, verliert aber an Bedeutung.

Wenn zum Beispiel die Informationen zu einem Aufsatz (bibliographische Angaben, Keywords, zitierte Literatur, methodische und theoretische Bezüge) nicht mehr verstreut auf verschiedene Aufzeichnungs- und Ordnungssysteme wie Text-Dokument, Tabellenkalkulation oder Literaturlistenbank, sondern gebündelt in einem durchlässigen Format abgelegt wären, kann sich der Umgang mit Informationen nachhaltig verändern. Gerade in den Geisteswissenschaften, wo aufgrund uneinheitlicher Terminologien, zahlreicher Sprachen und den unterschiedlichsten Zitationskulturen babylonische Zustände herrschen, käme eine solche rechner- und webgestützte Lösung zur Informationsverwaltung voll zur Geltung.

Soziale Vernetzung

Neben der maschinenunterstützten Strukturierung von Informationen ist mit Web 2.0 ein konkurrierendes Konzept populär geworden: Das chaotische, unkoordinierte Zusammenwirken vieler Individuen bei der Generierung und Strukturierung von Inhalten, wie wir es zum Beispiel bei Wikipedia kennen. Trotz vieler Vorbehalte gerade für die wissenschaftliche Nutzung solcher Dienste, lassen sich bei diesen Diensten auch durchaus nützliche und nutzbare Ergebnisse erzielen. Web 2.0 macht deutlich, dass eine soziale Vernetzung sich sehr gut mit dem Austausch fachbezogener oder gar wissenschaftlicher Informationen verbinden lässt – viel besser vermutlich, als in der vor-digitalen Zeit. Wer etwa die bislang eher spärlichen Erfahrungen mit wissenschaftlichen Weblogs analysiert, stellt fest, dass diese oftmals mit „schneller Feder“ geschriebenen Berichte aus dem wissenschaftlichen Alltag ganz neue Möglichkeiten eröffnen, um sich ein neues Feld gleichzeitig fachlich und sozial zu erschliessen. Weblogs bieten nämlich nicht nur die Möglichkeit zur Kommunikation über wissenschaftliche Fragen, sondern sie stellen auch eine Chance dar, in einem sozialen Netz wissenschaftlich Gleichgesinnter (in der Regel) unkoordiniert arbeitsteilig Informationsströme zu filtern und als relevant identifizierte Informationen in kommunizierbare Einheiten aufzubereiten.

Dennoch wird sich auch hier eine neue Balance entwickeln müssen: Was behält man „geheim“, weil man befürchtet, dass egoistische Kollegen einem gute Ideen abkupfern oder von der eigenen altruistisch gesinnten Vorarbeit profitieren, ohne selbst etwas Gleichwertiges zu bieten? Was stellt man vielleicht auch voreilig ins Netz und muss immer wieder die eigenen Informationen korrigieren? Hier stellen sich erneut die eingangs gestellten Fragen nach dem „Wert“ von Informationen und nach den urheberrechtlichen Ansprüchen. Wie werden sich hier die unterschiedlichen Anforderungen und Bedürfnisse, letztlich aber auch unterschiedliche Kulturen beim Umgang mit Wissen in der alltäglichen Praxis der wissenschaftlichen Kommunikation niederschlagen? Konkret: Können wir in einem wissenschaftlichen Kontext in naher Zukunft davon ausgehen, dass Einträge in einem Weblog den Charakter von akzeptierten Publikationen haben, sodass man Anspruch auf erstmalige Äußerungen eines Gedankens oder eines Forschungsergebnisses in einem Weblog-Eintrag geltend machen kann (wie das hier mit der in Fussnote 2 vermerkten Eintrag versucht wurde, wo erstmals der Ausdruck *Computer Aided Historiography* verwendet wurde)? Oder werden die Weblogs dazu führen, dass der individuelle Beitrag zur vielstimmigen, ständig aktiven wissenschaftlichen Kommunikation noch mehr an Gewicht verliert?

Digitale Schreibutensilien

Das Schreiben bleibt, wie wir weiter oben gezeigt haben, auch im digitalen Zeitalter eine zentrale Form der wissenschaftlichen Kommunikation im Fach Geschichte. Die Bedingungen des Schreibens verändern sich aber maßgeblich in Bezug auf die damit verbundenen sozialen Interaktionen und auf den strukturellen Aufbau der

Schriftstücke. Die Verbindungen zu den eben aufgeführten Veränderungen bei der sozialen Vernetzung und bei der Strukturierung der Information sind offensichtlich. Das digitale Schreibutensil der Zukunft müsste den unterschiedlichen Anforderungen durch Flexibilität Rechnung tragen: Der Einbezug unterschiedlicher Autoren mit unterschiedlichen Rechten und Autorschaften an unterschiedlichen Dokumenten muss ebenso möglich sein (wie das Google-Docs bereits anbietet) wie die Integration unterschiedlicher Daten (Bibliographie, Bilder, Zitate, Exzerpte, Transkripte, numerische Daten, Tabellen etc.) aus unterschiedlichen Ressourcen sowie deren Nachweis. Annotationen und Versionskontrollen gehören selbstverständlich dazu, um nachvollziehen zu können, wer für welche Änderungen oder Ergänzungen verantwortlich zeichnet. Das Schreibutensil soll ebenso die Erstellung klassischer, linearer Texte wie von Hypertexten ermöglichen.

Das gemeinschaftliche Erstellen von Texten wird in Zukunft an Bedeutung gewinnen und neue Formen der Zusammenarbeit und der Textproduktion bedingen. Dennoch werden in Zukunft nicht alle Texte zwangsläufig als kooperative Hypertexte publiziert werden. Wahrscheinlicher scheint uns eine funktionale und mediale Ausdifferenzierung entsprechend den verschiedenen Bedürfnissen und Ansprüchen unterschiedlicher Nutzungsszenarien von wissenschaftlichen Texten. Aber auch bei der Erstellung mono-aktorialer, linearer Texte wird die Praxis des gemeinschaftlichen Schreibens von modularen, vernetzten Texten einen Einfluss haben und die Rahmenbedingungen der Produktion prägen.

Fazit

Was uns mit dem *virtual H-Desk* vorschwebt, ist eine rechnergestützte Umgebung, bei der die verschiedenen medialen und funktionalen Ebenen des Systems eng miteinander verzahnt sind, als relevant identifizierte Informationen einfach und umfänglich zugänglich sind und verlustfrei in eigene Ordnungssysteme integriert und dort bearbeitet werden können. Die Zusammenarbeit und der Austausch mit Kolleg/innen soll sich einfach, produktiv und transparent gestalten. Gerade der letzte Punkt ist indes nicht primär eine Frage der Technik, sondern der Verfassung des sozialen Systems „Wissenschaft“. Aus diesem Grund sollte beim Einsatz digitaler Medien in der wissenschaftlichen Kommunikation nicht die zu erwartende Effizienzsteigerung im Mittelpunkt stehen, sondern eine nachhaltige Sozialverträglichkeit und eine Abstimmung auf die Bedürfnisse und Traditionen des jeweiligen wissenschaftlichen Feldes. Auch mit dem *virtual H-Desk* wird der gute alte Schreibtisch nicht ausgedient haben, wird der Gang in die Bibliothek und ins Archiv weiterhin zum Berufsalltag des Historikers gehören. Die wissenschaftliche Kommunikation wird aber vielschichtiger und vielleicht auch dynamischer werden. Und vielleicht wieder an gesellschaftlicher Relevanz gewinnen.

Konzepte zur Förderung der Wissenschaftskommunikation: Der Churer Ansatz zur Vermittlung von kollaborativen Kompetenzen

Lydia Bauer, Nadja Böller, Josef Herget, Sonja Hierl

Abstract deutsch

Die Wandlung der Anforderungen an das Individuum im Zuge der weitgehenden Digitalisierung der Arbeitsumgebungen hat dazu beigetragen, dass in der heutigen Informationsgesellschaft Informations- und Medienkompetenz zu den wichtigsten Schlüsselqualifikationen zählen.

Ebenso gewinnen Aspekte wie Wissenschaftskommunikation, Teamarbeit und Wissensaustausch im Berufsleben zunehmend an Bedeutung; kollaborative Arbeitsstile werden in der neuen, informationell vernetzten Arbeitsumgebung dominant. Die Vermittlung von grundlegenden Kompetenzen zur Teilnahme an wissenschaftlichen Kommunikationsprozessen sollte bereits zu Beginn eines Studiums beginnen. Aktuelle Konzepte wie Web2.0, die gezielte Nutzung von digitalen Bibliotheken und neuen Formen der kollaborativer Wissensentwicklung sollten Eingang in moderne Lehr- und Lernformen finden. Der Bibliothek kommt in diesem Prozess eine wesentliche Rolle zu, da sie mit ihren Dienstleistungen und Angeboten Studierende und Forschende maßgeblich unterstützen kann.

Das Konzept der Knowledge-enhancing Helix schafft einen Bezugsrahmen für eine in das Curriculum integrierte, umfassende und aktive Vermittlung von Methoden-, Sozial-, Fach-, Medien- und Informationskompetenz. Der zugrundeliegende Ansatz des Blended Learning wird hierbei in mehrwertschaffender Form mit diversen neuen Medien und Systemunterstützungen (Topic Map basierte eLearning-Webseite, Mediawiki, Lernplattform etc.) ergänzt, die in einem den Lehr- und Lernprozess unterstützenden Framework integriert sind und durch die Dienstleistungen von Bibliotheken abgerundet werden.

Abstract englisch

In an increasingly e-literate society, new forms of scientific communication and new technologies are being developed whereas traditional teaching approaches are challenged to meet new requirements. Further aspects such as teamwork and knowledge exchange are also becoming increasingly important both in the context of education and the daily working life. Collaborative work methods are more dominant in the networked working environment

The development of competencies is a crucial condition for being able to take part in the scientific communication process and needs to be taught from the very beginning of academical education. New concepts like Web2.0 and the goal oriented usage of

digital libraries as well as library services need to be embedded in the teaching process.

The objective of this paper is to show how the implementation of a holistic teaching approach following the model of a knowledge-enhancing helix including the idea of blended learning can be an effective way to train students in collaborative and academic working skills, methodological expertise, information competence, as well as making students aware of new developments in media literacy.

1. Wissen und Lernen – gegenwärtige Herausforderungen der Informationsgesellschaft

In der heutigen Informationsgesellschaft zählen Informations-, Medien-, Fach-, Methoden- sowie Sozialkompetenz zu den wichtigsten Schlüsselqualifikationen. Teamarbeit und Wissensaustausch sowie Wissensbewahrung erhalten eine immer größere Bedeutung. Kollaborative Arbeitsstile und netzwerkartige Strukturen werden in einer zunehmend vernetzten Arbeitsumgebung dominant und sind wesentlich geprägt durch die Entwicklungen der Informations- und Kommunikationstechnologien und neuen Medien.

Organisations- und Kommunikationsprozesse in der Wissenschaftskommunikation und Forschung wandeln sich aufgrund sich verändernder Arbeitsbedingungen in zeitlicher und räumlicher Dimension. Dem effizienten Umgang und Austausch von Wissen und Information kommt dabei auf organisationaler sowie individueller Ebene eine entscheidende Rolle zu (Mönnich 2006). Die beruflichen Anforderungen an das Individuum im Bereich der Wissensorganisation und -kommunikation verändern sich rasant und eine permanente Entwicklung und Anpassung bzw. Optimierung dieser persönlichen Fähigkeiten wird konstant gefordert.

Mit den aktuellen Entwicklungskonzepten des Web2.0 entstehen vielfältige neue Gestaltungsmöglichkeiten, die es gilt gewinnbringend in organisationale Prozesse einzubinden. Der Einsatz von Social Software wie beispielsweise Wikis, Chats oder Blogs hat sich in den vergangenen Jahren nicht nur im privaten Umfeld etabliert, sondern spielt auch in der Forschung als Instrument der Wissenschaftskommunikation sowie in Unternehmen eine zunehmend wichtigere Rolle (Chau & Maurer 2005). Dies stellt Institutionen wie Bibliotheken oder Dokumentationsabteilungen, die solche Ressourcen zur Verfügung stellen, vor neue Herausforderungen. Wiki-Tools sind zwar einfach zu installieren, ein effizienter und nachhaltiger Umgang mit dieser Art von Kollaboration und Kommunikation ist jedoch nur gegeben, wenn das Instrument auch gepaart mit den passenden Methoden in den Lernprozess eingebunden und entsprechend den Bedürfnissen der Benutzer angepasst wird. Studierende sowie Mitarbeiter müssen flexibel und bewusst auf diese Entwicklungen und Veränderungen reagieren können. Lernprozesse und Weiterbildung bedeuten unter dem Stichwort des lebenslangen Lernens dabei nicht nur den Erwerb neuen Fachwissens sondern gleichzeitig auch persönliche Entwicklung in Bezug auf neue Arbeitsstile, Betriebs- und Organisationsformen, dem Umgang mit neuen und oftmals unerwarteten Situationen und somit die

Herausbildung und Entwicklung von zukunftsfähigen Schlüsselqualifikationen (Mönnich 2006).

Die berufsqualifizierende Ausbildung an Hochschulen steht vor der Herausforderung auf diese Anforderungen der Arbeitswelt zu reagieren, indem bereits während des Studiums großer Wert auf die Herausbildung von Schlüsselkompetenzen, wie Informations- und Medienkompetenz, Kritikfähigkeit und analytische Denkfähigkeit sowie der Sensibilisierung für neue Technologien und dem effizienten Austausch und Erwerb von Wissen gelegt wird (Wagner & Kleinberger Günther 2004).

2. Lösungskonzepte aus der Hochschullehre

Das Konzept „DIAMOND“ (Didactical Approach for Media Competence Development) (Hierl et al. 2007) schafft einen Bezugsrahmen für eine in das Curriculum integrierte aktive und umfassende Vermittlung der Schlüsselkompetenzen, zudem liegt der Fokus auf dem effizienten Austausch und Erwerb von Wissen zur Gestaltung einer strukturierten Wissensordnung. Das über mehrere Jahre hinweg entwickelte und im Curriculum der Informationswissenschaft an der Hochschule für Technik und Wirtschaft in Chur eingesetzte Konzept folgt einem aus sechs Bausteinen bestehenden Modell:

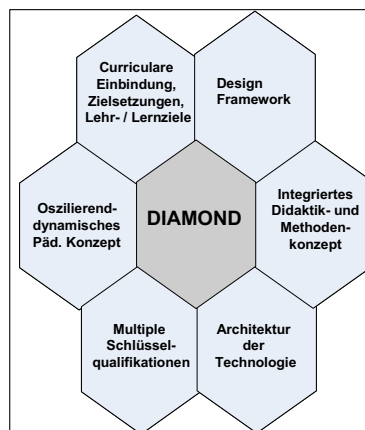


Abb.1: DIAMOND-Modell

Der zugrunde liegende Ansatz des Blended Learning wird hierbei in mehrwertschaffender Form mit diversen Medien und Systemunterstützungen (Topic Map basierte eLearning-Webseite, Mediawiki, Lernplattform etc.) ergänzt, die in einem Lehr- und Lernprozess unterstützenden Framework integriert sind.

Der pädagogisch-didaktische Schwerpunkt des Lehransatzes liegt dabei in der Förderung einer kollaborativen Arbeitsweise durch den Einsatz einer Wiki-Software. Das Konzept fußt auf dem Modell der Knowledge-enhancing Helix, das den theoretischen Grundlagen für gutes Online-lernen (Alley & Jansak 2001) folgt:

1. Wissen wird konstruiert.
2. Der Lernprozess wird effektiver, wenn der Lernende die Verantwortung für sein eigenes Lernen übernehmen kann.
3. Die Motivation des Lernenden ist ein einflussreicher Faktor für Lernerfolg und –auswirkungen.
4. Wissen wird konstruiert.
5. Der Lernprozess wird effektiver, wenn der Lernende die Verantwortung für sein eigenes Lernen übernehmen kann.
6. Die Motivation des Lernenden ist ein einflussreicher Faktor für Lernerfolg und –auswirkungen.
7. Anspruchsvolles Lernen erfordert Reflektion.
8. Der Lernprozess ist einzigartig für jedes Individuum.
9. Der Lernprozess basiert auf Erfahrung.
10. Der Lernprozess ist sowohl sozial als auch privat.
11. Rigorose erkenntnistheoretische Vorannahmen können anspruchsvolles Denken und Lernprozesse irreleiten.
12. Der Lernprozess erfolgt in Spiralen.
13. Der Lernprozess ist „chaotisch“.

Um diese Kriterien erfüllen zu können, sind als Basis verschiedene Kompetenzen und Fähigkeiten notwendig. Kompetenzen werden in diesem Zusammenhang ganzheitlich als ein Zusammenspiel verschiedener Kenntnisse, Können und Einstellungen verstanden (Mönnich 2006). Der Hauptfokus liegt dabei auf Aspekten wie Kollaboration, Kooperation und Kommunikation sowie individuellem und selbstgesteuertem Lernen und professioneller Betreuung (Fresen & Boyd 2005).

2.1. Prozess des kollaborativen Arbeitens

Ausgehend von dieser Erkenntnis wird im Folgenden aufgezeigt, wie innerhalb einer kollaborativen Arbeitsumgebung anhand von sechs Schritten der Knowledge-enhancing Helix die Studierenden die Fähigkeit erlangen, ihr Wissen zu erweitern, kritisch zu reflektieren und dabei Schlüsselkompetenzen herauszubilden:

1. Fokussieren

Die Studierenden beschaffen aktiv Informationen zu einem vorgegebenen Themenbereich. Hierbei identifizieren sie offene Forschungsfragen und diskussionswürdige Themenstellungen, indem sie deren Relevanz begründen und innerhalb einer Gruppe in einem Forum diskutieren.

2. Schreiben

Die Studierenden werden bei der selbständigen Bearbeitung eines Papers innerhalb der kollaborativen Arbeitsumgebung durch eine Wiki-Software in ihrem Arbeitsprozess unterstützt.

3. Präsentieren

Alle Arbeiten werden für alle Kommilitonen zugänglich gemacht. Die Studierenden entwickeln dabei Mut und Selbstvertrauen, ihre eigenen Ausführungen den Mitstudierenden zur Verfügung zu stellen.

4. Reflektieren

Die Studierenden lesen eine Auswahl der von den Kommilitonen erarbeiteten Ausführungen analytisch und kritisch durch. Sie sind dazu angehalten, das Gelesene zu hinterfragen und reflektieren und allenfalls weitergehende Recherchen zum Thema durchzuführen.

5. Kritisieren

Die Studierenden kommentieren und diskutieren die einzelnen Arbeiten konstruktiv mit Hilfe der Kommentarfunktion der Wiki-Software.

6. Kollaborative Optimierung

Auf Grundlage der Kommentare entscheiden die Studierenden, wo und wie sie ihre eigene Arbeit noch verbessern können und lernen mit positiver wie auch negativer Kritik umzugehen, indem sie ihre Arbeit aktiv reflektieren und eigene Problemlösungsstrategien entwickeln (Himpsl 2007).

2.2. Illustration der Knowledge-Enhancing Helix

Jeder Schritt in diesem Prozess integriert die in der vorhergehenden Phase erarbeiteten Fähigkeiten und Inhalte und führt somit zu kontinuierlicher Verbesserung. Gleichzeitig wird Wissen erworben, weitergegeben und eine Wissensbasis aufgebaut. Dieser repetierbare und kontinuierliche Prozess lässt sich plausibel mit folgender Helix darstellen (Abb2).

Die Spirale und somit stufenartige Erweiterung zeigt den Prozess, der Schritt für Schritt zur Wissensausweitung und Verbesserung führt.

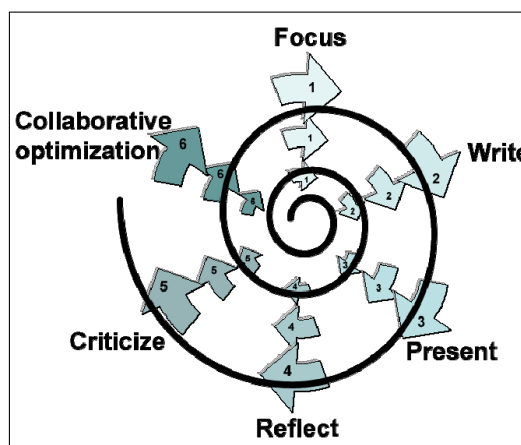


Abb. 2: Knowledge-enhancing Helix

Diese pädagogisch-didaktische Methode wird in der Hochschule für Technik und Wirtschaft im Curriculum Informationswissenschaft seit mehreren Jahren innerhalb der kollaborativen DIAMOND Umgebung eingesetzt (Böller, Herget & Hierl 2007). Hierzu wurde die Wiki-Software Mediawiki (Mediawiki 2007) ausgewählt. Im Gegensatz zu eher traditionellen Formen der Online-Kommunikation wie E-Mail, Forum oder Chat hat eine Wiki-Software offensichtliche Vorteile für die Online-Kollaboration (Abfalterer 2007). Weiter ist die Benutzung einer Wiki-Software sehr einfach zu erlernen und keine spezielle Client-Software ist nötig.

3. Bedeutung für die Wissenschaftskommunikation

Die Wissenschaftskommunikation ist stark durch netzwerkartige Strukturen geprägt. Kollaboration und Kommunikation spielen dabei innerhalb der wissenschaftlichen Community eine zentrale Rolle, um effizient arbeiten und langfristige Ergebnisse erzielen zu können. Wie das Konzept DIAMOND und insbesondere die didaktisch-pädagogischen Grundüberlegungen der Knowledge-enhancing Helix aufzeigen, führt eine systematische Entwicklung und Vermittlung von Schlüsselkompetenzen zu einer aktiven Förderung und Verbesserung der Arbeits- und Lernprozesse.

So werden in der Wissenschaftskommunikation Fähigkeiten gefordert, wie beispielsweise das Identifizieren von relevanten und weit reichenden wissenschaftlichen Fragestellungen, das analytische und reflektierende Lesen von wissenschaftlichen Ausarbeitungen sowie das Üben und Annehmen konstruktiver Kritik.

Es gilt folglich das Konzept zur Entwicklung von Schlüsselkompetenzen nicht nur innerhalb der ersten Jahre einer Hochschulausbildung zu etablieren, sondern die wissenschaftlichen Institutionen und Wissenschaftler selbst dazu anzuhalten, das Konzept kontinuierlich einzusetzen, um die Entwicklung der Schlüsselkompetenzen gewinnbringend, wie auch aktiv zu fördern und schließlich erfolgreiches Lernen zu ermöglichen.

3.1 Bibliotheken als Kompetenzvermittler

Diesen Überlegungen folgend lässt sich die Bibliothek als einer der wichtigsten Player im wissenschaftlichen Kommunikationsprozess identifizieren. Sie fungiert als Kompetenzzentrum im Netzwerk des Lernens (Lux & Sühl-Strohmenger 2004). Unter dem Stichwort der Informationsversorgung besteht die Aufgabe einer Bibliothek dabei nicht nur in der Bereitstellung digitaler Ressourcen sondern vielmehr auch in der aktiven Unterstützung im Lernprozess von Studierenden und Forschenden. Einerseits werden Informationen anhand zielgerichteter und adäquater Vorgehensweisen gewonnen, andererseits müssen diese aber auch bewertet, verstanden und verarbeitet werden (Lux & Sühl-Stromberger 2004). Ein weiterer entscheidender Schritt liegt in der Sicherstellung von qualitativ hochwertiger Information aus gesicherten und relevanten Quellen. An dieser Stelle lässt sich ein unmittelbarer Berührungspunkt zum Konzept der Knowledge-enhancing Helix erkennen: Die Bibliothek unterstützt durch ihr Angebot die Entwicklung der

Schlüsselqualifikationen, indem sie auf die Bedürfnisse der Benutzer zugeschnittene Dienstleistungen zur Verfügung stellt und die Anwender aktiv in der korrekten Nutzung schult.

3.2 Die Rolle Bibliothek innerhalb der Knowledge-Enhancing Helix

Die Bibliothek in ihrer bestehenden und neuausgerichteten Dienstleistungslandschaft wird geprägt durch eine Vielfalt von Anforderungen und neuen Technologien, die zu einem Wandlungsprozess der Angebote führt, um die sich ändernden Benutzerbedürfnisse laufend zu erfüllen. Folgende Tabelle beschreibt und klassifiziert bibliothekarische Serviceleistungen:

Traditionell	Neu
Bibliothekskatalog	Fachportale
Bibliografien	Suchmaschinen
Lexika / Nachschlagewerke	Mailinglisten
Statistiken	Newsletter
Biografien	Alerting-Dienste
Lehrmedien	Weblogs
Fachthesauri	Newsfeeds
Print-Zeitschriften	Chat
Dokumentlieferdienste	Elektronische Zeitschriften
Datenbanken (Literatur, Fakten, Volltext)	Elektronische Dokumentlieferdienste
Sondersammlungen	FAQs
Fernleihe	Wiki
Ausleihe	
Buchaufstellung	
Benutzerschulung	

Tab. 1: Traditionelle und neue Dienste der Bibliotheken

Im Folgenden wird skizziert, wie die Bibliothek mit Hilfe und auf Grundlage dieser Dienstleistungen und Angebote in den einzelnen Phasen im Prozess der Knowledge-enhancing Helix eine zentrale Rolle einnimmt zur Unterstützung und Ermöglichung des Lernprozesses.

1. Fokussieren

Der Umgang mit Informationen und auch die Bewertung und Verarbeitung von umfangreichen Informationsquellen ist eine wichtige Grundvoraussetzung für erfolgreiches wissenschaftliches Arbeiten. Zu Beginn jedes wissenschaftlichen Erkenntnisprozesses steht die Erforschung und Abgrenzung eines Themenbereichs. Dabei ist die Kenntnis aller vorhandenen Informationsquellen unerlässlich (Brändli 2007). Bibliotheken stellen hierfür Instrumente und Dienstleistungen wie digitale Ressourcen (beispielsweise Online-Zeitschriften, Datenbanken, Bibliothekskatalog, Lieferdienste etc.) zur Verfügung.

Studierende beschaffen aktiv Informationen unterschiedlichen Recherchestrategien folgend und beurteilen die Qualität und Zuverlässigkeit der verwendeten Quellen. In einem weiteren Schritt bewerten sie die Informationen selbst und betten sie in einen Gesamtkontext ein. Hierzu ist die Hilfestellung durch die Bibliothek notwendig, indem innerhalb von Benutzerschulungen aufgezeigt wird, wie die Angebote der Bibliothek effizient genutzt werden können. Wichtig erscheint zudem, dass die Bibliotheksbenutzer sowie auch Nicht-Nutzer mit den Dienstleistungen der Bibliothek vertraut gemacht werden und stets über Neuerungen informiert werden beispielsweise anhand von Mailinglisten oder Lieferdienste. Durch regelmäßiges Informieren und Schulen der Bibliotheksbenutzer sinkt die Hemmschwelle und es wird erkannt, wie wichtig es ist, hochwertig aufbereitete Informationen zu verwenden. Insbesondere im Zeitalter einer durch das World Wide Web und Google geprägten Gesellschaft ist es entscheidend, sich auf qualitativ hochstehende und valide Informationen zu konzentrieren. Hierbei gilt es zu lernen, wie beispielsweise zwischen primären und sekundären Informationsquellen unterschieden werden kann oder wie sich der Zweck und die Adressaten von Informationsquellen identifizieren lassen (Brändli 2007). Die Ergebnisse einer Umfrage bereits im Jahr 2002 zeigen auf, wie groß aber auch bedeutungsvoll die Herausforderung an Bibliotheken ist, Wissenschaftlern und Studierenden vor diesem Hintergrund bezüglich der Adressierung adäquater Informationsquellen im Rechercheprozess zu Schulen: So gaben lediglich 13,9% der befragten Personen an, sich bei der Suche nach Printzeitschriften an einen Bibliothekaren zu wenden, wovon nur 3,2% diese Konsultation als den bevorzugten Weg zur Identifizierung der Information bewerteten (Friedlander 2002).

2. Schreiben

Studierende bekommen bereits zu Beginn des Studiums eine Auswahl an methodischen Strategien und Fertigkeiten vermittelt, anhand derer sie sich im weiteren Studienverlauf eine effektive wissenschaftliche Arbeitsweise aneignen können. Die Bibliothek innerhalb einer Bildungsinstitution übernimmt auch hier eine Schlüsselrolle, indem sie Benutzerschulungen zum Erlernen und Vertiefen wissenschaftlicher Arbeitsweise anbietet. Hierzu gehören insbesondere Methoden wie beispielsweise das effiziente Lesen und Zusammenfassen von Texten, Abstracts schreiben, Thesen aufstellen und diskutieren, eine diskursive

Argumentationsführung, das Verfassen wissenschaftlicher Arbeiten mit korrekter Zitierweise sowie Techniken zur Aufbereitung von Information und die zielgerichtete Präsentation. Für Studierende stellen derartige Dienstleistungen eine Ergänzung zum regulären Studium dar. Dies erfordert eine aktive Zusammenarbeit und Austausch der Bibliothek mit der Lehrinstitution. Für Wissenschaftler wiederum zeigt sich damit eine optimale Möglichkeit auf, die aktuellen Entwicklungen in den Methoden des wissenschaftlichen Arbeitens zu verfolgen und vor allem zu vertiefen.

3. Präsentieren

Die Arbeitsweise im wissenschaftlichen Erkenntnisprozess ist stark geprägt durch die Möglichkeiten technischer Systemunterstützung. Nebst dem routinemässigen Bedienen von aktuellen Hard- und Softwareanforderungen, spielt die Kenntnis der aktuellen technologischen Entwicklungen eine große Rolle. Im Helix-Prozess erlernen die Studierenden das Schreiben und Präsentieren einer Seminararbeit innerhalb einer WIKI-Software. Dadurch werden die Studierenden mit neusten technologischen Entwicklungen vertraut gemacht und lernen gleichzeitig mit den Herausforderungen aktueller Konzepte wie Web2.0 umzugehen. Unter dem Stichwort Library2.0 wird auch für Bibliotheken der Einsatz aktueller Werkzeuge dadurch von grösserer Bedeutung (Danowski & Heller 2006). Zentral erscheint dabei der Aspekt, dass die Bibliotheksbenutzer genau darüber informiert werden, welche Informationssysteme zu welchem Zweck und Inhalt zur Verfügung stehen. Weiterhin wird der klassische Prozess des Publizierens im kleinen Rahmen nachvollzogen und den Studierenden die Relevanz der adäquaten Präsentation und Verbreitung von neuen Erkenntnisgewinnen vermittelt.

4-6. Reflektieren, Kritisieren und kollaborative Optimierung

Der kollaborative Helix-Ansatz führt zum aktiven Aneignen von Sozialkompetenz, hierdurch erhöhen Studierende die eigene Kritikfähigkeit und schulen ihre Fähigkeit im kritischen und analytischen Lesen und Denken. Die gegenseitige Strenge auch in formalen Angelegenheiten, wie dem korrekten Zitieren, führt zudem zur frühzeitigen Sensibilisierung dieser wichtigen wissenschaftsethischen Arbeitsform. Der Prozess der gegenseitigen Korrektur kann als Ansatz des Peer-Reviewing verstanden werden, wie er im wissenschaftlichen Publikationszyklus in der Regel stattfindet.

Um Texte kritisch lesen und beurteilen zu können, bedarf es vertiefter Kenntnisse und Fertigkeiten in weiterführender Literaturrecherche. Texte müssen in einen Gesamtzusammenhang gesetzt werden können, um sie beurteilen und bewerten zu können. Wie bei Schritt 1 erwähnt, besteht hier die Rolle der Bibliothek im zur Verfügung stellen der nötigen digitalen Ressourcen und Hilfestellungen bei der Recherchearbeit. Techniken wie das Durchführen beispielsweise von Schneeballrecherchen sind ein Werkzeug, das den Studierenden durch eine Bibliothek vermittelt werden sollte.

Das pädagogische Konzept der Knowledge-enhancing Helix greift auf vielen Ebenen ineinander. Der leicht verständliche und plausible Ansatz garantiert bei allen

Beteiligten eine nachhaltige Entwicklung der Schlüsselkompetenzen und trägt zu erfolgreichem Lernen auf individueller sowie organisatorischer Ebene bei. Das Aneignen von Schlüsselkompetenzen fokussiert sich schliesslich in der persönlichen Entwicklung eines jeden einzelnen und stellt somit einen wichtigen Impuls zum lebenslangen Lernen dar.

All die beschriebenen Punkte zeigen deutlich auf, dass die Bibliotheken im Lernprozess eine bedeutende Rolle spielen und die angebotenen Dienstleistungen, wie in folgender Abbildung dargestellt, den Prozess massgeblich unterstützen können.

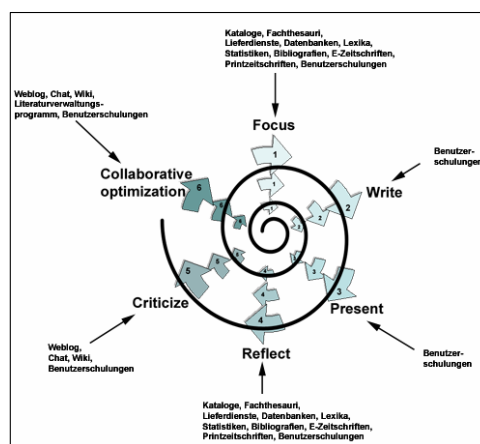


Abb. 3: Bibliotheksservices im Prozess der Knowledge-enhancing Helix

Der Stellenwert der Bibliothek im Prozess der Knowledge-enhancing Helix wird somit durch die obigen Ausführungen charakterisiert und definiert. Es handelt sich dabei um einen idealtypischen Prozess, der aufzeigt, wie sich Dienstleistungen und Instrumente wissenschaftlicher Bibliotheken als fester Bestandteil in Lehre und Forschung integrieren lassen. Eine enge Zusammenarbeit der Bibliothek mit Lehre und Forschung ist deshalb unbestritten. Hier gilt es zusätzlich zu beachten, dass Studierende je nach Studiengang aber auch Wissenschaftler selbst über unterschiedliche Kompetenzen verfügen und Benutzerschulungen entsprechend angepasst werden müssen (Brändli 2007). Schließlich bleibt die bedarfsgerechte Nutzung für die Qualität und Entwicklung der Schlüsselkompetenzen entscheidend (Brändli 2007).

Fazit

Die Schaffung einer umfassenden Wissenskultur in der Aus- und Weiterbildung ist eine unerlässliche Notwendigkeit, um Studierende optimal auf den künftigen Arbeitsmarkt vorzubereiten. Ebenso ist das Teilen und Verteilen von Wissen für Organisationen, wie auch für die Hochschulen von immenser Bedeutung. Durch diesen Prozess kann sichergestellt werden, dass wertvolles Wissen Verbreitung findet, wachsen kann, über den einzelnen Wissensträger hinaus zur Anwendung

kommt (Reinmann-Rothmeier 2001) und dadurch wieder Impulse für neues Wissen freisetzt. Die Vermittlung von Schlüsselqualifikationen, das Wissen um die richtige Auswahl von Informationen und die Befähigung zum Wissenstransfer in einer sich rasant verändernden, eng vernetzten Informationsgesellschaft verlangt eine enge Zusammenarbeit zwischen den einzelnen Wissensträgern und Key Playern im wissenschaftlichen Kommunikationsprozess. Der zyklische Ablauf der Knowledge-enhancing Helix gewährleistet nicht nur das Erfüllen der Kriterien für gutes Lernen, sondern bewirkt ebenso die konsequente, systematische Einbindung und die Stärkung der Position der Bibliotheken in der Hochschulausbildung. Bibliotheken besitzen wertvolle Ressourcen und Kompetenzen bezüglich einer optimalen Informations- und Wissensbeschaffung und der Churer Ansatz ermöglicht es, diese Fachkompetenzen gewinnbringend in den gesamten Prozess der Wissenskommunikation einzubinden. Das gesamte bibliothekarische Fachwissen wird systematisch weitervermittelt und gleichzeitig werden neue Synergien zwischen allen Beteiligten an der Wissenschaftskommunikation begünstigt und gefördert.

Literatur

Abfalterer, E. (2007): Foren, Wikis, Weblogs und Chat im Unterricht, Werner Hülsbusch, Innsbruck.

Alley, L.; Jansak, K. E. (2001): Ten keys to quality and assessment in online learning. In: Journal of Interactive Instruction Development, 13(3), 3-18.

Böller, N.; Herget, J.; Hierl, S. (2007): Knowledge-enhancing Helix: Herausbildung und Entwicklung von Informations- und Medienkompetenz durch systemgestütztes kollaboratives Arbeiten in der universitären Ausbildung. Eine Fallstudie. Beitrag auf der Konferenz IFLA 2007, Durban.

URL: http://www.ifla.org/IV/ifla73/papers/133-Boeller_Herget_Hierl-de.pdf

[zuletzt besucht am 16.8.2007]

Brändli, L. (2007): Gesucht – gefunden? Optimierung der Informationssuche von Studierenden in wissenschaftlichen Bibliotheken. In: Herget, J.; Hierl, S. (Hrsg.): Churer Schriften zur Informationswissenschaft – Schrift 21, Chur.

URL:

http://www.informationswissenschaft.ch/fileadmin/uploads/pdf/csi/CSI__21_Br_ndli_FINAL.pdf, [zuletzt besucht am 16.8.2007]

Chau, T.; Maurer, F. (2005): A case study of wiki-based experience repository at a medium-sized software company. In: International Conference on Knowledge Capture, Canada, S. 185-186.

Danowski, P.; Heller, L. (2006): Bibliothek 2.0 - Die Bibliothek der Zukunft? In: Bibliotheksdienst, 40, 1250–1271.

URL: http://www.zlb.de/aktivitaeten/bd_neu/heftinhalte2006/DigitaleBib011106.pdf [zuletzt besucht am 16.8.2007]

Fresen, J.W. & Boyd, L.G. (2005) "Caught in the web of quality" International Journal of Educational Development, 25, pp. 317-331.

Friedlander, A. (2002): Dimensions and Use of the Scholarly Information Environment. Introduction to a Data Set Assembled by the Digital Library Federation and Outsell, Inc. Digital Library Federation and Council on Library and Information Resources, Washington, D.C.

URL: <http://www.clir.org/PUBS/reports/pub110/contents.html>, [zuletzt besucht am 16.8.2007]

Hierl, S.; Bauer, L.; Boeller, N.; Herget, J. (2007): DIAMOND: Didactical Approach for Media Competence Development in a Blended Learning Environment, Proceedings for ECEL 2007, Kopenhagen.

Himpsl, K. (2007): Wikis im Blended Learning. Ein Werkstattbericht, Werner Hülsbusch, Innsbruck.

Lux, C.; Sühl-Strohmeier, W. (2004): Teaching Library in Deutschland. Vermittlung von Informations- und Medienkompetenz als Kernaufgabe für Öffentliche und Wissenschaftliche Bibliotheken. Dinges&Frick, Wiesbaden

MediaWiki (2007): Social Software

URL: www.mediawiki.org [zuletzt besucht am 16.8.2007]

Mönnich, M. (2006): Soziale Kompetenz in der postindustriellen Gesellschaft. Einschätzung ihrer praktischen Relevanz und gesellschaftlichen Funktion, Norderstedt, Books on Demand GmbH.

Reinmann-Rothmeier, Gabi (2001): Wissen managen: Das Münchener Modell (Forschungsbericht Nr. 131). Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, Ludwig-Maximilians-Universität, München.

URL <http://epub.ub.uni-muenchen.de/archive/00000239/> [zuletzt besucht 15.08.2007]

Wagner F.; Kleinberger Günther. U. (2004): Was ist neu an den Kompetenzen für neue Medien? In: Wagner F.; Kleinberger Günther U. (Hrsg.): Neue Medien - Neue Kompetenzen? Texte produzieren und rezipieren im Zeitalter digitaler Medien. Bonner Beiträge zur Medienwissenschaft, Band 3, Lang, Frankfurt.

WEB 2.0: Beispiele aus der Praxis

Wissenschaftskommunikation in der Entwicklungsforschung/Entwicklungszusammenarbeit Web2.0 und Communities of Practice - ein Beitrag aus der Praxis

Susanne von Itter

Der Beitrag beschäftigt sich mit der Wissenschaftskommunikation in der Entwicklungsländerforschung/Regionalstudien (Asien, Afrika, Lateinamerika) und Entwicklungszusammenarbeit. Es werden folgende Fragen aufgeworfen und beantwortet: Welche Rolle spielt das Web2.0 bei der Kommunikationsarbeit europäischer Entwicklungsforschungsorganisationen? Welche Anwendungen können zum Publizieren und Kommunizieren von Forschungsergebnissen genutzt werden? Welche „Communities of Practice“ haben sich im europäischen Raum in den letzten drei bis vier Jahren herausgebildet? Welche Zielgruppen werden angesprochen und neu erschlossen? Der Beitrag stellt verschiedene Projekte im Bereich Web2.0 und Entwicklungsforschungseinrichtungen vor und skizziert anhand von Praxisbeispielen Möglichkeiten und Grenzen bei der Anwendung von social software, Blogs und RSS-Diensten. Vorgestellt werden europäische Initiativen wie EADI Portal, Focuss.eu, Euforic.org, EL@ND (European Libraries Network for Development).

This article deals with research communication in the context of development research and area studies (Africa, Asia, Latin America) and development co-operation and answers the following questions: What role does the Web2.0 play with regard to the communication of Europe-based development research organisations? What technologies can be used to publish and disseminate research results? What communities of practice have emerged during the last three to four years? What target groups will be reached? The article presents various projects in development research organisations using web 2.0 and demonstrates the possibilities and limitations of social software, blogs and RSS feeds. European initiatives like EADI portal, Focuss.eu, Euforic and EL@ND (European Libraries Network for Development) will be presented.

Entwicklungsforschung in Europa und Vernetzung von Forschungseinrichtungen

Im Europäischen Dachverband der Entwicklungsforschungsinstitutionen EADI haben sich 150 Institutionen der Entwicklungsländerforschung und Regionalstudien zusammengeschlossen. Der Verband besteht seit 30 Jahren und arbeitet inter- und multidisziplinär. Seine Hauptaufgaben bestehen in der Vernetzung von Institutionen und Wissenschaftlern europa- und weltweit, der Förderung des Dialogs und Austausches der verschiedenen Disziplinen der Entwicklungsforschung, und der Kommunikation von Forschungsergebnissen innerhalb des Netzwerkes, an die

interessierte (Fach)Öffentlichkeit und Politik. EADI hat seinen Sitz seit dem Jahr 2000 in Bonn.

Definitionen von Entwicklungsforschung unterscheiden sich je nach wissenschaftstheoretischer Position und politischem Hintergrund. So zitiert das Paulo Freire Zentrum in Wien in seinem Internetdossier „Entwicklungsforschung 2007“ die Internet-Enzyklopädie Wikipedia, um diesen facettenreichen Forschungsgegenstand zu beschreiben (Paulo Freire Zentrum für transdisziplinäre Entwicklungsforschung 2007). "Unter Entwicklungsforschung ist die wissenschaftliche Beschäftigung mit den Ursachen, Aspekten, Kennzeichen und Folgen von Entwicklung und Unterentwicklung zu verstehen. Mit dem aus ihr resultierenden Erkenntnisgewinn soll ein Beitrag zum Verständnis der Probleme von Entwicklungsländern und zu deren Lösung geleistet werden."¹ . Verschiedene universitäre Disziplinen in Deutschland haben diese Aspekte in ihre Studiengänge integriert, wie etwa die Politikwissenschaften, Internationale Beziehungen, Wirtschaftswissenschaften, Gesellschafts- und Kulturwissenschaften, Agrarwissenschaften, Geographie, Regionalwissenschaften. In England und den USA gibt es klar umrissene Curricula für die sog. „development studies“ (Sumner 2006).

Für die Kommunikationsarbeit eines Verbandes wie EADI bedeutet dies, Forschungsergebnisse, die aus unterschiedlichen Disziplinen stammen, in einer kohärenten Weise darzustellen. Wichtigste Medien sind hierbei die Website und die entsprechenden elektronischen Newsletter, neben den traditionellen Medien, wie eine wissenschaftliche Zeitschrift und eine Buchserie² . Strategische Leitfragen sind derzeit: Wie organisieren wir unsere im Verband versammelte Expertise, mit dem Ziel, Ergebnisse besser zu kommunizieren? Wie kommunizieren wir diese Ergebnisse, so dass politische Entscheidungen in komplexen Gesellschaften in der Gesamtheit ihrer Facetten vernünftig gestaltet werden können (vgl. hierzu auch Ihne 2002)? Eine dritte Leitfrage ist: Wie kommunizieren wir mit Partnern in Nord und Süd, Ost und West und wie kommunizieren wir mit Wissenschaftlern in Entwicklungsländern, wie stellen wir das Wissen unserer Mitglieder zur Verfügung?

Der European Consensus on Development (Council of the European Union 2005/EC2006/C46/01) fordert, dass die EU die internationale Debatte über Entwicklung fördern und Best-practice Beispiele bekannt machen solle. Zudem brachte die Europäische Kommission im März 2006 eine Kommunikation „EU Aid: Delivering more, better and faster“ heraus. Hier wird die Notwendigkeit betont, Forschungsergebnisse, die in Europa erstellt werden, besser zu vernetzen und vor allem an die Entscheidungsträger der Entwicklungszusammenarbeit und in der internationalen Politik zu kommunizieren, damit sie in die globale Debatte einfließen können und letztendlich dazu beitragen, globale Probleme zu lösen. Ähnliche

¹ siehe auch Nuscheler 2004 für die verschiedenen Entwicklungstheorien

² European Journal of Development Research, EJDR, 19 Jahrgänge, 4 Ausgaben pro Jahr

Diskussionen finden zur Zeit auch in anderen wissenschaftlichen Netzwerken statt (Dabelko 2006, S. 1).

Die Rolle des Web2.0 bei der Kommunikationsarbeit europäischer Entwicklungsforschungsorganisationen und -netzwerke

Die technischen Komponenten des Web2.0 werden zunehmend von Entwicklungsforschungseinrichtungen für die Kommunikation von Forschungsergebnissen genutzt. Hierbei sind insbesondere RSS-Dienste, Blogsoftware und social bookmarking services zu nennen. Ein Viertel der in EADI vernetzten Institute bieten mittlerweile einen RSS Service an³. Gute Beispiele sind die Internetseiten des Institute of Social Studies in Den Haag⁴, oder die des Overseas Development Institute in London⁵, und des Informationsdienstes ELDIS⁶. Es existieren bisher noch keine systematischen Evaluierungen im Verband darüber, in welcher Weise und wie oft diese Instrumente von Besuchern dieser Websites genutzt werden. Mit seiner neuen Webpräsenz wird ELDIS, ein Informationsservice des Institute of Development Studies in Brighton, eine systematische Evaluierung für seine Website durchführen, so dass auf diesen Erkenntnissen aufgebaut werden kann. Web2.0 zeichnet sich jedoch auch durch den sog. „user generated content“ aus. Diese Möglichkeit wird zur Zeit in einigen Projekten genutzt, vor allem zur Konferenzdokumentation.

Netzwerke wie EADI und auch Euforic (Europe's Forum on International Cooperation) arbeiten dezentral und deren Mitglieder und Organisationseinheiten wie beispielsweise Arbeitsgruppen sind räumlich voneinander getrennt. Sie haben schmale Budgets, wenig Personal und müssen große Anstrengungen unternehmen, um Projektmittel zu akquirieren. Dies sind die wichtigsten Gründe für die Wahl von allgemein verfügbaren, nicht-kommerziellen Web2.0 Technologien. Bestimmte Formen der Kommunikation wären ohne das Web 2.0 für diese Art von Organisationen überhaupt nicht möglich geworden oder nur mit großen Anstrengungen im Bereich des Fundraising. Andererseits steigen die Erwartungen an einen wissenschaftlichen Verband, bestimmte Informationsangebote bereitzuhalten. Zudem zeigte eine Nutzerbefragung der Webdienste von EADI im Frühjahr 2007, dass die Nutzer neben dem starken Interesse an Forschungsergebnissen auch an der interaktiven Vernetzung und Kommunikation über Ergebnisse interessiert ist. Vorteilhaft ist somit bei dem Einsatz von Web2.0 Technologien der geringe Mittelaufwand, da Entwicklungskosten entfallen.

Entwicklungsforschungsinstitutionen, besonders im anglophonen Raum, investieren zunehmend strategisch in die Kommunikationsarbeit. So hat das englische DFID (UK

³ eine Auflistung befindet sich unter www.eadi.org.

⁴ www.iss.nl

⁵ www.odi.org.uk

⁶ www.eldis.ids.ac.uk

Department for International Development), eine dem deutschen Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung vergleichbare Institution, in seinen Förderrichtlinien für Forschungsvorhaben festgelegt, dass ein Teil des beantragten Projektbudgets (mindestens 10%) für die Kommunikation der Forschungsergebnisse aufzuwenden ist. (Barnard et al 2007, S. 33). Diese Auflage hat zur Folge, dass Institutionen sich diesen Anforderungen anpassen, so hat etwa das Overseas Development Institute in London jeder Organisationseinheit einen/e Experten/in für Kommunikation zugeordnet.

Communities of Practice in Europa

In den letzten vier bis fünf Jahren ist ein Trend festzustellen, dass einzelne Netzwerke, die sich mit Entwicklungspolitik und insbesondere der Wissenschaftskommunikation/Wissensmanagement in der Entwicklungszusammenarbeit widmen, die neuen Möglichkeiten, die die Anwendungen des Web2.0 bieten, nutzen, anpassen, weiterentwickeln und Anwender schulen. Dieses sind neben EADI die europäische Kooperative Euforic, Europe's Forum on International Co-operation⁷. Diese hat in den letzten zwei Jahren eine „Community of Practice“ aufgebaut, um die Anwendungen des Web2.0 unter den Mitgliedern zu verbreiten, mit dem Ziel kollaborativ arbeiten zu können. Weitere Initiativen in diesem Bereich ist das Netzwerk KM4Dev, eine Community von internationalen Entwicklungsexperten, die sich im Bereich Wissensmanagement spezialisieren und eine eigene Zeitschrift herausgeben. Die Information Management Working Group (IMWG) von EADI, und das European Libraries Network for Development EL@ND (ein von EADI geführtes Konsortium) sind weitere Beispiele. Die Jahresversammlung der Arbeitsgruppe mit circa 50-60 Teilnehmern aus den Bereichen Bibliothekswesen und Kommunikation/Informationsmanagement widmete sich 2006 dem Thema e-libraries und 2007 steht das Thema Library2.0 auf der Tagesordnung⁸. Hier werden auch Trainingseinheiten zu Web2.0 Anwendungen angeboten. Somit erfüllt diese europäische Arbeitsgruppe eine wichtige Multiplikatorenfunktion. Alle Initiativen treffen sich regelmäßig zum Austausch. Die in diesen Netzwerken aktiven Personen sind wiederum in disziplinär orientierten Netzwerken integriert, so dass hier ein reger europaweiter Austausch zwischen den verschiedenen Fachdisziplinen und den Querschnittsthemen wie Kommunikation oder Dokumentation/Bibliothekswesen erreicht wird.

Anwendungen zum Publizieren und Kommunizieren von Forschungsergebnissen mit Praxisbeispielen

Weblogs

Insecurityforum.org war ein Konferenz-Weblog, der 2005 im Zuge der 11. Generalkonferenz von EADI in Bonn eingesetzt wurde. Wissenschaftler, die nicht die

⁷ www.euforic.org

⁸ s. aktuelles Programm unter http://www.iss.nl/conferences/eadi_imwg_conference_2007

Möglichkeit hatten, an der Konferenz teilzunehmen, wurden eingeladen, die Ergebnisse der Panels und Workshops zu kommentieren. Dies erforderte eine zeitnahe Berichterstattung von der Konferenz. Auch bei der Konferenz anwesende Journalisten beteiligten sich an dem Weblog. Insgesamt wurde das Angebot gut angenommen. Das Overseas Development Institute führt einen Weblog, um Stellungnahmen und Meinungen zu aktuellen entwicklungspolitischen Themen aus dem Institut zu publizieren. Sowohl die wissenschaftlichen Mitarbeiter als auch die Leitungsebene sind dazu aufgerufen, in regelmäßigen Abständen kurze Stellungnahmen zu veröffentlichen. Zielgruppen sind hier die allgemeine interessierte (Fach)öffentlichkeit, aber auch politische Entscheidungsträger und NGOs. Euforic setzt Blogs ein, um Tagungen und Konferenzen zu dokumentieren. Um Redebeiträge nachzuhören, werden dort auch Podcasts eingesetzt. So ist eine virtuelle Teilnahme auch für Teilnehmer aus Entwicklungsländern möglich.

RSS Newsfeeds

RSS-Dienste erlauben die schnelle und zielgerichtete Verbreitung von zumeist schnelllebigem Informationen. Sie stellen eine neue Kombination von Information und Kommunikation dar. Newsfeeds erlauben auch die Kombination von Informationen aus unterschiedlicher Quelle (Addison 2006, S. 624). Im Zuge der Nutzung moderner Content Management Systeme verbreiten sich RSS Dienste immer mehr, so auch bei Institutionen der Entwicklungsforschung. Beispielhaft seien die sehr informativen Dienste vom Overseas Development Institute, vom European Centre for Development Policy Management in Maastricht, vom Royal Tropical Institute in Amsterdam und vom OECD Development Centre in Paris genannt. Euforic listet als besonderen Service Newsfeeds von Mitgliedern auf, so auch EADI⁹.

Google und Co

Focuss.eu¹⁰ ist eine Initiative von mittlerweile 24 Organisationen in der Entwicklungsforschung und Entwicklungszusammenarbeit. Focuss.eu ist eine fachbezogene Suchmaschine, die Ressourcen durchsucht, die von den mitarbeitenden Initiativenmitgliedern ausgewählt und bereitgestellt wurden. So durchsucht diese Suchmaschine beispielsweise die Websites von wichtigen entwicklungsorientierten Bibliotheken, aber auch von einigen Nichtregierungsorganisationen. Die Suchmaschine basiert auf Google Search. Das zweite interaktive Element von Focuss.eu ist die Möglichkeit, interessante Quellen zu der Suchmaschine mittels Google Marker hinzuzufügen. Dieses collaborative tagging ist erst in der Anfangsphase, und soll weiter ausgedehnt werden. Das Resultat ist schon jetzt eine spezialisierte Suchmaschine, die weitaus schneller qualitätsgeprüfte Ergebnisse liefert, als wenn man eine Suche mit der allgemeinen Google-Suchmaschine durchführt. Das Angebot richtet sich an Studenten, Wissenschaftler und Praktiker.

⁹ s. http://www.euforic.org/detail_page.phtml?page=feedfinder_members, www.eadi.org

¹⁰ www.focuss.eu

Social bookmarking

Sowohl EADI als auch Euforic nutzten den Dienst del.icio.us, um ihre Webseiten zu aktualisieren. Das European Centre for Development Policy Management (ECDPM) setzt ebenfalls del.icio.us ein, um mehrere Informationsportale zur Europäischen Afrikapolitik zu speisen¹¹. EADI und Euforic, sowie ECDPM nutzen miteinander abgestimmt einzelne tags, um gemeinsam Ressourcen auf ihren Websites darzustellen. Diese Art des kollaborativen Arbeitens in objektzentrierten sozialen Netzwerken (Danowski und Heller, 2006, S. 3-4), wo das öffentliche Entdecken und Erschließen von Informationsressourcen im Mittelpunkt steht, ist erst in Ansätzen sichtbar, soll aber längerfristig und systematisch ausgebaut werden, um die Potentiale, die diese Form des Arbeitens bietet, auszuschöpfen.

Projekte im Bereich Web2.0

EADI Portal: Europäische Expertise im Bereich Entwicklungsforschung sichtbar machen

Das Projekt DevExpertise/DevResearch hat zum Ziel, Ergebnisse von Forschungsprojekten europaweit zu dokumentieren und sichtbar zu machen, und einen Austausch hierüber zu fördern. So können die Ergebnisse sowohl von Praktikern genutzt werden als auch als Politikempfehlungen in den politischen Diskurs einfließen. Als Basis für die Forschungsdokumentation dient ein umfangreicher del.icio.us account, der nach Themen, Ländern, Institutionen gegliedert ist und auch eine Archivfunktion bietet. Zusätzlich wird eine Blogsoftware eingesetzt, um Forschungsergebnisse (Research Highlights) vorzustellen, die dann diskutiert werden können. Hiermit ist eine direkte Interaktion möglich. Weiterhin werden Newsfeeds aus eigenen del.icio.us accounts genutzt, um Newsfeeds auf der Portalstartseite zu positionieren, und auch Newsfeeds, die von Mitgliedern bereitgestellt werden, integriert. Eine angepasste Google-Suchmaschine bietet weitere Funktionen, um nach Themen, Ländern zu suchen. In einem Bereich, der nur für Mitglieder zugänglich ist, werden weitere Serviceleistungen zur Verfügung gestellt, so dass das Portal und das Redaktionsteam, das dahinter steht, die Rolle eines information brokerage service einnimmt. Dieses Projekt wird von mehreren Partnern gefördert, z.B. dem niederländischen Außenministerium und dem Research Council of Norway. Zielgruppen dieses Projektes sind Wissenschaftler, Forschungsförderungseinrichtungen und nationale und europäische Entscheidungsträger aus Politik sowie Administration. Das Projekt befindet sich zur Zeit in der Pilotphase.

European Libraries Network for Development EL@ND

Das European Libraries Network for Development¹² ist ein Projekt innerhalb der EADI Arbeitsgruppe zum Thema Information Management (IMWG, s.o.). 10 auf

¹¹ www.ecdpm.org

¹² www.eadi.org/eland

Entwicklungsforschung spezialisierte, online verfügbare Kataloge können hier mit einer Suchfunktion durchsucht werden. Bibliotheken, die sich auf das Thema Entwicklungsforschung und verwandte Themen spezialisiert haben, sind oft klein und daher weniger sichtbar oder oft auch überhaupt nicht eingebunden in größere Projekte wie beispielsweise Vascoda¹³ in Deutschland. Auch europaweit sind sie weniger vernetzt. Hier setzt dieses Projekt an. Zum einen wird eine Plattform geboten, auf der von spezialisierten Bibliothekaren empfohlene Literatur dargestellt wird, und andererseits werden auch Bibliothekare ermuntert, neue Services im Bereich Web2.0 zu benutzen. Weiterhin verschafft diese Website Zugang zu Open Access Zeitschriften und zu digitalen Repositorien im Bereich Entwicklungsforschung.

Emergent Issues in Information and Knowledge Management (IKM) and International Development

Dieses auf fünf Jahre angelegte Forschungsvorhaben von EADI hat das Ziel, Wissensmanagementprozesse im Entwicklungsbereich zu untersuchen. Eine kontinuierliche Diskussion mit Wissenschaftlern, Praktikern, Politikern und Informations- und Wissensmanagern soll in einen Forschungsplan münden, der eine Debatte und Reflexion über das Management und den Gebrauch von Wissen in der Entwicklungszusammenarbeit erlaubt und Innovationen initiiert. Die Verbindungen zwischen Wissenschaft, Politik und Umsetzung sollen untersucht werden, insbesondere im Hinblick auf die Rolle der Kommunikation in Forschungsprogrammen. Die Rolle der Anwendung von Wissen im Entwicklungsprozess soll untersucht werden, insbesondere im Hinblick auf das Wissen in Entwicklungsländern. Ein eigener Bereich in diesem Forschungsprogramm ist die Untersuchung der Anwendung von neuen Handwerkszeugen für die Wissensmanagement- und Kommunikationsarbeit. Es wird erwartet, dass Tools aus dem Web2.0 -Bereich eingesetzt und untersucht werden, also hier ein Erkenntnisgewinn über den Einsatz von Web2.0 Technologien im Entwicklungskontext zu erwarten ist. Eine starke Einbindung von Partnernetzwerken aus dem Süden ist vorhanden¹⁴.

Bei den dargestellten Projekten geht der Trend weg von reinen Informationsplattformen, die auf das Aufbereiten und Sammeln von Informationen spezialisiert sind hin zu Plattformen, wo Informationen ausgetauscht werden, durch Besucher ergänzt und kommentiert werden und somit eine Verbindung zwischen Menschen/Autoren und Informationen hergestellt wird.

Grenzen

Wenn sich Organisationen auf Web2.0 Technologien verlassen, müssen sie sich bewusst sein, dass sie trotz der geschilderten Vorteile auf Services angewiesen sind und von deren Funktionieren abhängig bleiben, die sie nicht vollständig unter

¹³ www.vascoda.de ist ein Internetportal für wissenschaftliche Information

¹⁴ Zur Zeit wird an einer Website für das Projekt gearbeitet, siehe www.ikmemergent.net

Kontrolle haben (Addison 2006, S. 624). Serverausfälle oder Verlangsamungen z.B. bei den social bookmarking services sind keine Seltenheit.

Die Rolle von „Knowledge Intermediaries“

Innerhalb von EADI gibt es zur Zeit noch wenige Wissenschaftler, die bsp. social bookmarking services nutzen. Gründe sind die geringe Bekanntheit, aber auch der Zeitaufwand, den die Pflege erfordert, z.B. bei Blogs. Es gibt wenige Wissenschaftler mit eigenen aktuellen Blogs. Entscheidet sich eine Organisation dafür, diese Web2.0-Tools als Wissensmanagement-Instrumente einzusetzen, erfordert dies auch ein Umdenken bei Arbeitsweisen, Strukturen und Zeiteinsatz bei Wissenschaftlern, was eine Frage der Organisationsentwicklung ist (vgl. hierzu auch Peters und Stock 2007, Schnitzer und Zinnbauer, 2007). Die Integration von Kompetenzen zur Informationsverarbeitung und kollaborativem Arbeiten in die Curricula von Studiengängen wäre ein Ansatz auf der Ebene der Lehre.

Knowledge Intermediaries und von ihnen angebotenen Dienste versorgen Wissenschaftler, Politik und Öffentlichkeit mit aktuellen Informationen über Entwicklungszusammenarbeit und Entwicklungsforschung und werden daher auch in Zukunft diese Lücke schließen. ELDIS, EADI Portal, oder auch das deutschsprachige Portal EPO¹⁵ (Entwicklungspolitik online) leisten hier wichtige Dienste bei der Aufbereitung, Verfügbarmachung und Zusammenstellung von Informationen. Sie bieten zudem Unabhängigkeit und Kontinuität in der Kommunikation auch über das Projektende von Forschungsprojekten hinaus (vgl. auch Barnard et al 2007, S. 15)¹⁶. Eine Herausforderung wird auch die Multidisziplinarität dieser Informationsportale bleiben, das heißt, Informationen aus unterschiedlichen Quellen so auszuwählen und zu kombinieren, so dass sie ein breites Spektrum abbilden, und auch Informationen von Organisationen mit einzubeziehen, die nicht so starke Kapazitäten in der Öffentlichkeitsarbeit haben. Dies erfordert eine redaktionelle Recherche nach diesen Informationen, da sie nicht automatisch per Newsfeed geliefert werden.

Qualitätssicherung

Eine weitere Fragestellung, die immer wieder aufkommt, ist die Frage der Qualitätssicherung. Hier ist eine beständige Kontrolle der Webinhalte und die Einführung geeigneter peer review Mechanismen unerlässlich.

Archivfunktion, Schnellebigkeit der Webinhalte

Erfahrungsgemäß veralten Links und Nachrichten schnell. Wenn man eine Internetseite mit Newsfeeds speist, die jeweils nur die fünf neuesten Nachrichten dynamisch darstellen, entsteht das Problem, dass eine Website die Archivfunktion verlieren kann. Auf ältere Ressourcen kann dann nur noch mit Mühe zugegriffen werden. Daher wurde bei EADI Portal in del.icio.us eine Verschlagwortung nach

¹⁵ www.epo.de

¹⁶ Eine Auflistung der derzeit aktiven Knowledge Intermediaries in der Entwicklungszusammenarbeit befindet sich unter <http://www.ids.ac.uk/research-comms>

Erscheinungsdatum der Nachricht/Publikation zur Archivierung eingeführt. Hier wird zur Zeit noch mit einer Taxonomie gearbeitet, die sich noch nicht an traditionellen Dokumentations- und Bibliothekssystemen orientiert, sondern eher an pragmatischen Gesichtspunkten. Im Zuge der Weiterentwicklung des Projektes wird hier jedoch die Zusammenarbeit mit den im Netzwerk aktiven Bibliothekaren der Mitgliedsinstitute angestrebt.

Die digitale Kluft zwischen Nord und Süd

Wie erreicht man die Menschen in Regionen mit schlechter Anbindung an das Internet, wie es beispielsweise in vielen Regionen Afrikas der Fall ist? Auch im Kontext der Open-Access Diskussionen ist dies eine wichtige Frage, damit die Wissenschaftler, Studenten und Praktiker in Entwicklungsländern überhaupt von all diesen Bemühungen, Zugang zu Wissen zu schaffen, profitieren können. De Beer nennt einige vielversprechende Initiativen, die den Zugang zum Internet von Universitäten fördern, wie bspw. die National Research and Education Networks (de Beer 2007). Dieser Punkt der fehlenden ICT Infrastruktur bleibt jedoch, nach wie vor eine große Herausforderung. Fairer Zugang sowohl zu Wissen als auch Technologie muss garantiert werden (Priestley 2007).

Ausblick

Das Thema Web2.0 wird zur Zeit in der Entwicklungsgemeinschaft stark diskutiert. So findet im September in Rom die Konferenz „Web2. for Dev: Participatory Web for Development“ statt, organisiert von mehreren Mitgliedern und Partnern von EADI, und der Welternährungsorganisation FAO. Auch im September 2007 findet das 31. Jahrestreffen der EADI Information Management Working Group statt. An der Universität Maastricht wird am 28. November 2007 die Konferenz „The Changing Role of Knowledge for Development: Implications for Higher Education and Research Institutions“ stattfinden.

Das Web2.0 bietet trotz der geschilderten Grenzen zahlreiche Möglichkeiten, die in Entwicklungsforschungseinrichtungen zur Kommunikation von Forschungsergebnissen genutzt werden können. Hier setzen die geschilderten Projekte an: Informationen schnell und gezielt so zu verbreiten, so dass wissenschaftliche Ergebnisse aus Entwicklungsforschungseinrichtung im Norden und im Süden die erforderliche Wirkung erzielen können, d.h. umgesetzt werden und in die politischen Entscheidungsprozesse einfließen.

Literatur

Alle Weblinks zum Artikel: <http://del.icio.us/wissskom/wissskom2007>

Addison; C. (2006): Web2.0: a new chapter in development in practice? In: Development in Practice, Vol. 16, No. 6, November 2006, S. 623-627

Barnard, G. Carlile, L. und D. Basu Ray (2007): Maximising the Impact of Development Research. Brighton (<http://www.ids.ac.uk/research-comms>)

Council of the European Union (2005): Joint Statement by the Council and the representatives of the Governments of the Member States meeting within the Council, the European Parliament and the Commission on European Union Development Policy: The European Consensus on Development, 14820/05.

Dabelko, G. (2006): Turning Up the Volume on the Science-Practice Interface in Global Change. In: IHDP Update 3/4/2006. (http://www.ihdp.uni-bonn.de/html/publications/update/pdf-files/IHDPUpdatePolicyRelevance3_06.pdf)

Danowski, P. und L. Heller (2006): Bibliothek 2.0: Die Zukunft der Bibliothek? In: Eprints in Library and Information Science (<http://eprints.rclis.org/archive/00007618/>)

de Beer, J.A. (2007): Beyond First-Generation Open Access for Africa. In: EADI Newsletter 2007, S. 10-11 (<http://www.eadi.org>)

Europäische Kommission (2006): Communication from the Commission: EU Aid: Delivering more, better and faster (02.03.2006) (http://ec.europa.eu/commission_barroso/michel/Policy/key_documents/index_en.html)

Ihne, H. (2002): Entwicklungsforschung und Politik
Warum lassen sich die Politiker zu wenig beraten? In: Entwicklung und Zusammenarbeit, Nr. 3, März 2002, S. 68 - 69

Nuscheler, F. (2004): Lern- und Arbeitsbuch Entwicklungspolitik. Bonn

Paulo Freire Zentrum für transdisziplinäre Entwicklungsforschung und -bildung (2007): Schwerpunkt Entwicklungsforschung 2007 (<http://www.pfz.at/index.php>)

Peters, I. und W.G. Stock (2007): Web 2.0 im Unternehmen. In: Wissensmanagement, H. 4, S. 22-25

Powell, M. (2006): Which knowledge? Whose reality? An overview of knowledge used in the development sector. In: Development and Practice, Vol. 16 No. 6, S. 518-532

Priestley, C. (2007): Access to and Dissemination of Research Results in Developing Countries and Emerging Economies. In: EADI Newsletter 1-2007, S. 8-9 (<http://www.eadi.org>)

Schnitzer, T. und Zinnbauer, M. (2007): Wikis und Blogs: Die Bottom-Up-Revolution im Wissensmanagement. In: Wissensmanagement H. 5, S. 18-20

Sumner, A. (2006): What is Development Studies? In: Development in Practice 16 (6), 644-650.

Glossar

Blog: Blogs oder Weblogs sind digitale Tagebücher, die im World Wide Web veröffentlicht werden. Sie werden von Unternehmen und Organisationen als sogenannte „Corporate Blogs“ als Kommunikationsinstrument genutzt.

RSS: Steht für Really Simple Syndication und bezeichnet ein XML-basiertes Format zum plattformunabhängigen Datenaustausch. RSS-Feeds oder Newsfeeds werden mit Feed-Readern abonniert und eignen sich besonders für eine schnelle und zielgerichtete Distribution von schnelllebigem Webinhalten. Darüber hinaus kann man fremde RSS-Feeds auch in seine Internetseite integrieren und somit mehrere Informationsströme auf einer einzigen Seite zusammenführen.

Social bookmarking: Social bookmarking Dienste ermöglichen das gemeinschaftliche Abspeichern und Wiederfinden von Web-Adressen, die von den Benutzern zudem frei verschlagwortet („getaggt“) werden können.

Web2.0: Gesamtheit aller web-basierten Tools und Dienste, die kollaborativ verwendet werden (können) und den Benutzer in die Erstellung von Webinhalten einbeziehen. Der Begriff wurde 2005 von Tim O'Reilly geprägt (s. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>)

Wissenschaftliche Informationsversorgung am modernen digitalen Arbeitsplatz am Bsp. des Corporate Information & Research Management der DaimlerChrysler AG

Steffen Leich-Nienhaus

Abstract deutsch

Die aktuellen Entwicklungen der web-basierten Kommunikation, beispielsweise im Umfeld des Web 2.0, prägen die Versorgung der Mitarbeiter mit wissenschaftlichen und anderen arbeitsrelevanten Informationen ähnlich stark wie das ursprüngliche Aufkommen der digitalen Medien:

- Verfügbarkeit nahezu sämtlicher Quellen der arbeitsbezogenen Informationen in digitaler Form
- Unterstützung der Informationsautonomie des Benutzers
- Verstärkter Bedarf nach Verfügbarkeit eines Rückkanals in den Medien um Meinungen und Strukturierung beizusteuern
- Veränderung der Suche-Techniken durch Web 2.0-Prinzipien (Social Search, Mashup, etc.) - Benutzer werden such- und navigationstechnisch versierter.
- Verstärkte Bewegung der Benutzer auch parallel zu den „üblichen“ Hauptmedien wie in Newsfeeds
- Zunahme des Angebots frei verfügbarer Angebote im Internet, die parallel zu den organisationsinternen Quellen genutzt werden
- Verschmelzung der arbeitsbezogenen Informationsbeschaffung mit arbeitsnaher Informationsbeschaffung, da alle Bereiche von den gleichen Medien abgedeckt werden

Aus diesen Entwicklungen ergeben sich neue Herausforderungen für eine organisationszentrale Informationsversorgung:

- Anpassung der internen Informationsversorgung an aktuelle Funktionsweisen des Internet, wie die Personalisierbarkeit von Inhalten
- Anpassung von Prozessen für die Nutzung des Rückkanals
- Erreichen des Benutzers in den neuen Medien und Schaffen von Aufmerksamkeit für organisationsinterne Informationen
- Integration von adäquaten externen und internen Angeboten

Während beispielsweise die Web 2.0-Techniken nicht immer sinnvoll für die relevanten organisationsinternen Prozesse sind und nicht unbedingt ihrer eigentlichen Funktion entsprechend direkt verwendet werden können, demonstrieren sie doch neue Möglichkeiten der Kommunikation und geben Anregung für Weiterentwicklungen. So verlagert sich die frühere Frage, wie der Transport der papiernen wissenschaftlichen Informationen an den Arbeitsplatz bewerkstelligt wird, dahingehend, wie der Benutzer nun in seiner digitalen Arbeitsumgebung und seinen eigenen (digitalen) Prozessen erreicht werden kann.

Der Vortrag behandelt verschiedene Effekte der neuen web-basierten Techniken aus Sicht der betrieblichen Informationsversorgung und gibt einen Einblick in hier umgesetzte Integration und Nutzung.

Abstract englisch

The current developments of web based communication, for example in the surrounding field of the Web 2.0, are influencing the supply of internal organisation members with scientific and other work relevant information just as strongly as the initial arising of the digital media:

- Availability of almost all sources of the work-referred information in digital form
 - Support of the informational autonomy of the user
 - Intensified need for availability of a backward channel in the media to contribute opinions and structuring
 - With the change of the search techniques by Web 2.0-techniques (Social Search, Mashup, etc.) - users become search and navigation-technically more experienced.
 - Increased usage of meta media by the users in parallel to the main media as in newsfeeds
 - Increased offer of freely available resources on the Internet, which are used in parallel to the organization internal sources
 - Fusion of the work referred provision of information with worknear provision of information, since all ranges are covered by the same media
- From these developments new challenges for a organization central information supply result:
- Adjustment of the internal information supply to current function modes of the Internet, like the personalization of contents
 - Adjustment of processes for the use of the backward channel
 - Reaching the user in the new meta media and generation of attention for organization internal information
 - Integration of adequate external and internal informational resources

While for example the Web 2.0-techniques are not always useful for the relevant internal processes and can not be used directly, they nevertheless demonstrate new possibilities of communication and give suggestion for advancements. Thus the earlier question of how the delivery of paper based scientific information to the employee is managed, is shifted to how the user can be reached now in its digital work environment and its own (digital) processes.

The lecture treats different effects of the new web based techniques from view of the internal information supply and gives an insight in practical integration and use.

1. Einordnung der Wissenschaftskommunikation in die Informationsorganisation im Unternehmen

Der Begriff Wissenschaftskommunikation lässt sich zunächst in zwei unterschiedlichen Ebenen betrachten: der eigentlichen inhaltlichen und der technischen Ebene.

1.1. Wissenschaftskommunikation inhaltlich

Die inhaltliche Wissenschaftskommunikation ist grundlegend die Vermittlung von wissenschaftlichen Erkenntnissen unter allen Interessierten, ihre Verständlichmachung und Verbreitung, die Schaffung von Verständnis und Akzeptanz in der Bevölkerung und somit insgesamt die Grundlage für den wissenschaftlichen Diskurs¹.

Aus Sicht der Forschung wären klassische Fragestellungen neben der Verbreitung an das entsprechende Fachpublikum auch das „Verkaufen“: Wie bringe ich meine Forschungsergebnisse in die Köpfe der Menschen, wie erringe ich Aufmerksamkeit? Wie werden die Ergebnisse von der breiten Leserschaft über die Massen- und andere Medien aufgenommen?

Die Wissenschaftskommunikation an sich dürfte von technischen Entwicklungen zunächst unberührt sein. Was die Forschung kommuniziert, ist medien- bzw. technikneutral. Die eigentliche Wissenschaftskommunikation steckt in den transportierten Inhalten selbst.

1.2. Wissenschaftskommunikation aus technischer Sicht

Diese inhaltliche Wissenschaftskommunikation wird ermöglicht durch die technische Unterstützung der Erzeugung wissenschaftlicher Ressourcen (wissenschaftlicher Erkenntnisprozess) und der Informationsversorgung durch die Bibliotheken. Grundlegende Funktionen sind hierbei Transport, Bereitstellung, Speicherung, Zugänglichmachung, Verwertung u.a. von wissenschaftlichen Ressourcen.

Durch technische Entwicklungen können sich neue Berührungspunkte mit dem aufnehmenden Leser ergeben. Neue Techniken erlauben der Wissenschaft, ihre Inhalte schneller und direkter zum Leser zu transportieren. Dies wiederum kann einen Einfluss auf die Entwicklung unterschiedlicher Stärken von Aufmerksamkeit haben. Durch neue Strukturierungsmöglichkeiten, Suchfunktionalitäten, Visualisierung und andere Techniken können Wissenstransporte besser organisiert ablaufen. So kann ein Leser den Inhalten gewogener sein, wenn er übersichtlich und schrittweise an News herangeführt wird, z.b. zunächst einen Teaser lesen kann, bevor er in die Details einsteigt.

1.3. Rolle der Fachinformationen in Organisationen

Oben genannte Funktionen kommen einer organisationsinternen Fachinformation zu. Sie lassen sich wie folgt in spezifische Aufgaben untergliedern²:

¹ Vgl. dazu auch [Wissenschaft im Dialog]

² Zu einer Definition von Fachinformationen vgl. auch [Kuhlen: Fachinformation] unter „Fachinformation zur Absicherung professionellen und öffentlichen Handelns“

- Organisation des gesamten Versorgungsbetriebs
- Beschaffung, Bereitstellung, Aufbereitung, Pflege und Aktualisierung wissenschaftlicher Ressourcen
- Gestaltung und Strukturierung von Informationsangeboten
- Filterung, Auswahl, Kanalisierung von externen Quellen nach Wirtschaftlichkeitsprinzipien und der Entsprechung interner Vorgaben (i.d.R. wird innerhalb einer Organisation eine eigenständige Informations- und Meinungswelt gepflegt)
- Erklärung und Kommentierung von Angeboten

Diese Grundfunktionen müssen oftmals unter den Maßgaben infrastruktureller Besonderheiten und anderer technischer Herausforderungen erfüllt werden:

- z.T. eingeschränkter Internet-Zugang in der Organisation
- Vorgaben für nutzbare Internet-Quellen
- Standardisierte Infrastruktur an Werkzeugen und Online-Systemen. Dem gegenüber steht eine Vielzahl neuer Internet-Technologien und Kommunikationstechniken, die evtl. intern abgebildet werden sollen.

Die Erfüllung der Aufgaben der Fachinformation innerhalb der Organisation enthält einige Funktionen, die auch von Systemen im Internet wahrgenommen werden, wie Gestaltung, Strukturierung, Filterung, Auswahl, Erklärung, Kommentierung verteilter, externer Inhalte.

Daneben konvergieren Technik für Informationsmanagement und Internet-Technik zunehmend. Bislang bezog sich diese Konvergenz hauptsächlich auf die Infrastruktur, also Transport- und Applikationsebene. Mittlerweile jedoch vollzieht das interne Informationsmanagement auch hinsichtlich der Anwendungsszenarien und Kommunikationsprinzipien eine Anpassung in Richtung der aktuellen Themen im Internet.

Somit ist die technische Seite der Wissenschaftskommunikation, der Wissenstransport, weiterhin verbunden mit den Entwicklungen im Internet.

2. Technische Trends und Entwicklungen

Im Mittelpunkt der Betrachtungen sollen hier nicht ausschließlich die aktuellen Web 2.0-Entwicklungen stehen. Ebenso wenig die bibliotheksorientierten Techniken im Rahmen der Library 2.0-Initiativen. Relevant ist hier der direkte oder indirekte Bezug zur wissenschaftlichen Informationsversorgung in Organisationen sowie die Kommunikationsprinzipien, die hinter diesen Systemen stehen. Diese sollen auf ihre Verwendungsmöglichkeit für eine Weiterentwicklung von bestehenden Kommunikationstechniken untersucht werden.

2.1. Verfügbarkeit von Quellen in digitaler Form

Mittlerweile ist, wo die Möglichkeit dazu besteht, die Umstellung der Lieferung von Quellen von gedruckter auf digitale Form quasi vollzogen, sämtliche arbeitsrelevanten Informationen sind in digitaler Form für den eigenen Arbeitsplatzrechner und

für beliebige Weiterverarbeitungen (Umwandlung in Sprache, Druck auf Papier oder andere digitale Formate) verfügbar.

Druckereien arbeiten mit digitalen Druckvorlagen, wodurch für sämtliche Schriftstücke digitale Ausgangsformate entstehen. Automatisiertes Scannen von großen Dokumentbeständen inkl. der automatischen Texterkennung sind inzwischen ausgereifte Techniken.

Die letzten noch in Papier verfügbaren Quellen, die bestehenden Bücher, werden wohl nach und nach in digitale Form umgesetzt, sei es von Google oder einem anderen Digitalisierungs- und Suchdienst.

Der stark verbreitete Umgang mit den digitalen Ressourcen und den damit verbundenen Vorteilen in Beschaffung und Handhabung, sowie der Flexibilität der Weiterverarbeitung, lassen kaum mehr eine Nutzung anderer Materialien zu.

2.2. Freie Verfügbarkeit von Angeboten

Für alle Inhaltsarten einer Fachinformation gibt es neben den intern verwendeten kostenpflichtigen und professionell zusammengestellten Ressourcen das freie, öffentlich zugängliche Pendant im Internet. Neben spezialisierten Quellen für enger fokussierte Themen existieren die übergreifenden Content-Container: Online-Zeitungen, Wikipedia, Google und Weblogs.

Darin enthalten sind auch zunehmend popularisierte wissenschaftliche Quellen. Die Internet-Inhalte, momentan noch entfernt von einer etablierten Open Access-Welt, sind bestimmt von Quellen aller Art mit jeglicher wissenschaftlicher Güte.

Demonstriert wird diese Entwicklung unter anderem durch zwei Phänomene, die bereits intensiv in der Literatur behandelt werden:

- Google als Einstiegspunkt³ : Jegliche Informationsbeschaffung wird über die Eingabe von Anfragen in ein Google-Suchfeld zumindest begonnen.
- Copy-Paste Wissenschaft⁴ : Motiviert durch breite Verfügbarkeit und leichte Auffindbarkeit werden bereits formulierte Inhalte in wissenschaftliche Publikationen integriert.

Aus der verstärkten Nutzung der frei verfügbaren Inhalte entstehen verschiedene Probleme für die interne Informationsversorgung: Schwierigkeiten beim Review der Qualität und der Verifizierung von Belegen und Quellen, evtl. mangelnde Vollständigkeit und Aktualität.

Gleichzeitig werden dabei auch für die interne Nutzung wertvolle seriöse Quellen erschlossen, die in die internen Informationsprozesse einbezogen werden können.

2.3. Verstärkte Informationsautonomie der Benutzer

Eine derartige Verfügbarkeit von Inhalten sorgt u.a. auch für eine verbesserte Informationsautonomie.⁵ Der Nutzer hat das Wissen um die Möglichkeit, sich mit Fachinformationen u.a. auch mehr oder weniger selbst versorgen zu können.

³ Vgl. [Palm: Google]

⁴ Vgl. [Weber: Copy-Paste]

⁵ Vgl. [Kuhlen: Informationsassistenten], S. 171 ff.

Er ist im Zuge der Entwicklung der digitalen Welt technisch versiert geworden und besitzt ein reichhaltiges Wissen über den Einsatz von digitalen Informationen zum optimalen eigenen Nutzen. Er hat einen gewissen Grad an Medienautonomie und Medieneffizienz entwickelt und arbeitet routiniert mit über Personalisierung angepassten Informationsquellen.

Neben dieser Entwicklung von Fähigkeiten stehen dem Benutzer mittlerweile auch sehr mächtige Werkzeuge zur Verfügung, die ihm sämtliche Möglichkeiten der Konfiguration und Filterung von Informationsströmen ermöglichen. Dadurch erreicht er eine gesteigerte Effizienz bei der Informationssuche und -aufnahme.

2.4. Nutzung eines Rückkanals

Teil der informatorischen Selbstbestimmung ist es auch, die genutzten Informationsquellen anhand von Rückkanal-Eingaben (Kommentare, Konfigurationen) anzupassen und im Sinne einer effizienten Nutzung zu optimieren.

Über solche Kanäle können Meinungen, Strukturierungen, Filterungen und Empfehlungen abgegeben werden. Möglich ist auch die Erstellung von Daten und Suchsystemen und deren Konfiguration nach dem Wiki-Prinzip. Motivation für die Nutzung solcher Rückkanäle ist meist die Optimierung der eigenen Kommunikation und Informationsversorgung bzw. ein Geben und Nehmen in Form des Austausches von Optimierungsarbeit.

2.5. Neue Suchtechniken

Neue Suchtechniken im Internet sind mit Blick auf die Bibliotheksentwicklung sehr interessant, da zum einen Suchfunktionalitäten eine immer zentralere Rolle in allen Arten von Anwendungssystemen einnehmen und zum anderen Suche die grundlegende Technik für die Informationsrecherche liefert.

Auffällig ist z.B. die aktuelle Orientierung einiger Hersteller von Suchsystemen:

So sieht „fast search“ Suchsysteme als neuen, höchsten und alle Anwendungen durchziehenden neuen „OSI-Layer“⁶. Dieser eher marketingbasierte Ansatz betont die grundlegende Relevanz von Suchtechniken für alle Arten von Anwendungen. Suche kann demnach eine Querschnittsfunktion für die Applikationsebene übernehmen.

Daneben sind folgende aktuelle Entwicklungen im Bereich der Suche-Technik relevant, von denen einige unter dem Stichwort Search 2.0 zusammengefasst werden:

- Consumerization: Einer der Trends nach Aussagen von Vertretern großer Suchsysteme ist die sog. Consumerization⁷ der Enterprise-Search⁸. Die einfach zu bedienenden und ergonomisch optimierten Internet-Suchen dienen als Vorbild für die Weiterentwicklung der Enterprise-Suchsysteme.

⁶ Siehe [fast: Enterprise Search], S. 20, „The New „OSI Layer Model““

⁷ Siehe [Gartner: Cosumerization]

⁸ Siehe [Google: Enterprise Search] und [Exalead: Human-Machine Interaction]

- Personal Search: Bsp. Eureka⁹: Eine persönliche Suche, die neben statischen Seiten auch andere Search-Engines einbindet.
- Bsp. Rollyo¹⁰: Der Benutzer kann sich ein Suchportal aus verschiedensten von der Community angelegten Such-Engines zusammenstellen. Browser-Plugins erlauben die Mitnahme des Rollyo-Portals bei der Informationssuche.
- Human Powered Search¹¹: Ausgangspunkt sind fehlendes Vertrauen in Precision und unnötig ausuferndes Recall der maschinellen Suchmaschinen. Redakteure oder die Community erstellen spezielle Ergebnisseiten oder optimieren Ergebnis-Rankings.
- Neue Techniken der Suche- und Ergebnisnavigation: Unterschiedliche Navigationswerkzeuge, wie facettierte Suche, grafische Darstellung von Ergebnisgruppen und anderes ergänzen aktuelle Suchsysteme.
- Flexibler Zugriff auf Suchfunktionen: Der Zugang zu den präferierten Such-Engines kann ständig im Browser mitgeführt werden.
- Suchergebnisse als News-Stream: Die Notwendigkeit der Ausführung wiederholter Suchanfragen zum gleichen Thema entfällt. Das System meldet sich bei Auffinden neuer Ergebniselemente per News-Meldung beim Suchenden.
- Universal Search: Ansatz von Google und anderen, nicht nur Textlisten mit Links zu WWW-Seiten als Ergebnisse zurückzuliefern, sondern daneben auch Prozesseinstiege und alle Facetten von verschiedenen Ergebnisverwertungen. Desktop-Suchen binden mit lokalen, Intranet- und Internet-Quellen unterschiedliche Bereiche in die Ergebnisse ein.

2.6. Modularität von Ressourcen

Die Modularität von Ressourcen, also die Bereitstellung von Informationsangeboten in Form von atomaren Einzeldiensten, die in externe, verteilte Systeme integriert werden können, dürfte einer der am stärksten beachteten Trends im Internet sein. Diese unter den Begriffen Mashup und Service-Orientierung betriebene Entwicklung wird allerdings bereits unter dem Begriff „Library 2.0“¹² ge- und behandelt und soll hier nicht weiter verfolgt werden. Dies ist somit bereits eine speziell bibliotheksorientierte Entwicklung und kein Trend im Internet, der auf den Arbeitsplatz übertragen werden kann.

Allgemeine Modularisierungsentwicklungen werfen jedoch die Frage auf, ob eine Mitnahme der Bibliotheksdienstleistungen und Personalisierungen innerhalb der Organisation in den Arbeitsalltag der Mitarbeiter möglich ist.

⁹ Vgl. <http://eureka.com/>

¹⁰ Vgl. <http://rollyo.com/>

¹¹ Vgl. z.B. <http://chacha.com>, <http://www.mahalo.com> und <http://prefound.com>,

¹² Vgl. [Casey: Library 2.0]

2.7. Parallel- oder Metamedien

Neben den eigentlichen Content-Transportern existieren mittlerweile einige als Meta- oder auch Parallelmedien zu bezeichnende Dienste, die dem Benutzer zur Nutzungsverwaltung von Hauptmedien dienen. In ihnen werden Neuigkeiten angekündigt, Änderungen verfolgt, die Hauptmedien referenziert oder zusammengefasst. Metamedien sind Social Search, Social Bookmarking, Wiki-Journalismus und umfassend das Medium Newsfeeds, das wiederum auch aus den anderen Metamedien generiert werden kann. So werden bereits Newsticker, Inhaltsverzeichnisse, Suchergebnisse, Änderungs- und Entwicklungsprotokolle als Stream-Lieferanten eingesetzt. Gerade die extreme Verbreitung von RSS zeigt die Relevanz der Parallelmedien in einem immer umfangreicheren webbasierten Internet.

Immer mehr Benutzer bewegen sich zunehmend im Bereich der Parallelmedien, die z.T. bereits eine ausreichende Informationsquelle darstellen oder sogar in einer vereinfachten Fassung den eigentlichen Inhalt darstellen können (Überschriften, Teaser oder Vollinhalte von Artikeln). Auch bei einer einfachen Referenzierung von Quellen werden die Benutzer ausschließlich und direkt auf die betreffende Inhaltsseite geleitet und kommen so nicht mehr über die eigentlichen Einstiegspunkte von Portalen und Wege, die ein Besucher nimmt, der „browser“-weise den Auftritt besucht.

2.8. Konvergenz von arbeitsbezogener und arbeitsnaher Informationsbeschaffung

Während „früher“ ein Bibliotheksnutzer innerhalb einer Organisation für die arbeitsbezogene Informationsbeschaffung spezielle elektronische Recherche-Werkzeuge einsetzte, die hochspezialisiert und proprietär für singuläre Funktionen notwendig waren und er daneben für die arbeitsnahe Informationsbeschaffung eher neutrale Werkzeuge nach Internet-Standards benutzte, haben sich diese beiden Tätigkeitsbereiche hinsichtlich der dabei benutzten Werkzeuge mittlerweile angeglichen.

So finden sich dann lediglich noch ein WWW-Browser, ein Groupware-System mit Kalender, Tasklist und E-Mail, sowie ein Newsfeed-Reader auf dem Arbeitsplatzrechner. Der WWW-Browser kann noch mit separaten Bookmark-Verwaltungen für unterschiedliche Themen ausgestattet werden, ein Newsfeed-Reader kann für die verschiedenen Aufgaben verschiedene Filter oder Feed-Kategorien einrichten. Während der Benutzung verschwimmen hier jedoch die Tätigkeitsbereiche. Vor allem die gefilterten Newsfeeds können relativ heterogen in einer Ansicht gelistet werden. Im WWW-Browser können so nun gleichzeitig das Recherche-System einer Online-Bibliothek neben allen anderen arbeitsbezogenen oder arbeitsnahen Systemen geöffnet sein.

Dabei sind evtl. Fachinformationseinhalte in der heterogenen Menge der sonstigen Informationen nicht ausreichend kenntlich. Es entsteht eine lediglich geringe

Aufmerksamkeit, da im gleichen Medium alle anderen Arbeits- sowie die arbeitsnahen Inhalte vorhanden sind.

Ein zu hinterfragendes Extrem-Beispiel wird im Zusammenhang mit der Entwicklung von Library 2.0 diskutiert: mit MashUp Biblio-Info in TV-Screen Corner¹³.

3. Anregungen für die Entwicklung der Organisations-kommunikation

Aus den aktuellen Entwicklungen ergeben sich Anregungen für die Kommunikation in der Organisation, sei es für die interne Kommunikation der Informationsabteilung oder zwischen der Informationsabteilung und ihren Kunden und Lieferanten.

Zum einen entstehen beim Benutzer, der im Internet Erfahrungen sammelt und Gewohnheiten entwickelt Ansprüche an Informations- und Recherche-Systeme, werden sie nun im privaten oder im Arbeitsumfeld genutzt. Der Benutzer bringt seine gewohnten Arbeitsumgebungen und Funktionen des Internet als Bedürfnis in die Organisation bzw. an den Arbeitsplatz.

Zum anderen ist es für interne Organisationen aus Gesichtspunkten der Effizienz, Nutzerakzeptanz und anderen Prozessverbesserungen sinnvoll, aus diesem Entwicklungspool Anregungen und Anleitungen zu übernehmen. Im Internet erfolgt die Entwicklung solcher Funktionen und entsprechender Systeme stets für sehr breit skalierte Systeme, unter dem Einfluss vieler Beteiligter. Sie werden in unterschiedlichen Systemen unter unterschiedlichen Bedingungen getestet und weiterentwickelt und bieten somit optimale Demonstrationsszenarios.

Notwendig für eine Ausrichtung an den Internet-Trends ist ein Monitoring der Internet-Techniken und deren Entwicklungen, also eine entsprechende Trendanalyse.

3.1. Ermittlung der relevanten Kommunikationsebenen

Zunächst sind Überlegungen darüber sinnvoll, an welchen Stellen diese Anregungen Eingang in die eigenen Prozesse finden sollen. Sie können in unterschiedlichen Schichten der internen Organisation aufgenommen werden:

- zwischen einer Informationsabteilung und ihren Lieferanten (z.B. ein Newsfeed für das Änderungsmanagement)
- als Vorgabe der Informationsabteilung für die Kommunikation zwischen den Lieferanten
- intern in der Informationsabteilung
- zwischen der Informationsabteilung und ihren Kunden

3.2. Anpassung der Angebote an Benutzerbedürfnisse

Die verschiedenen Ansätze zur Personalisierung von Informationsangeboten gehen von unterschiedlichen Motivationen aus:

Der Informationszugang kann durch Anpassung und Filterung effizienter werden, die Benutzerzufriedenheit steigt. Es kann hier also Effizienz für das Informationsmanage-

¹³ Siehe [Tebbut: Service Game], S. 15.

ment des Benutzers verbessert werden. Allerdings ist zu berücksichtigen, dass eine massive Anpassbarkeit von Portalen, die zahlreiche einzelne Informationselemente enthalten, auch zu einer Überforderung des Benutzers führen kann.

Gleichzeitig kann ein Informationsanbieter die impliziten Rückmeldungen in Form von Personalisierungen immer auch zur Optimierung der Angebote oder ihrer Strukturierung nutzen. Hier kann also Effizienz für die Informationsabteilung entstehen, die damit ihre Ressourcen wirtschaftlich einsetzen und planen kann. Bspw. können Angebote entfernt werden, die nicht oder nur selten genutzt oder in der Personalisierung eines Portals nach unten sortiert werden.

Durch implizite Rückkopplungen in Form von Personalisierungs-Einstellungen ist eine Form von Rückkanal für die Optimierung von Inhalten und bei der Einstellung von Suche-Ergebnissen für Ranking, Recall und Precision gegeben.

3.3. Inhalt und Nutzung von Search & Find

Für eine Entwicklung der eigenen Such-Infrastruktur (ist es eine eigene Infrastruktur oder sind es Applikationsinseln?) ergeben sich auf der Grundlage der vorgestellten Entwicklungen verschiedene Anregungen:

- Vereinheitlichung der Suchsysteme auf moderne Abfragetechniken
- Sind Definition von Kontexten und Teilthemen möglich oder ist die Suche lediglich in einer breiten Auslegung nutzbar?
- Kann die Suchekonfiguration in die Personalisierung einbezogen werden? Gibt es Auswahlmöglichkeiten von Suchmaschinen-Favoriten oder Konfigurationsmöglichkeiten einzelner Suchmaschinen?
- Sind die für ein Thema relevanten Suchdienste optimal z.B. in einer Metasuche verknüpft und darin erscheinende Ergebnisgruppen wiederum sinnvoll kategorisiert?
- Lassen sich Suchergebnisse flexibel verwenden? Sind die Suchausgaben multipel, enthalten also verschiedene Inhalts-Typen und Anwendungsgebiete in einer Ausgabe (Abwägung mit der Kontext-Definition)?
- Kann die Informationsabteilung bei ungenügender automatischer Leistung evtl. manuell unterstützend in den Sucheprozess eingreifen, z.B. mit Rankingdefinitionen, der Auswahl möglicher Quellen oder der Erstellung von Informationsseiten für häufig gesuchte/angefragte Themen?

3.4. Präsenz in mehreren Kanälen und Schaffung von Aufmerksamkeit

Die Informationsabteilung benötigt für ihre Dienstleistungen verschiedene Ausgabe-kanäle, um in allen wesentlichen Medien, eben auch den Parallelmedien und damit Lesewerkzeugen vertreten zu sein.

Denkbar ist hier die RSS-Wandlung sämtlicher Angebote, die einen zeitlichen Aktualitätsaspekt bzw. News-Charakter besitzen. Die eigenen Inhalte, falls geeignet, können als interne Newsfeeds ausgegeben werden. Zusätzlich können externe Angebote von News-Feeds nach intern transportiert und kontrolliert weitergegeben

werden. Dabei sollten die Fachinformations-Inhalte eindeutig gekennzeichnet sein, um von anderen Inhalten separiert werden zu können.

Bei der Ausgabe von Meta-Inhalten neben dem eigentlichen Auftritt ist zu bedenken, dass zwar Nutzer in ihren Parallelmedien erreicht, jedoch u.U. auch zum Besuch der Originalseiten angeregt werden sollen. Somit ist zu klären, in welchem Umfang die Inhalte in Newsfeeds gegeben werden und welche Referenzen auf die eigentlichen Quellen enthalten sind. Ist keine interaktive Nutzung von Angeboten notwendig, so kann u.U. der Informationsfluss vollständig über das Parallelmedium abgewickelt werden um informationsautonomen Benutzern entgegenzukommen.

3.5. Integration und Pflege externer Inhalte

Externe Inhalte, die auch für arbeitsnahe Informationsarbeit nutzbar wären, werden immer zahlreicher bei steigender Qualität. Gleichzeitig werden die Inhalte, ihre Struktur sowie ihre Verteilung und Auffindbarkeit im Netz immer vielfältiger und komplexer.

Hier sind zwei Extreme zu bedenken:

- Der Benutzer kann sich innerhalb der Organisation ausschließlich innerhalb der internen Quellen bewegen. Er weiß um die Existenz der externen und empfindet dieses Fehlen als arbeitsbehindernden Mangel.
- Der Benutzer bewegt sich lediglich innerhalb externer Quellen, findet dort jedoch arbeitsrelevante interne Ressourcen nicht.

Notwendig ist die Integration beider Quellen-Bereiche. Suche und Verlinkung können transparent beide Zonen einbeziehen. Hier kann auch die Notwendigkeit für die Informationsabteilung entstehen, einführend als „Guide“ für diese Quellen zu agieren. Für den Nutzer hilfreich können dabei aktive Auffindung, Kategorisierung, Zugänglichmachung und eine Pflege¹⁴ über ein entsprechendes Monitoring sein.

4. Umsetzungen in der Organisation

An dieser Stelle soll ein knapper Überblick über die Implementierungen bei Corporate Information & Research Management (IR) im Zusammenhang mit den oben genannten Anregungen aus technischen Entwicklungen im Internet gegeben werden. Die zentrale Präsenz von IR ist eingebettet in ein konzernzentrales Mitarbeiter-Portal, aus dem heraus an zahlreichen Stellen auf verschiedene Bereiche des Auftrittes von IR verlinkt wird.

4.1. Community-Bildung für Spezialthemen

Innerhalb des Auftrittes von IR werden zu besonderen Themen und für spezielle Benutzergruppen sogenannte „Fachportale“ oder „Channels“ geführt. Hierfür gibt es meist eine definierte Gruppe von Nutzern, die aktiv an der Gestaltung des Fachportals mitwirken bzw. über ihr Fachwissen Inhalte beisteuern, strukturieren und pflegen.

¹⁴ Zur organisierten Pflege und Beobachtung von Web-Inhalten vgl. auch [Alpar, Leich: Web Farming], S. 272

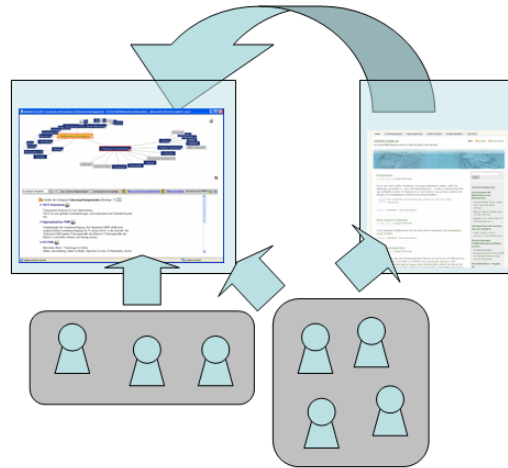


Abb. 1: Community-Bildung für Fachportale

Zur Unterstützung der Kommunikation bzgl. Entwicklung und Pflege bestehen für einige der Fachportale Weblogs, die als Datenspeicher und Kommunikationswerkzeug fungieren.

Bewegen sich die zuvor definierten Nutzer im Fachportal, so haben sie zunächst den gleichen Zugang zu Funktionen wie andere Nutzer außerhalb der definierten Gruppe. Zusätzlich jedoch steht für sie das Kommunikationswerkzeug Weblog kontextsensitiv zur Verfügung (**Abb. 1**). Wird im Fachportal eine Suche ausgeführt, so kann diese auch auf den Weblog ausgedehnt werden. Wird über einen Navigationsbaum in den Inhalten per Browsing gesucht, so steht per Link jeweils die entsprechende Weblog-Kategorie zur Ansicht oder zum Einfügen eines neuen Beitrages zur Verfügung.

Wesentlich ist hierbei, dass die Community, die für die Entwicklung eines Fachportals zuständig ist, nicht wie bei entsprechenden Internet-Werkzeugen frei gebildet, sondern gesteuert aufgebaut wird.

4.2. Social Search und Social Bookmarking

Im Auftritt von IR werden, organisiert in einer umfangreichen hierarchischen Struktur, Verlinkungen auf interne aber in größerem Maß auch externe Informationsressourcen gepflegt. Daraus ergibt sich ein durchsuch- und durchblätterbarer Datenbestand. Entsprechend einem Directory mit mehreren Redakteuren werden hier von den Information Professionals der Informationsabteilung interessante Quellen zusammengetragen.

Die Gruppe, die hier für Benutzer die Suche-Organisation und das Bookmarking übernimmt, ist dabei ebenfalls definiert. Beiträge gruppenexterner Benutzer werden von der Gruppe moderiert.

4.3. Unterstützung unterschiedlicher Ansätze der Ressourcen-Navigation

Allgemein werden sämtliche Informationsressourcen im Auftritt von IR nach mehreren Navigationsarten organisiert.

Bereits der Eintritt in die Oberfläche von IR aus dem konzernzentralen Mitarbeiter-Portal bietet neben einem zentralen Link auf die personalisierbare Einstiegsseite andere, bereits nach Fachinformations-Oberthemen getrennte Seiteneinstiege, über die Benutzer schneller an die von ihnen gesuchten Informationen gelangen.

Im IR-Auftritt selbst sind alle Ressourcen inkl. der Internet-Links nach Oberthemen geordnet. Gleichzeitig jedoch sind alle Quellen über eine auftrittübergreifende Suche auffindbar. Die Suche arbeitet dabei rollenbezogen und zeigt lediglich Informationen an, für die der Benutzer ein Nutzungsrecht hat.

4.4. Social Bookmarking/Harvesting innerhalb der Information Professionals

Für die Kommunikation der internen Gruppe von Information Professionals findet ein Austausch von arbeitsrelevanten Neuigkeiten per Weblog mit angebundenem Newsfeed statt. Eine entsprechende Kommunikation wäre ansonsten nur per E-Mail-Liste umsetzbar, was den Aufwand der Kommunikationsabwicklung immens erhöhen würde. Die sonst beim Weblog übliche 1:n-Kommunikation kann bei einer internen Nutzung beliebig abgewandelt werden. So können hier z.B. alle Leser des Weblogs in diesem auch schreiben.

4.5. Präsenz in den Parallelmedien

Die Parallelmedien werden in IR an verschiedenen Stellen mit Input versorgt. Neben den internen Quellen werden auch Ressourcen externer Dienstleister in dieser Form angeboten.

- In den Auftritt sind bereits externe Informationsdienstleister eingebunden, die ihre Spezialsuchen mit Alerting in Form von E-Mail oder Newsfeed ergänzen.
- Newsfeeds von externen Informationsanbietern werden direkt per Link oder auch über Gateways in den webbasierten Auftritt als News eingebunden.
- News der Abteilung für die Konzernmitarbeiter werden als E-Mail und
- als Newslisting auf dem Web-Auftritt angeboten. Dieses wiederum kann von Mitarbeitern als Newsfeed bezogen werden.

4.6. Personalisierung und Ranking von Inhalten



Abb. 2: Personalisierung mittels "myIR"

Eine personalisierte Anzeige sämtlicher Inhalte von IR ist im Auftritt über eine spezielle, vom Benutzer bestückbare „myIR“-Seite implementiert (**Abb. 2**). Diese ist dann nicht mehr oder nur noch im abstrahierten Sinn nach der Portalstruktur gestaltet. Die Navigationsstruktur wird insgesamt vereinfacht, wodurch ein abgekürzter Einstieg in die Online-Ressourcen entsteht. Trotzdem ist jedoch die Sichtbarkeit der allgemeinen System-News auf dieser Seite gewährleistet. Somit „verpasst“ der Benutzer keine wichtige Neuerung dadurch, dass er sich lediglich auf seiner persönlichen Seite bewegt.

Ein Ranking der Quellen ergibt sich aus dem Feedback, das die Benutzer implizit mit den Applikationsaufrufen und explizit mit der Platzierung auf ihren myIR-Seiten abgeben. Die Gruppe von Information Professionals wertet dieses Ranking aus und passt die Informationsauslieferung entsprechend an.

5. Beurteilung der Anwendung neuer Internet-Techniken in der Fachinformation

Bei allen neuen technischen Entwicklungen, die einen Einfluss auf die interne Kommunikation bzw. Infrastruktur haben können, stellt sich die Frage, in welchem Umfang diese übernommen bzw. angewendet werden können. Entschieden werden muss dabei, ob nur Teilaspekte oder das Konzept in vollem Umfang, in welchen Ebenen es intern zum Einsatz kommen soll, welche Zielgruppen angesprochen werden sollen und wer die Beteiligten bei z.B. kollaborativen Ansätzen sein sollen.

- Eine direkte Anwendung der im Internet verfügbaren Systeme ist innerhalb einer Organisation meist nicht möglich, da die Einrichtung externer Benutzerkonten mit internen Daten aus Sicherheitsgründen nicht möglich ist. Wegen der infrastrukturellen Besonderheiten der internen Organisation müssen die Entwicklungen bei der Übernahme ohnehin abgewandelt und angepasst werden.
- Eine Community-Bildung kann anders ablaufen, als dies bei offenen Systemen im Internet der Fall ist. Hier tauchen die Fragen auf, wer die beteiligten „Peers“

sind. Mögliche Bereiche sind Nutzer, Lieferanten und Informationsabteilung. Die Communities können innerhalb der Bereiche oder bereichsübergreifend eingerichtet werden.

- Die Übernahme bzw. Referenzierung von externen Quellen im internen Angebot kann lediglich in kontrollierter, kanalisierter Form ablaufen, da sonst die Befriedigung des Informationsbedarfs vollständig in einem bzgl. des Arbeitsthemas unstrukturierten Feld stattfindet.
- Beiträge von Nutzern, die über sog. Rückkanäle wieder in das Angebot der Fachinformation einfließen sollen, bedürfen einer Moderation. Eine vollständig freie Kommunikation in einer sich frei findenden Community ist nicht übertragbar.

Aus allen Entwicklungen und daraus entstehenden potenziellen Anregungen ergibt sich die Frage, wie die wissenschaftliche Informationsversorgung in 5 bis 10 Jahren aussieht: Nach Meinung des Autors ist sie geprägt von der technischen Ausrichtung auf das Internet oder sogar teilweise mit der Internet-Informationstechnik verschmolzen. Die Trends Consumerization in Bezug auf sämtliche Informations- und Kommunikationstechniken sowie die Suche-Fundierung in allen Anwendungsgebieten dürften sich dann endgültig etabliert haben.

Mit Blick auf die aus den neuen Entwicklungen gewonnenen Anregungen kann sich eine neue, erweiterte Rolle der Fachinformation ergeben, die von den folgenden Zusatzaufgaben geprägt sein würde:

- Bereitstellung von Plattformen für eine erweiterte Kommunikation. Die Plattformen lassen sich schnell und flexibel an verschiedene Kommunikationsarten anpassen.
- Moderation von verteilter Kommunikation. Zentral gesteuerte Moderation verschiedener Themen, Gruppen, Kommunikationsprinzipien
- Unterstützung von gruppenbasierten Prinzipien. Übernahme eines „Peer to Peer“-Parts
- Guide für organisationsübergreifende Informationsressourcen
- Competence Center für Suchetechniken (vor allem f. Query und Bedienung)

Insgesamt würde die interne Informationsabteilung zur Anlaufstelle für informationsautonome Benutzer als gleichwertige Kommunikationspartner.

Literatur

[Alpar, Leich: Web Farming] Alpar, Paul; Leich, Steffen: Web Farming, in Wirtschaftsinformatik, 42 (3), 2000, S. 271-274

[Casey: Library 2.0] Casey, Michael: Working Towards a Definition of Library 2.0, Abruf: 25.07.2007, Erstellung: 21.10.2005, http://www.librarycrunch.com/2005/10/working_towards_a_definition_o.html

[Exalead: Human-Machine Interaction] Exalead: Enabling Effective Human-Machine Interaction, Präsentation für Search Engine Meeting 2007, 04/2007, Boston, USA

[fast: Enterprise Search] fast: Enterprise Search: Bringing Order from Chaos, Präsentation für Search Engine Meeting 2007, 04/2007, Boston, USA

[Gartner: Consumerization] Gartner: Gartner Says Consumerization Will Be Most Significant Trend Affecting IT During Next 10 Years, Abruf: 25.07.2007, Erstellung: 20.10.2005, http://www.gartner.com/press_releases/asset_138285_11.html

[Google: Enterprise Search] Google: Secrets of the High-Powered Business: Staying Ahead with Enterprise Search, Präsentation für Search Engine Meeting 2007, 04/2007, Boston, USA

[Kuhlen: Fachinformation] Kuhlen, Rainer: Zukunft der Fachinformation - Fachinformation und Fachkommunikation in der Informationsgesellschaft, Abruf: 25.07.2007, letzte Änderung 18.03.2000, http://www.nethics.net/nethics/de/themen/fachinfo/fachinfopolitik_kuhlen.html

[Kuhlen: Informationsassistenten] Kuhlen, Rainer: Die Konsequenzen von Informationsassistenten, 1999, Suhrkamp Verlag, 1. Aufl.

[Palm: Google] Palm, Goedart: Die Welt ist fast alles, was Google ist – Zur Macht der Suchmaschinen, Abruf: 25.07.2007, Erstellung: 28.03.2002, <http://www.heise.de/tp/r4/artikel/12/12187/1.html>

[Tebbut: Service Game] Tebbut, David: Playing a new Service Game, in: Information World Review, März 2006, S. 13-15

[Weber: Copy-Paste] Weber, Stefan: Das Google-Copy-Paste-Syndrom - Wie Netzplagiate Ausbildung und Wissen gefährden, 2006, Heise Verlag, 1. Aufl.

[Wissenschaft im Dialog] Wissenschaftsorganisationen: Wissenschaft im Dialog, gefördert vom Bundesministerium für Bildung und Forschung, <http://www.wissenschaft-im-dialog.de/index.php4>

Collaborative Tagging als neuer Service von Hochschulbibliotheken

Christian Hänger, Christine Krätzs

Abstract in deutscher Sprache

Collaborative Tagging bietet Bibliotheken die Möglichkeit, ihr Wissensmanagement stärker partizipatorisch auszurichten, indem es die Kundinnen und Kunden an der Erschließung von Medien beteiligt. Die Universitätsbibliothek Mannheim möchte untersuchen, welchen Beitrag Collaborative Tagging für die inhaltliche Erschließung von bisher nicht erschlossenen und daher der Nutzung kaum zugänglichen Dokumenten auf Volltextservern oder in elektronischen Zeitschriften leisten kann, wie sie z. B. im Rahmen der durch die Deutsche Forschungsgemeinschaft finanzierten Nationallizenzen zur Verfügung stehen. Die Tags werden hinsichtlich ihrer Qualität und ihrer Struktur analysiert. Diese Ergebnisse werden anschließend mit einer ebenfalls an der UB Mannheim durchgeführten automatischen Indexierung von Volltexten verglichen. Ziel ist es, eine Empfehlung auszusprechen, wie bisher nicht erschlossene Dokumente am besten indexiert werden können: automatisch, mit Tags oder durch eine Kombination von beiden Verfahren.

Abstract in englischer Sprache

Collaborative Tagging offers a chance to change the knowledge management in academic libraries by involving the users in the process of creating and aggregating metadata. The Mannheim University Library would like to survey if Collaborative Tagging could be an instrument to organize and enable access to non-categorized digital documents. That could be useful especially for the huge repositories in context of the national licences for electronic text collections and journals which are financed by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). For that purpose the tags are analyzed regarding their quality and their structure. The findings are subsequently compared to the results of another project at Mannheim University Library, where index terms are automatically generated for full text documents. As a result of the project a recommendation will be given, in which way non-categorized documents can be indexed best: automatically, with tags or by a combination of both methods.

Tagging als Ausdruck einer neuen Lebenswirklichkeit

Vor dem Hintergrund einer anhaltend gewaltigen digitalen Informationsflut gestaltet sich die Orientierung und Navigation im Internet immer schwieriger. In Reaktion darauf hat sich mit dem Tagging eine neue Form der Organisation von (großen Mengen an) Informationen etabliert, die im Vergleich zur bisher gängigen digitalen Informationsverwaltung völlig anders strukturiert ist.

Tagging kann als freie Verschlagwortung von digitalen Ressourcen verstanden werden, bei der die Nutzerinnen und Nutzer auf der Basis von verschiedenen Social Software Anwendungen Webseiten (del.icio.us), bibliographisches Material (BibSonomy), Fotos (Flickr) und andere Internetressourcen mit Hilfe einer beliebigen Zahl von Schlagwörtern - sogenannten Tags - kennzeichnen, um sie innerhalb eines gemeinsam genutzten Datenpools später (besser) wiederzufinden.

Die wachsende Popularität des (Selbst-)Indexierens basiert auf der praktikablen Unterstützung der Organisation von eigenem Material (Fotos bei Flickr) oder von verwendeten Ressourcen Dritter (Webseiten z.B. bei del.icio.us) im Sinne eines besseren Retrievals dieser Informationen im Netz. Darüber hinaus bietet die gemeinschaftliche Sammlung auf der Grundlage der Verknüpfung gleich verschlagworteter Ressourcen einen quantitativ und qualitativ neuartigen und offenen Zugang zu Informationen von anderen. Diese (sozial) geteilten Informationen eröffnen ein umso größeres Reservoir an möglicherweise relevanten Daten, je mehr Personen sich in solch einem System beteiligen.

Im akademischen Bereich haben sich die Dienste BibSonomy, Connotea und CiteULike herausgebildet. BibSonomy wird von der Knowledge & Data Engineering Group der Universität Kassel betrieben. Im vorliegenden Projekt wurde zur Prüfung der Akzeptanz von Tagging durch die Bibliotheksnutzerinnen und -nutzer BibSonomy ausgewählt, da es die umfangreichsten Strukturierungs- und Netzwerkooptionen anbietet. Connotea und CiteULike konzentrieren sich vorrangig auf ein naturwissenschaftliches Fächerspektrum und sind daher weniger für die wirtschafts- und sozialwissenschaftliche Ausrichtung der Universität Mannheim geeignet.

Aktueller Forschungsstand zu Collaborative Tagging

Vor dem Hintergrund der wachsenden Bedeutung von Tagging im Internet wächst das Interesse, dieses Phänomen für den Bereich der Informations- und Bibliothekswissenschaft nutzbar zu machen. In der fachlichen Diskussion werden neben einer Reihe von differenzierten aber in erster Linie theoretischen Auseinandersetzungen mit den Vor- und Nachteilen von Collaborative Tagging im Vergleich zu herkömmlicher Sacherschließung, einzelne empirische Untersuchungen von Collaborative Tagging verhandelt.

Collaborative Tagging versus traditionelle Sacherschließung

Im Unterschied zur traditionellen Sacherschließung entfallen beim Tagging die Übersetzungsprobleme zwischen verschiedenen Vokabularen von Autoren, professionellen Sacherschließern und Nutzern (vgl. Mathes 2004). Die direkte Reflektion von Sprache und Vokabular der Anwenderinnen und Anwender wird als entscheidende Stärke des Collaborative Tagging angesehen (vgl. ebd./Shirky 2005). Insbesondere bei der zunehmend auch für Bibliotheken wichtiger werdenden Erschließung von Inhalten im Internet stößt die bibliothekarische Sacherschließung an Grenzen, da sie der hohen Dynamik von digitalen Inhalten nicht gewachsen ist.

Das beim gemeinschaftlichen Indexieren akkumulierte Vokabular kann dagegen den schnellen Wandel von Sprache und Begrifflichkeiten im Kontext des Internets abbilden (vgl. Shirky 2005), da die Erschließung der Verwendung von Ressourcen nicht vorgelagert ist, sondern parallel stattfindet. Das nutzergenerierte Vokabular entspricht den kognitiven Modellen der Nutzerinnen und Nutzer (vgl. ebd.) wodurch die schwierige konzeptionelle Antizipation ihrer Denkweisen entfällt (vgl. Quintarelli 2005).

Auf der anderen Seite weisen Taggingssysteme mit fehlender Synonymkontrolle, geringer semantischer Präzision und abwesender hierarchischer Eindeutigkeit zahlreiche Probleme auf. So benennen u. a. Guy/Tonkin (2006) die hohe Fehlerhaftigkeit (Synonyme, lexikalische Abweichungen, persönliche Färbung) von Tags und konstatieren, dass sie sich für eine effektive Suche im Sinne einer Suchanfrage nicht eignen. Die Wahrscheinlichkeit, bei einer Recherche viele irrelevante Treffer (sog. Ballast oder Noise) zu erhalten, ist relativ groß (vgl. Macgregor/McCulloch 2006). Dieses Problem tritt zwar in den Hintergrund, wenn sich eine genügend große Zahl von Nutzern beteiligt, das System bleibt jedoch ungenau und unspezifisch (vgl. Krosky 2005).

Tags definieren eher die Beziehung zwischen einer Ressource und dem Konzept, das der Nutzer mit ihm verknüpft, als Dokumente wie in Klassifikationen exakt durch einen Indexterm zu bezeichnen (vgl. Guy/Tonkin 2006). Anders als beim Klassifizieren werden durch die Vergabe von Tags eben keine systematischen Beziehungen zwischen Objekten hergestellt, sondern Informationen verschiedener Nutzerinnen und Nutzer auf der Grundlage von Ähnlichkeiten miteinander verbunden (vgl. Mathes 2004).

Mathes vermutet, dass insbesondere die als Schwächen identifizierten Eigenschaften von Taggingssystemen deren wachsenden Erfolg begründen, da sie die Barrieren der Kooperation verringern (vgl. ebd.). Auch Shirky (2005) sieht in den Ungenauigkeiten von Taggingssystemen eine Stärke. Indem sie sprachliche Mehrdeutigkeiten zulassen, können sie die feinen Unterschiede in Hinblick auf die Interessen und das Wissen derjenigen, die Inhalte taggen, viel genauer repräsentieren.

Collaborative Tagging kann für die Suchstrategien der Nutzerinnen und Nutzer auch deshalb von Vorteil sein, weil eine große Menge an verschiedenen Tags eine größere Zahl an Einstiegspunkten für die Suche bietet und zugleich mehr Möglichkeiten schafft, zufällige Entdeckungen zu machen (vgl. Mathes 2004/ Quintarelli 2005). Tagging unterstützt weniger die gezielte Suche nach relevanten Inhalten zu einer konkreten Fragestellung, sondern vor allem das Browsen und die ungerichtete Suche nach interessanten Inhalten. Im sofortigen Feedback bei der Eingabe und dem dadurch möglichen Zugriff auf alle Objekte, die mit dem gleichen Begriff getaggt wurden, liegt nach Ansicht von Mathes (2004), neben der Reduktion von Komplexität und zeitlichen wie kognitiven Kosten, der entscheidende Vorteil von Taggingssystemen.

Taggingssysteme unterscheiden sich in ihrer Dynamik von Interaktion und Partizipation ebenso wie in ihrer Architektur. Im Hinblick auf den Gebrauchswert der

in diesen Systemen gesammelten Informationen ist neben dem konkreten Design auch die Motivation der Nutzerinnen und Nutzer von großer Bedeutung. Motivationen können aus persönlichem Nutzen, aber auch aus sozialen Interessen erwachsen (vgl. Marlow et. al 2006).

Empirische Untersuchungen von Collaborative Tagging

In den vorliegenden empirischen Studien wurden einige der diskutierten Annahmen überprüft. Guy/Tonkin (2006) konnten in ihrer Untersuchung von zwei Social Software Anwendungen (del.icio.us und flickr) die Vermutung bestätigen, dass die von den Anwenderinnen und Anwendern selbst generierten Systeme hohe Fehlerquoten bezüglich der verwendeten Schlagwörter aufweisen. So waren 28% bzw. 40% der Tags bei del.icio.us und Flickr als Suchbegriffe unbrauchbar, da sie fehlerhaft oder stark persönlich gefärbt waren.

Golder/Huberman (2005) stellen in einer empirischen Untersuchung des Bookmarkdienstes del.icio.us fest, dass das Verhalten der Nutzerinnen und Nutzer in Bezug auf die Anzahl und die Frequenz der Tag-Vergabe stark variiert – einige vergeben wenige, andere viele Tags. Auch die Tags selbst und die Inhalte, die sie beschreiben, unterliegen einem Veränderungsprozess. Hinsichtlich der Größenverhältnisse lassen sich dennoch stabile Muster erkennen: Bereits nach einer verhältnismäßig geringen Anzahl von Bookmarks für eine URL (ca. 100) beginnt sich ein Konsens bei der Tagvergabe herauszubilden und die Variation der verwendeten Tags sinkt. Insgesamt sind Tags mit allgemeinen Bedeutungen stärker vertreten als Tags, die mit persönlichen Bedeutungen verknüpft sind.

Voss (2006) vergleicht in seiner Untersuchung die Kategorisierung von Inhalten in der Wikipedia mit Klassifikationen wie der Dewey Decimal Classification (DDC) und Collaborative Tagging bei del.icio.us. Seiner Meinung nach handelt es sich bei der Wikipedia um einen Thesaurus, der Tagging und hierarchisches Systematisierung von Wissen vereinigt. Weil diese Systeme hinsichtlich der Verteilung von Indextermen und Datensätzen interessante Vergleichsmöglichkeiten bieten, plädiert Voss dafür, den Blick stärker auf die konzeptionellen Eigenschaften zu richten und zu überlegen, wie sich diese Systeme sinnvoll kombinieren lassen, statt Collaborative Tagging und Klassifikation als Gegensatz zu begreifen.

Abgesehen von den Problemen des Collaborative Tagging, die sich vor allem aus der Vielfalt der Sprache ergeben, stellt gemeinschaftliches Indexieren ein Instrument dar, das den Nutzerinnen und Nutzern ermöglicht, sich aktiv an der Wissensorganisation in einer Bibliothek zu beteiligen. Collaborative Tagging kann die traditionelle Inhaltserschließung erweitern und verbessern und zugleich ein wichtiger Schritt hin zu einem "konsequenten Perspektivwechsel zur Anwendersicht" (DFG 2004: 13) sein. Allerdings sollten die Vorteile kontrollierten Vokabulars für eine effiziente Suche nicht einfach aufgegeben, sondern den Nutzern verstärkt vermittelt werden (vgl. Macgregor/McCulloch 2006).

Heller verweist auf den unterschiedlichen Ansatz dieser zwei Arten der Sacherschließung und auf die Chance einer "neuen Ebene der Kommunikation"

(2007: 8) zwischen Autoren, Nutzern und Bibliothekaren, die entstehen könnte, wenn es Bibliotheken gelänge, auf den Kundennutzen fokussierte Tagginganwendungen für ihre Recherchesysteme zu adaptieren.

Zukünftig könnten sich der klassische Erschließungsansatz (Indexieren durch ausgebildetes Bibliothekspersonal) und der neue soziale Ansatz (Indexieren durch die Kundinnen und Kunden) ergänzen. Denkbar wäre beispielsweise, die Schlagwortnormdatei (SWD) in ein Taggingsystem einzubinden und den Kundinnen und Kunden genormtes Vokabular zum Taggen anzubieten um so die jeweiligen Vorteile zu verbinden.

Ziele des Forschungsprojekts

1. Optimierung der Fachrecherche durch den Einsatz partizipatorischer Elemente

Die Präsentation digitaler Inhalte im Internet stellt für die wissenschaftliche die gleiche Herausforderung wie für die alltägliche Recherche dar: Recall und Precision von Recherchen verschlechtern sich, da die digitalen Inhalte nicht bibliothekarisch erschlossen sind. Ein signifikantes Beispiel hierfür sind die im Rahmen der Nationallizenzen bereit gestellten elektronische Dokumente, die eine deutliche Erweiterung der Angebote der Hochschulbibliotheken darstellen, häufig aber von den Nutzerinnen und Nutzern nicht gefunden werden, da sie sich aufgrund ihrer heterogenen Struktur nur schwierig in die bibliothekarischen Recherchesysteme einbinden lassen und zudem inhaltlich nicht erschlossen sind.

Das bewusst experimentell angelegte Projekt möchte untersuchen, ob durch den Einsatz von kollaborativem Tagging die Fachrecherche in nicht inhaltlich erschlossenen Volltexten optimiert werden kann. Im Rahmen des Vorhabens sollen die digitalen Inhalte des Volltextservers der Universitätsbibliothek Mannheim und die elektronischen Zeitschriften von Elsevier aus dem Bereich Social Sciences and Humanities erschlossen werden.

2. Erweiterung eines Recherchesystems durch Taggingfunktionen

Im Rahmen eines gemeinsamen DFG-Projekts von Prof. Dr. Heiner Stuckenschmidt und der Universitätsbibliothek Mannheim werden elektronische Dokumente aus den Sozial- und Wirtschaftswissenschaften automatisch erschlossen¹. Dabei verwendet das System die relative Häufigkeit von Begriffen aus dem Thesaurus in den zu indizierenden Dokumenten, um die Relevanz eines Begriffes im Hinblick auf das Thema eines Dokuments zu bestimmen. Die Suchergebnisse werden in der graphischen Clustermap-Technologie präsentiert. Die Universitätsbibliothek Mannheim möchte exemplarisch die graphische Oberfläche des Recherchesystems Collexis der Firma SyynX um Taggingfunktionen erweitern. Dabei können die Nutzerinnen und Nutzer gefundene Ressourcen über einen konstanten Link im

¹ Es handelt sich um das DFG-Projekt "Verbesserung der Fachrecherche in großen Volltextsammlungen mit Methoden des Semantic Webs" (kurz Semantic Web).

Taggingssystem BibSonomy speichern und mit Tags kennzeichnen. Gleichzeitig wird im Index der Suchmaschine ein Verweis auf die eingegebenen Tags eingefügt, mit dessen Hilfe die Nutzer über die grafische Oberfläche auch verwandte Dokumente finden können.

3. Evaluierung des Nutzerverhaltens bei Tagging in Hochschulbibliotheken

Die Universitätsbibliothek Mannheim möchte evaluieren, welche Akzeptanz Tagging bei den Studierenden sowie den Wissenschaftlerinnen und Wissenschaftlern der Universität Mannheim hat. Im Rahmen des Projekts sollen für bestimmte Zeiträume Taggingssysteme in unterschiedlichen Ausprägungen an die Suchmaschine Collexis angebunden werden. In den jeweiligen Untersuchungsphasen wird die Akzeptanz der Angebote quantitativ und qualitativ ausgewertet. Bei der quantitativen Erhebung wird die Anzahl der Zugriffe auf die verschiedenen Taggingangebote im jeweiligen Untersuchungszeitraum erfasst und in Relation gesetzt.

4. Untersuchungen zur Qualität der Tags

Die durch die Nutzerinnen und Nutzer vergebenen Tags werden hinsichtlich ihrer Qualität untersucht. Die Qualität der Tags wird durch Nutzung der im Projekt Semantic Web erarbeiteten Methoden ermittelt. Dabei werden die von den Kundinnen und Kunden vergebenen Tags in Relation zu automatisiert bzw. von Fachkräften vergebenen Schlagwörtern gesetzt. Die Indexterme der Fachkräfte werden als Goldstandard definiert und anschließend der Grad der Ähnlichkeit zu den automatisiert vergebenen und durch Nicht-Fachkräfte vergebenen Indextermen ermittelt. Anschließend wird untersucht, inwiefern sich Recall und Precision durch die von den Nutzerinnen und Nutzern vergebenen Tags verändern.

Im Rahmen des Projekts soll untersucht werden, auf welche Weise nicht erschlossene Dokumente am besten indexiert werden können – automatisch, durch Tags oder durch eine Kombination der beiden Verfahren. Damit hat das Projekt einen hohen Stellenwert für die tägliche Arbeit in den Hochschulbibliotheken, da es nach Instrumenten zur Ordnung innerhalb unüberschaubarer Informationsmengen sucht (vgl. KROSKI 2005).

Methodisches Vorgehen

Evaluierung des Nutzerverhaltens bei Tagging durch quantitative Verfahren

Für den Vergleich von verschiedenen Angeboten im Recherchesystem der Universitätsbibliothek Mannheim wird das quantitative Verfahren der Logfile Analyse gewählt. Logfiles haben den Vorteil, dass sie vom Webserver automatisch generiert und gespeichert werden, ohne dass den Nutzerinnen und Nutzern bewusst wird, dass ihr Verhalten Gegenstand von Beobachtung ist. Von den in diesen Textdateien enthaltenen Informationen über die Kommunikation von Server und Netz sind für das Vorhaben vor allem die Daten interessant, die die Untersuchung der Navigation von

Kundinnen und Kunden auf den Bibliotheks-Webseiten ermöglichen. Konkret soll die Auswahl verschiedener Informations- und Navigationsmöglichkeiten in den Recherchesystemen der UB Mannheim durch die statistische Auswertung des Klickverhaltens der Kundinnen und Kunden analysiert werden.

Logfiles enthalten allerdings viele Daten, die eine Analyse verzerren können. Um einzelne Seitenaufrufe sowie Nutzer und Sitzungen identifizieren zu können, ist ein Preprocessing zur Aufbereitung der Daten notwendig (vgl. Hippner et al. 2002b). Für die gezielte Untersuchung des Klickverhaltens müssen die aufgezeichneten Daten einzelnen Nutzer-Sessions zugeordnet werden. Es bleibt noch zu prüfen, ob diese Identifikation auf der Basis von Session-IDs oder durch die Vergabe von Cookies erfolgen soll. Session-IDs liefern die genaueren Daten, erfordern aber die Anmeldung der Nutzerinnen und Nutzer im OPAC, was eine zusätzliche Hürde und möglicherweise eine Beeinflussung der Nutzer bei ihrer Navigation zur Folge haben könnte. Da Cookies Rechnern und nicht Personen zugeordnet werden, wird eine Identifikation von OPAC-Recherchen an von vielen Personen genutzten Rechnern erschwert (vgl. Säuberlich 2002).

Ziel ist es, zusammenhängende Nutzungsvorgänge hinsichtlich folgender Kenngrößen auszuwerten: Wie verteilen sich die Hits (Anfragen) eines Page Views (Sichtkontakt mit HTML-Seite) innerhalb eines Visits (Session) auf die jeweils angebotenen Links zur herkömmlichen Sacherschließung (RVK; RSWK/SWD) und auf die Links zu den verschiedenen Tagging Angeboten in den einzelnen Arbeitsschritten? In jeder Untersuchungsphase wird zudem geprüft, wie sich die Reaktionen der Kundinnen und Kunden auf die jeweiligen Angebote im Zeitverlauf verändern (vgl. Englbrecht 2002).

Untersuchungen der Qualität der Tags durch Ähnlichkeitsfunktion

Die Qualität der Tags wird mit Hilfe der Ähnlichkeitsfunktion gemessen, ein Verfahren, das die Universitätsbibliothek Mannheim bereits erfolgreich für die Qualitätskontrolle automatisiert vergebener RVK-Notationen eingesetzt hat². Dabei wurden aus dem OPAC zufällig 1000 Titel ausgewählt, die intellektuell und automatisiert vergebene Notationen haben. Die durch bibliothekarische Fachkräfte vergebenen Notationen wurden als Goldstandard gesetzt und der Grad der Ähnlichkeit zu den automatisiert erschlossenen Notationen gemessen. Vollständige Ähnlichkeit bestimmt sich in der exakten Übereinstimmung zwischen den beiden Indextermen, der Grad der Abweichung in der Entfernung vom jeweiligen Knoten im Klassifikationsbaum.

Zur Qualitätskontrolle der Tags muss das Verfahren modifiziert werden. Die erste Modifikation betrifft den Umstand, dass das oben beschriebene Verfahren bei Klassifikationen Verwendung findet, es sich beim Tagging jedoch um eine Form der verbalen Sacherschließung handelt. Im beantragten Projekt werden zunächst die Tags mit den Schlagwörtern der SWD (Schlagwortnormdatei) verglichen. Gibt es

² Dieses Verfahren hat Magnus Pfeffer, Universitätsbibliothek Mannheim entwickelt.

eine Übereinstimmung, wird das Ergebnis notiert. Bei Abweichungen werden die frei vergebenen Schlagwörter in den Registereinträgen der RVK-Klassen gesucht. Dadurch wird untersucht, ob das frei vergebene Schlagwort in den richtigen Zweig des Klassifikationsbaumes eingeordnet worden ist und in welchem Abstand zum relevanten Knoten der Indexterm steht. Dieses Verfahren wird bei der Auswertung der Tags im Dokumentenservers MADOC angewandt, dessen Titel Schlagwörter beinhalten, die von bibliothekarischen Fachkräften und Autorinnen und Autoren der Beiträge vergeben sind.

Die zweite Modifikation betrifft den unterschiedlichen Grad der Erschließung der zugrunde liegenden Datenquellen. Da die elektronischen Zeitschriften von Elsevier noch nicht erschlossen sind, muss zur Ermittlung des Goldstandards zu einem Hilfsmittel gegriffen werden. Diese Dokumente werden automatisiert durch die Software Collexis erschlossen und diese Indexterme – gekürzt um einen durchschnittlichen Abstandswert zu den intellektuell ermittelten Indextermen – als Goldstandard gesetzt.

Untersuchung der Struktur der Tags

Im Rahmen des beantragten Projekts soll ein Vorschlag erarbeitet werden, welche Kombination von Merkmalsausprägungen ein Indexsystem haben muss, um einen optimalen Service für die Nutzerinnen und Nutzer bei der Recherche nach wissenschaftlichen Dokumenten zu bieten. Tabelle 1 zeigt die möglichen Strukturelemente der verschiedenen Indexsysteme.

Tab. 1: Vergleich von RVK, RSWK/SWD und BibSonomy

	RVK	RSWK/SWD	BibSonomy
Struktur des Indexsystems	Klassifikation	Thesaurus	keine
Elemente/Index-terme	Klassen	Schlagwörter	Tags
Art der Relation zwischen den Indextermen	hauptsächlich Hierarchie	Äquivalenz/ Hierarchie/ Assoziation	keine Relation
Indexat/Objekte/ Dokumente	physische Medien	physische Medien	Webseiten
Quelle der Ressourcen/ Dokumente	Fremde	fremde	hauptsächlich fremde (eigene möglich)
Indexierungsrechte	Redaktion	Redaktion	alle Nutzer (Registrierung erforderlich)
Normierung/Kontrolle	Hoch	Hoch	keine
automatische Unterstützung bei der Indextermvergabe	möglich	Nein	auf der Grundlage eigener und zu einem Dokument vergebener Tags
Recall	Hoch	Hoch	hoch
Precision	Gering	mittel bis hoch	gering
Noise/Ballast	Gering	Gering	hoch
Flexibilität	Gering	Mittel	hoch
Ziel	Groberschließung	Feinerschließung	individuelle und kollaborative Erschließung
Motivation	Bestandserfassung	thematische Suche unterstützen	subjektive und kollaborative Wiederauffindbarkeit erhöhen
Synonymkontrolle	-	Ja	nein
Homonymkontrolle	-	Ja	nein
unterstützte Recherchestrategie	Browsen	Suche	Browsen
Personalisierungsmöglichkeit	Nein	Nein	ja

Ausblick

Das vorgestellte Projekt steht im Kontext vergleichbarer Projekte an der Universitätsbibliothek Mannheim, die sich mit Formen alternativer Sacherschließung beschäftigen. Die Universitätsbibliothek Mannheim hat ein Programm zur automatischen Erschließung von Regensburger Verbundklassifikationen entwickelt und führt gemeinsam mit Prof. Dr. Heiner Stuckenschmidt das bereits erwähnte Projekt mit dem Titel „Verbesserung der Fachrecherche in großen Volltextsammlungen mit Methoden des Semantic Webs“ durch. Angesichts einer stetig wachsenden Menge von wissenschaftlichen digitalen Informationen, die nicht mehr allein intellektuell erschlossen werden können, geht es nicht darum, die traditionelle intellektuelle Sacherschließung zu ersetzen, sondern vielmehr darum, sie sinnvoll zu ergänzen und zu optimieren. Durch den Einsatz von Tagging-Funktionen soll untersucht werden, auf welche Weise nicht erschlossene Dokumente am besten indexiert werden können – automatisch, durch Tags oder durch eine Kombination der beiden Verfahren. Das Projekt hat einen hohen Stellenwert für die tägliche Arbeit in Hochschulbibliotheken, da es Möglichkeiten der Ordnung unüberschaubarer Informationsmengen erforscht.

Literatur

BibSonomy.org <http://www.bibsonomy.org> (Zugriff 16.04.2007).

CiteULike.org <http://de.citeulike.org> (Zugriff 13.03.2007).

Connotea.org <http://www.connotea.org> (Zugriff 13.03.2007).

del.icio.us <http://del.icio.us> (Zugriff 13.03.2007).

DFG: (2004): Aktuelle Anforderungen der wissenschaftlichen Informationsversorgung. Empfehlungen des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme und des Unterausschusses für Informationsmanagement vom 11./12.März 2004.
http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/strategiepapier_wiss_informationsvers.pdf (Zugriff 26.02.2007).

Englbrecht, Andreas (2002): Deskriptive Logfile-Analysen. Durchführung und Einsatzpotentiale. In: Hippner, Hajo et al. (Hrsg.): Handbuch Web Mining im Marketing. Konzepte, Systeme, Fallstudien. Braunschweig u. a.: Vieweg, S. 125-139.

flickr.com <http://www.flickr.com> (Zugriff 13.03.2007).

Golder, Scott A./Huberman, Bernardo A. (2005): The Structure of Collaborative Tagging Systems. Information Dynamics Lab: HP Labs, Palo Alto, USA.
<http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf> (Zugriff 01.02.2007).

Guy, Marieke/Tonkin, Emma (2006): "Folksonomies: Tidying up Tags?" In: D-Lib Magazine, Vol. 12 No. 1. <http://www.dlib.org/dlib/january06/guy/01guy.html> (Zugriff 01.02.2007).

Heller, Lambert (2007): Bibliographie und Sacherschließung in der Hand vernetzter Informationsbenutzer. In: Preprint des Sonderhefts "Bibliothek2.0" der Zeitschrift Bibliothek Forschung und Praxis (erscheint im Sommer 2007). http://www.bibliothek-saur.de/preprint/2007/ar_2448_heller.pdf (Zugriff 26.07.2007).

Hippner, Hajo/Merzenich, Melanie/Wilde, Klaus D. (Hrsg.) (2002a): Handbuch Web Mining im Marketing. Konzepte, Systeme, Fallstudien. Braunschweig u. a.: Vieweg.

Hippner, Hajo/Merzenich, Melanie/Wilde, Klaus D. (2002b): Grundlagen des Web Mining – Prozess, Methoden und praktischer Einsatz. In: Dies. (Hrsg.): Handbuch Web Mining im Marketing. Konzepte, Systeme, Fallstudien. Braunschweig u. a.: Vieweg, S. 3-31.

Kroski, Ellyssa (2005): The Hive Mind: Folksonomies and User-Based Tagging. Blogeintrag auf InfoTangle. <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/> (Zugriff 01.02.2007).

Macgregor, George/McCulloch, Emma (2006): Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. http://eprints.rclis.org/archive/00005703/01/CollaborativeTaggingToolPaperGmEm_preprint.pdf (Zugriff 05.02.2007).

Marlow, Cameron/Naaman, Mor/Boyd, Danah/Davis, Marc (2006): Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. <http://www.rawsugar.com/www2006/29.pdf> (Zugriff 01.02.2007).

Mathes, Adam (2004): Folksonomies - Cooperative Classification and Communication Through Shared Metadata. <http://adammathes.com/academic/computer-mediated-communication/folksonomies.pdf> (Zugriff 05.02.2007).

Quintarelli, E. (2005): Folksonomies: power to the people. <http://www.iskoi.org/doc/folksonomies.htm> (Zugriff 13.02.2007).

Säuberlich, Frank (2002): Vorverarbeitung von Web-Daten – Pre-Processing. In: Hippner, Hajo et al. (Hrsg.): Handbuch Web Mining im Marketing. Konzepte, Systeme, Fallstudien. Braunschweig u. a.: Vieweg, S. 107-123.

Shirky, Clay (2005): Ontology is Overrated: Categories, Links, and Tags. http://shirky.com/writings/ontology_overrated.html (Zugriff 01.02.2007).

Voss, Jakob (2006): Collaborative thesaurus tagging the Wikipedia way. <http://arxiv.org/abs/cs/0604036> (Zugriff 05.02.2007).

Mobile IT in kleinen und mittleren Medienunternehmen zur Prozess- und Serviceverbesserung

Christoph Bläsi

Das Projekt TECH-CYCLE verschiedener Institute der Friedrich-Alexander-Universität Erlangen-Nürnberg geht von den beiden Grundannahmen aus, dass erstens der Einsatz von Informationstechnologie (IT), insbesondere mobiler IT, die Innovationskraft von Unternehmen stärkt und dass zweitens der Einsatz solcher Geräte durch kleinere und mittlere Unternehmen (KMUs) verschiedenen Hürden wie beschränkten Finanzmitteln und Personalressourcen gegenübersteht. Das Projekt hat es sich darauf aufbauend zur Aufgabe gestellt, im Rahmen eines Second-Product-Lifecycle-Ansatzes, d.h. im Rahmen eines Ansatzes, der den Einsatz von bei den Erstanwendern im Rahmen eines Generationswechsels ersetzter mobiler Hardware bei KMUs vorsieht, Transfermodalitäten zu entwickeln und zu beschreiben, die die Übergabe mobiler Hardware für den zweiten Teil ihres Produkt-Lebenszyklusses an KMUs effektiv gestalten lassen. Dabei ist offensichtlich, dass es nicht reicht, wenn nur Geräte den Besitzer wechseln, sondern dass es darüber hinaus unabdingbar ist, dass auch Wissen über die zweckgerichtete Einführung und Anwendung dieser Geräte, möglicherweise auch die bei deren Anwendung eingesetzten Applikationen (Software) an die KMUs weitergegeben werden (ggf. entsprechend angepasst). In dem hier im Zentrum stehenden Teilprojekt wird die Erlanger Buchwissenschaft zunächst herausarbeiten, welchen Beitrag der Einsatz mobiler IT in kleinen und mittleren Unternehmen der Medienbranche an im Branchendiskurs und / oder der wissenschaftlichen Reflexion der Branchenentwicklung bereits identifizierten und dokumentierten Ansatzpunkten für Prozess- und Serviceverbesserungen sowie an durch Experteninterviews und eine Befragung lokalisierten Punkten leisten kann. Im Anschluss daran wird sie in Feldversuchen mit exemplarischen Unternehmen aus dem Unterbereich Print und digitale Entsprechungen nachweisen, dass die vorhergesagten Verbesserungseffekte tatsächlich erzielt werden können, sowie herausarbeiten bzw. ggf. herbeizuführen helfen, dass die in vorgängigen Teilprojekten des Projektes entwickelten Transfermodalitäten für Hardware und Anwendungswissen um mobile IT, möglicherweise einschließlich Applikationen, für Unternehmen der Medienbranche effektiv und effizient sind.

A very common assumption is that Information Technology (IT), especially mobile IT, enable innovation. The TECH-CYCLE project of several institutes of the University of Erlangen-Nuremberg aims to develop and validate processes of mobile IT transfer in the second part of its life cycle. The project will examine both mobile IT hardware that is being replaced by its primary users (typically large scale enterprises) in the course of adopting newer generations of technology, and also knowledge about the

introduction and use of these systems, possibly including its software applications. The recipients of the transfer in terms of this project are small and medium-size enterprises (SMEs) that typically lack the financial and human resources to be early adopters of mobile IT. In a subproject, the research faculty of the University of Erlangen-Nuremberg who study the publishing industry, and their industry partners, will demonstrate the potential of mobile IT for SMEs in the media sector. In order to establish this potential, the project will investigate how mobile IT helps solve problems in business processes and service requirements that are either a) previously identified in this economic sector; and b) identified by industry deciders in a survey conducted in the framework of this subproject. The study will also evaluate the transfer modalities for hardware and knowledge around it, particularly its application software, as mobile IT is transferred from its first users to SMEs. The faculty for the study of publishing has chosen as industry partners for this study first, a publisher, and second, a book retailer. Operation scenarios to be considered for the impact of mobile IT include the publisher's editorial process and the book retailer's sales presentations.

Das Gesamtprojekt

Beim Projekt „Second-Product-Lifecycle-Strategien für Hightech-Geräte zur Steigerung der Innovationsfähigkeit kleiner und mittelständischer Unternehmen sowie nichtstaatlicher Organisationen“ (TECH-CYCLE) geht es darum, die Erfolgsfaktoren dafür herauszuarbeiten, zu beschreiben und für die Anwendbarkeit der daraus zu ziehenden Schlüsse aufzubereiten (in Form von Transfermodalitäten), mobile Hardware nebst dem zugehörigen Anwendungswissen, möglicherweise einschließlich der eingesetzten Applikationen, von Erstanwendern an kleine und mittlere Unternehmen weiter zu geben. Es wird dabei davon ausgegangen, dass die weiter gegebene Hardware bei den Erstanwendern, typischerweise Großunternehmen, im Rahmen des Wechsels auf die nächste Gerätegeneration frei wird und nach der Weitergabe – für ein weiteres Stadium in ihrem Produkt-Lebenszyklus – bei KMUs Dienst tut, die die Möglichkeiten eines Ersteinsatzes solcher Hardware typischerweise nicht haben. Am Projekt werden drei Wirtschaftsinformatik-Lehrstühle der Universität Erlangen-Nürnberg sowie die Lehrstühle für Wirtschaftspädagogik und für Gesundheitsmanagement und die Buchwissenschaft dieser Universität beteiligt sein. Aufgabe der letztgenannten drei Institutionen ist es dabei v.a., die Anwendung der erarbeiteten Transfermodalitäten exemplarisch in verschiedenen Bereichen (Gesundheitsmanagement, Bildung, Medien) vorzubereiten, zu begleiten, den Erfolg des daraus resultierenden Einsatzes mobiler IT zu überprüfen sowie die sich ergebenden Erkenntnisse in die Optimierung der vorgeschlagenen Transfermodalitäten einfließen zu lassen. Als Entwicklungspartner werden an dem Projekt noch Research in Motion als Hersteller mobiler Hardware, T-Mobile als Anbieter und Betreiber von Lösungen mit solcher Hardware und Dr. Städtler als einschlägiges Beratungsunternehmen beteiligt sein.

sowie für jeden der abgedeckten Bereiche mindestens ein mittelständisches Unternehmen als Feldpartner.

Die Stränge des Projektes

Die verschiedenen Stränge des Projektes fügen sich folgendermaßen zu einem Ganzen:

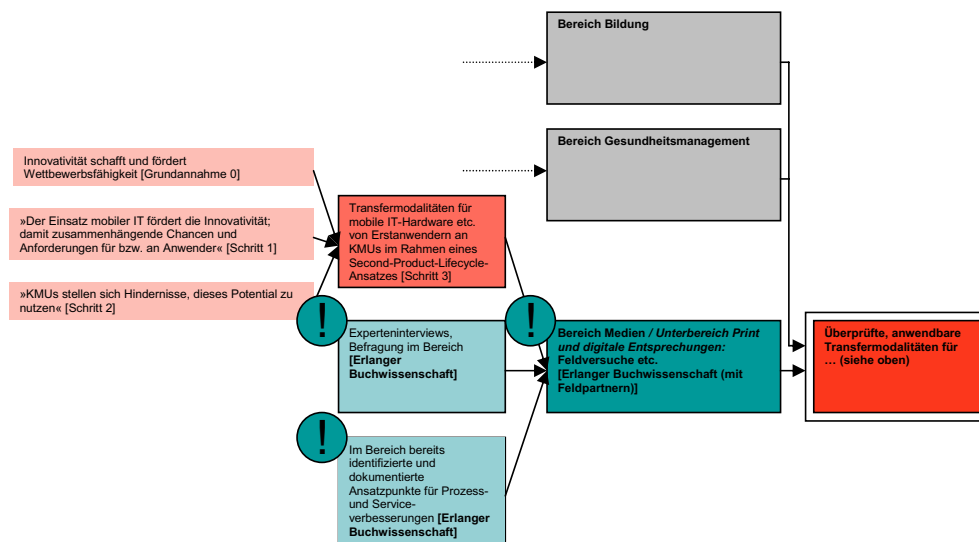
- Grundannahme 0: Innovativität schafft und sichert Wettbewerbsfähigkeit – vgl. dazu z.B. (Leonard-Barton 1995), (Leifer/McDermott/Colarelli/Peters/Rice/Veryzer 2000), (Tidd/Bessant/Pavitt 2005), zum weiteren Umfeld z.B. auch (Vahs/Burmester 2006) und für den Medienbereich z.B. (Habann 2003)¹ ; diese Annahme wird im Rahmen des Projektes noch einmal belegt
- Grundannahme 1: IT, insbesondere mobile IT, unterstützt die Innovativität – das wird im Rahmen des Projektes für den typischen Fall von Großunternehmen als Erstanwendern noch einmal belegt sowie die daraus resultierenden Chancen und Anforderungen für bzw. an Anwender herausgearbeitet: »Der Einsatz mobiler IT fördert die Innovativität; damit zusammenhängende Chancen und Anforderungen für bzw. an Anwender« (Schritt 1)
- Grundannahme 2: KMUs haben – in Bezug auf Finanzmittel, Personalressourcen, etc. – eingeschränkte Möglichkeiten, das Potential mobiler IT zu nutzen: »KMUs stellen sich Hindernisse, dieses Potential zu nutzen« (Schritt 2)
- erste originäre Ergebnisse: Im Projekt werden in Vorbereitung auf den hier im Zentrum stehenden Projektteil der Erlanger Buchwissenschaft Modalitäten entwickelt, wie mobile Hardware, das Anwendungswissen um diese, möglicherweise einschließlich der eingesetzten Applikationen, von Großunternehmen / Erstanwendern an KMUs transferiert werden können: »Transfermodalitäten für mobile IT von Großunternehmen / Erstanwendern an KMUs« (Schritt 3)
- Projektteil der Erlanger Buchwissenschaft (Bereich Medien; exemplarisch: Unterbereich Print und digitale Entsprechungen):
 - Es wird der Zusammenhang hergestellt zwischen bereits identifizierten und dokumentierten Erkenntnissen zu möglichen und notwendigen Prozess- und Serviceverbesserungen im bearbeiteten Bereich einerseits und einem an den entsprechenden Stellen sinnvollen Einsatz mobiler IT andererseits: »Bereichsspezifische Szenarien, bei denen

¹ Die Erlanger Buchwissenschaft erarbeitet im Moment zusammen mit Steinröder Publishing Consulting, München eine umfragebasierte Studie über das Innovationsmanagement in Verlagen; die Ergebnisse dieser Studie, die sich bereits im fortgeschrittenen Stadium der Auswertung befindet, werden für das Teilprojekt der Erlanger Buchwissenschaft relevant sein.

mobile IT in identifizierten und dokumentierten Problembereichen bei Innovationen bzw. Prozess- und Serviceverbesserungen unterstützen kann« (Erlanger Buchwissenschaft, Teil 1)

- Es werden mit Partnern aus der Wirtschaft Feldversuche unter Anwendung der in Schritt 3 erarbeiteten Transfermodalitäten durchgeführt, um zu überprüfen, ob diese Modalitäten zielführend sind (und nicht zuletzt der dadurch ermöglichte Einsatz mobiler IT auch tatsächlich die vorhergesagten Effekte hat): »Der intelligenter Einsatz von IT, insbesondere mobiler IT, unterstützt nachprüfbar an den identifizierten und dokumentierten Ansatzpunkten für Prozess- und Serviceverbesserungen in der Medienbranche«; »Die entwickelten Modalitäten des Transfers funktionieren« (Erlanger Buchwissenschaft, Teil 2)

- Anschließend werden die Ergebnisse der Erlanger Buchwissenschaft (und ihrer Feldpartner) mit den Ergebnissen aus den zwei anderen abgedeckten Bereichen Bildung und Gesundheitsmanagement zusammengeführt und die in Schritt 3 erarbeiteten Transfermodalitäten ggf. entsprechend optimiert und zu einem verallgemeinerten, in der Praxis exemplarisch überprüften und anwendbaren Vorgehensmodell für eine Second-Product-Lifecycle-Strategie im Rahmen der Aufgabenstellung fortentwickelt: »Überprüfte, anwendbare Modalitäten für den Transfer mobiler IT auf KMUs im Rahmen einer Second-Product-Lifecycle-Strategie«



Grafik 1: „Second-Product-Lifecycle-Strategien für Hightech-Geräte zur Steigerung der Innovationsfähigkeit kleiner und mittelständischer Unternehmen sowie nichtstaatlicher Organisationen“ – Stränge des Projektes mit Fokus auf die Beiträge der Erlanger Buchwissenschaft

Schritt 1: »Der Einsatz mobiler IT fördert die Innovativität; damit zusammenhängende Chancen und Anforderungen für bzw. an Anwender«

In zu unserem vorgängigen Projektteilen stellen sich die Projektbeteiligten – auf der Basis der dazu breit vorhandenen Literatur – v.a. die Frage, welchen Einfluss mobile Hightech-Geräte auf die Innovationsfähigkeit von Unternehmen haben. Diese Frage setzt dabei bei Unternehmen an, die eine solche Technologie bereits im Einsatz haben. Ziel ist es, eine Wissensbasis über Erfolgs- und Misserfolgskriterien anhand von Fallstudien zu generieren. Diese Ergebnisse können von späteren Empfängern der Technologie genutzt werden, um damit selbst erfolgreich Innovationen anzustoßen.

Schritt 2: »KMUs stellen sich Hindernisse, dieses Potential zu nutzen«

In zu unserem vorgängigen Projektteilen stellen die Projektbeteiligten noch einmal die besonderen Restriktionen kleiner und mittlerer Unternehmen (sowie auch von sozialen Trägern) im Innovationsprozess zusammen. Diese werden nicht nur im Bereich fehlender Finanzmittel und Personalressourcen, sondern auch in Form eines Mangels an für den Einsatz notwendigem Wissen (in einem weiten Sinne) erwartet.

Schritt 3 / erste originäre Ergebnisse: » Transfermodalitäten für mobile IT von Großunternehmen / Erstanwendern an KMUs«

In zu unserem vorgängigen Projektteilen gehen die Projektbeteiligten der Forschungsfrage nach, wie der Transfer von Technologie und Wissen in einer Second-Product-Lifecycle-Strategie gestaltet werden kann. Hierbei geht es zum einen um den physischen Übertragungsprozess von Hightech. Zum anderen muss aber auch untersucht werden, wie Wissensressourcen erfolgreich transferiert werden können.

Im Detail stellen sich dabei u.a. die folgenden Fragen:

- Welche Treiber und Hemmnisse können im Hardware-Transferprozess festgestellt werden? Lassen sich dabei bestimmte Hightech-Klassen identifizieren, welche sich besonders gut für Second-Product-Lifecycle-Strategien eignen? Wie können Prozessmodelle für die effiziente Rücknahme und gezielte Überführung von Hightech-Geräten aussehen? – Aufbauend auf etablierten Prozessen und unter Berücksichtigung der neuen Strategie sollen hier Lösungen mit möglichst niedrigen Kosten entwickelt werden.
- Welche Treiber und Hemmnisse können im Wissens-Transferprozess identifiziert werden? Gibt es Hemmnisse von Seiten der Unternehmen, Auskunft über Lösungen zu geben? Welche Institutionen könnten gegebenenfalls KMUs dieses Wissen zur Verfügung stellen?
- Wie sehen Software-Applikationen aus, die einen erfolgreichen Wissenstransfer an das Zielunternehmen ermöglichen? – Dies spricht in besondere Weise Lösungen für ein unternehmensübergreifendes Knowledge Management an.

Erlanger Buchwissenschaft, Teil 1: »Bereichsspezifische Szenarien, bei denen mobile IT in identifizierten und dokumentierten Problembereichen bei Innovationen bzw. Prozess- und Serviceverbesserungen unterstützen kann«

Ziel des Gesamtprojektes ist es, überprüfte und anwendbare Vorgehensmodelle zu formulieren für die erfolgreiche Weitergabe von mobiler IT an KMUs. Aufgabe der Erlanger Buchwissenschaft und ihrer Feldpartner ist es dabei, die Entwicklung dieser Vorgehensmodelle mit ihrer Kenntnis der Branche, insbesondere der in ihr laufenden Diskurse und der über sie stattfindenden wissenschaftlichen Reflexion sowie zusätzlich a) mit Hilfe von Experteninterviews und einer Befragung in Branchenunternehmen zu den von diesen gesehenen Einsatzmöglichkeiten für mobile IT und b) mit Hilfe von der Buchwissenschaft begleiteten Feldversuchen in der Branche zu unterstützen. Aufgrund des Anspruches, dass die Ergebnisse (nämlich eben diese Transfermodalitäten) insgesamt weitgehend branchenunabhängig sind, stellt es keine unbillige Einschränkung dar, wenn die Erlanger Buchwissenschaft bei der Betrachtung des Bereiches Medien paradigmatisch den Unterbereich Print und digitale Entsprechungen wählt und konsequenterweise für die Feldversuche dann Szenarien aus eben diesem Bereich. Bei den gewählten Szenarien handelt es sich v.a. um eines aus dem Wertschöpfungsblock² „Inhalte erstellen“ und eines aus dem Wertschöpfungsblock „Inhalte distribuieren“; im ersten Falle werden wir Unterstützungsmöglichkeiten der redaktionellen Arbeit auf der Wirtschaftsstufe Verlag durch mobile Hardware betrachten, im anderen entsprechende Unterstützungsmöglichkeiten der Kundenberatung auf der Wirtschaftsstufe Buch(einzel)handel. Zu jedem dieser beiden Anwendungsszenarien gibt es einen Feldpartner; im ersten Fall ist das ein mittelständischer Verlag, im zweiten eine mittelständische, inhabergeführte wissenschaftliche Buchhandlung.

Wie in vielen anderen Branchen bzw. Teilbranchen gibt es auch im gewählten Bereich Print und digitale Entsprechungen Ansatzpunkte, die schon in der wissenschaftlichen Literatur und im Branchendiskurs (wie er z.B. in Fachzeitschriften³ geführt und dokumentiert wird) tendenziell einvernehmlich als solche identifiziert sind, an denen – aus verschiedenen Gründen – Prozess- und Serviceverbesserungen Not taten. Die Tatsache, dass im Rahmen des Gesamtprojektes sowohl Methoden der strategischen Prozess- und Serviceportfolioplanung als auch die operative Gestaltung, Einführung, Ausführung und das Controlling von Prozessen und Services abgedeckt sein sollen, schlägt sich im Bereich Print und digitale Entsprechungen darin nieder, dass die Erlanger Buchwissenschaft sowohl Ansätze zu Verbesserungen betrachtet und auswertet, die

² Zur Einteilung der spezifischen Wertschöpfungsbeiträge von Medienunternehmen in Haupt“blöcke“ vgl. z.B. Schumann/Hess 2006, S. 12.

³ Vgl. dabei zum Beispiel „Börsenblatt des Deutschen Buchhandels“, „BuchMarkt“ oder „buchreport“.

einen stärker operativen Fokus haben⁴, als auch solche, die im Zusammenhang mit der strategisch angeratenen Konzentration auf Prozess- und Servicebereiche zu sehen sind, die durch die starken Disintermediationstendenzen in der Branche weniger grundsätzlich in Frage gestellt und deshalb als Kernbereiche besonders zu entwickeln sind⁵. Für jeden der in wissenschaftlicher Literatur und / oder Branchendiskurs identifizierten Ansatzpunkte für Prozess- und Serviceverbesserungen werden wir überprüfen, inwieweit der Einsatz mobiler IT bei der „Implementation“ solcher Verbesserungen unterstützen kann.

Szenario 1 (Verlag, „strategisch“)

Eines der Szenarien, bei denen wir die Einsatzmöglichkeit von mobiler Hardware (sowie natürlich entsprechender Front- und Backend-Applikationen) bei Verlagen untersuchen werden, ist das folgende: Die mobile Hardware wird – gewissermaßen für eine möglicherweise zukunftsweisende Zwischenstufe zwischen professionellen Redakteuren / Reportern einerseits und Laien / Amateuren (wie sie im Zusammenhang mit dem sogenannten Web 2.0 eine zunehmende Rolle als Content-Lieferanten spielen) andererseits – an ausgewählte Nutzer eines Portals zu Berufswahl-Themen im Web ausgegeben. Die ausgewählten Nutzer sollen mit dieser mobilen Hardware aktuellen Reportage-Content „von unterwegs“ (z.B. von Events) liefern. Die konzeptionelle Besonderheit ist hier erstens, dass diese ausgewählten Nutzer vom Betreiber des Portals systematisch ausgebildet und gecached werden, und zweitens, dass der Betreiber systematisch Qualitätssicherung des gelieferten Contents betreibt; ein möglicher Begriff für dieses Konzept könnte „value-added mobile user generated content“ sein.

Szenario 2 (Buchhandlung, „operativ“)

Eines der Szenarien, bei denen wir die Einsatzmöglichkeit von mobiler Hardware (sowie natürlich entsprechender Front- und Backend-Applikationen) bei Buchhandlungen untersuchen werden, ist das folgende: Ein wissenschaftlicher Fachbuchhändler identifiziert zusammen mit einem Kunden beim Beratungsgespräch am Bücherregal in der Buchhandlung eine ganz bestimmte Seite in einem bestimmten, gerade begutachteten Buch als im Sinne des Kundenwunsches besonders einschlägig. Diese Seite wird nun mit dem PDA oder Smartphone des Buchhändlers fotografiert und die entstehende Datei anschließend einem Optical-Character-Recognition-Prozess unterzogen; anschließend wird dessen Ergebnis (ein ASCII-Text) mit Hilfe der Internet-Connectivity des Gerätes an eine Web-Applikation übermittelt. Diese Web-Applikation gleicht den Textausschnitt mit Hilfe von

⁴ Vgl. dazu z.B. für den Bucheinzelhandel (Marx 2006).

⁵ Vgl. dazu z.B. für das Verlagswesen (Bläsi 2006) oder den Vortrag „The end-to-end information chain could well bypass its supposed commander – publishing houses and the service revolution of the second order“, 2. Juni 2006, Università della Svizzera italiana, Lugano (BIT [Business and Information Technologies] International Conference 2006: Innovating Markets and Organizations Through Information and Communication Technologies – Academic Perspectives) desselben Autors.

Textähnlichkeits-Algorithmen mit anderen, in der Volltext-Datenbank VTO des Börsenvereins des Deutschen Buchhandels⁶ repräsentierten Büchern ab und identifiziert aufgrund des Ergebnisses dieses Abgleiches andere Bücher als „ähnlich“ (und damit ähnlich einschlägig im Sinne des Kundenwunsches). Diese Bücher kann der Buchhändler nun dem potentiellen Kunden (ebenfalls) zum Kauf anbieten – mit einem starken, nachprüfaren Argument.

Erlanger Buchwissenschaft, Teil 2: »Der intelligente Einsatz von IT, insbesondere mobiler IT, unterstützt nachprüfbar an den identifizierten und dokumentierten Ansatzpunkten für Prozess- und Serviceverbesserungen in der Medienbranche«; »Die entwickelten Modalitäten des Transfers funktionieren«

Für jede Einsatzmöglichkeit im Rahmen der Feldversuche ist anschließend erstens zu überprüfen, ob die mit dem Einsatz mobiler IT systematisch vorhergesagten Verbesserungen der angegangenen Prozesse und Services unter den gegebenen Umständen (KMU, Nutzung von Hardware in einem zweiten Teil ihres Produkt-Lebenszyklusses, etc.) tatsächlich eingetreten sind; zweitens ist zu überprüfen, ob die Anwendung der erarbeiteten Modalitäten zur Übertragung von Hardware (und Anwendungswissen, möglicherweise einschließlich Applikationen) von den Erstanwendern auf die an den Feldversuchen beteiligten KMUs im Bereich Medien die aufgrund der Ergebnisse von Schritt 3 angestrebte Effektivität und Effizienz erreicht hat.

Gesamtergebnis: »Überprüfte, anwendbare Modalitäten für den Transfer mobiler IT auf KMUs im Rahmen einer Second-Product-Lifecycle-Strategie«

Abschließend führen alle Projektpartner, einschließlich der Institute und Feldpartner, die in den Bereichen Bildung und Gesundheitsmanagement Ergebnisse erarbeitet haben, die analog zu denen sind, die die Erlanger Buchwissenschaft für den Bereich Medien erarbeitet wird, die nicht zuletzt auch in iterativen Schritten optimierten Erkenntnisse zusammen. Damit können branchenübergreifende, in verschiedenen exemplarischen Bereichen überprüfte und ggf. aufgrund dessen optimierte Modalitäten für den Transfer von mobiler Hardware und dazugehörigem Anwendungswissen von Erstanwendern auf kleine und mittlere Unternehmen ausgewiesen werden.

⁶ Vgl. dazu <http://www.volltextsuche-online.de>.

Literatur

Bläsi, Christoph (2006): Herausforderungen Wikipedia und Open Access – können Verlage etwas lernen von den Strategien angesichts Linux & Co.? In: Knowledge eXtended. Die Kooperation von Wissenschaftlern, Bibliothekaren und IT-Spezialisten. 3. Konferenz der Zentralbibliothek (2.–4.11.2005). Schriften des Forschungszentrums Jülich. Band 14. Jülich 2005, S. 73–89.

Habann, Frank (2003): Innovationsmanagement in Medienunternehmen. Theoretische Grundlagen und Praxiserfahrungen. Wiesbaden 2003.

Leonard-Barton, D. (1995): Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation, Boston.

Leifer, R.; McDermott, C.; Colarelli, G.; Peters, L.; Rice, M.; Veryzer, R. (2000): Radical Innovation – How Mature Companies Can Outsmart Upstarts, Boston, MA. Harvard Business School Press.

Marx, Gaby (2006): Profilbildung im stationären Sortiment als Chance im Wettbewerb. In: Sortimenter Ausschuss des Börsenvereins des Deutschen Buchhandels (Hg.): Informationen für den Buchhandel, Nr. 8. Web: http://www.boersenverein.de/de/69181?rubrik=69177&seite=30&dl_id=114426, 16.6.2006.

Tidd, J.; Bessant, J.; Pavitt, K. (2005): Managing Innovation. Integrating Technological, Market and Organizational Change, Wiley & Sons, Hoboken.

Schumann, Matthias / Hess, Thomas (2006): Grundfragen der Medienwirtschaft. Eine betriebswirtschaftliche Einführung. Berlin [u.a.] 2006.

Vahs, Dietmar / Burmester, Ralf (2006): Innovationsmanagement. Von der Produktidee zur erfolgreichen Vermarktung. Stuttgart 2005.

Herausforderung Primärdatenmanagement

PANGAEA® als vernetztes Verlags- und Bibliothekssystem für wissenschaftliche Daten

Michael Diepenbroek, Hannes Grobe

Abstract

Since 1992 PANGAEA® serves as an archive for all types of geoscientific and environmental data. From the beginning the PANGAEA® group started initiatives and aimed at an organisation structure which – beyond the technical structure and operation of the system – would help to improve the quality and general availability of scientific data. Project data management is done since 1996. 2001 the *World Data Center for Marine Environmental Sciences* (WDC-MARE) was founded and since 2003 – together with other German WDC – the group was working on the development of data publications as a new publication type. To achieve interoperability with other data centers and portals the system was adapted to global information standards. PANGAEA® has implemented a number of community specific data portals. 2007 – under the coordination of the PANGAEA® group – an initiative for networking all WDC was started. On the long-term ISCU plans to develop the WDC system into a global network of publishers and libraries for scientific data.

Abstract

Seit 1992 werden mit PANGAEA® wissenschaftliche Daten aus allen Bereichen der Geowissenschaften und Biologie archiviert. Von Beginn an wurden Initiativen gestartet und eine Organisationsstruktur angestrebt, welche über das technische System hinaus helfen die Qualität und Verfügbarkeit von wissenschaftlichen Daten zu verbessern. Seit 1996 wird intensiv Projektdatenmanagement betrieben. 2001 wurde das *World Data Center for Marine Environmental Sciences* (WDC-MARE) gegründet und ab 2003 zusammen mit den weiteren deutschen WDC an der Entwicklung der Datenpublikation als neuem Publikationstyp gearbeitet. Um eine Interoperabilität mit anderen Datenzentren und Datenportalen zu erreichen wurde das System gleichzeitig an globale Informationsstandards angepasst. PANGAEA® hat in der Folge selbst eine Reihe von community spezifischen Datenportalen implementiert. Koordiniert von der PANGAEA® Gruppe wurde 2007 eine Initiative zur Vernetzung der WDC gestartet. Langfristig ist vorgesehen das WDC System in einen globalen Verlags- und Bibliotheksverbund für wissenschaftliche Daten zu entwickeln.

Einleitung

Datenzentren, eine neuzeitliche Erfindung, entstanden zumeist mit der Motivation, wissenschaftliche Daten, die regelmäßig in einem gewissen institutionellen oder projektgebundenen Rahmen erhoben werden, langfristig zu sichern. So war das Geophysikalische Jahr 1957 Ausgangspunkt für die Gründung einer Reihe von global verteilten Welt Datenzentren, die sich der in diesem und in den nachfolgenden Jahren

produzierten wissenschaftlichen Daten annehmen und dauerhaft archivieren sollten. Das System der *World Data Center* (WDC) ist an den *International Council for Scientific Unions* (ICSU) gebunden und feierte in diesem Jahr seinen 50. Geburtstag. Seit einigen Jahren unterliegt das System einem zunehmenden Erneuerungsdruck. Die exponentiell steigende Datenflut und die Entwicklung des Internet führte zum Aufbau vieler neuer Daten haltender Systeme, eines davon ist das 1992 entstandene *Publishing Network for Geoscientific and Environmental Data* (PANGAEA® – www.pangaea.de). 2001 gründete die PANGAEA® Gruppe das ICSU *World Data Center for Marine Environmental Sciences* (WDC-MARE – www.wdc-mare.org). PANGAEA® war von Anfang an als breitspektral arbeitendes System gedacht. Die Heterogenität und Dynamik der Geowissenschaften und Biologie erforderte ein möglichst flexibles System zur Erfassung, Bearbeitung und Archivierung der vielfältigen Daten. Dennoch wurde bereits in der Aufbauphase klar, dass ein gutes technisches System zwar Grundvoraussetzung ist, nicht aber die prinzipiellen Probleme der Verfügbarkeit und Qualität wissenschaftlicher Daten lösen kann. Wissenschaftliche Primärdaten sind, neben den Publikationen, das zweite wichtige Ergebnis, das langfristig und in nachnutzbarer Form nach dem Prinzip des offenen Zugangs (DFG 1998, ESF 2000, Berliner Erklärung 2003, OECD 2004) verfügbar sein muss. In den 70er Jahren des vorigen Jahrhunderts war es noch üblich, Primärdaten direkt in einer Publikation oder im Anhang einer Zeitschrift abzdrukken. Mit steigenden Datenmengen und dem Übergang zur elektronischen Publikation wurde diese Praxis, primär aufgrund der Kosten, aufgegeben. Wissenschaftliche Verlage erlauben zwar die Ablage von zu einer Publikation gehörenden Primärdaten, die Archivierung folgt jedoch keinerlei Standards oder einheitlichen Strukturen und ist vom peer review ausgeschlossen, kann also nicht als Vorbild für eine allgemeine Lösung des Problems gelten.

Viele Datenzentren, auch ein guter Teil der ICSU WDC sind dem gegenüber technisch gut vorbereitet. Dennoch folgt auch hier die Archivierung zumeist keinen globalen Standards.

Die Trennung von wissenschaftlicher Publikation und zugrunde liegenden Primärdaten kann als gravierendes strukturelles Problem in den empirischen Wissenschaften gesehen werden. Nicht nur die Evaluierung einer Publikation auch die Nachnutzung der Ergebnisse ist erheblich eingeschränkt. Es gibt keine wirklich autorisierten und authentifizierten Orte für die langzeitliche Aufbewahrung von wissenschaftlichen Daten, keine Korrelation zwischen archivierten Daten und wissenschaftlicher Publikation und keine Vernetzung zwischen den Datenzentren. Gebraucht wird ein globales Verlags- und Bibliothekssystem für die Erfassung, Archivierung und Publikation wissenschaftlicher Daten. Hier spielen die ICSU WDC eine aktive Rolle. Die drei deutschen WDC (WDC-Climate, WDC-RSAT, WDC-MARE) und das GFZ haben – zusammen mit der Technischen Informationsbibliothek (TIB), Hannover - in den letzten drei Jahren ein praktikables System zur Publikation wissenschaftlicher Daten geschaffen (Schindler et al 2005, www.std-doi.de). WDC-MARE mit dem Informationssystem PANGAEA® und seinem redaktionellen System

kann dabei bereits jetzt als Referenz für ein Verlags- und Bibliothekssystem für wissenschaftliche Daten gesehen werden und soll im weiteren näher beschrieben werden.

Von der Datenerfassung zur Publikation

WDC-MARE / PANGAEA® wird als permanente Einrichtung vom Zentrum für Marine Umweltwissenschaften (MARUM) der Universität Bremen und der Stiftung Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI) in Bremerhaven betrieben. 3 Wissenschaftler sind für die grundsätzliche Organisation und Entwicklung zuständig. Ein Team von durchschnittlich 6-8 Wissenschaftlern übernimmt seit 1996 das Projektdatenmanagement (www.pangaea.de/projects). Die daraus gewonnenen Drittmittel tragen zu einem erheblichen Teil zur Finanzierung des Betriebes bei.

Datenerfassung, Qualitätssicherung, redaktionelle Bearbeitung und Archivierung

Die Akquise wissenschaftlicher Daten ist ein gravierendes Problem. Eigenen Schätzungen zufolge werden nur wenige Prozent der global produzierten wissenschaftlichen Daten auch in geeigneten Datenbanken langfristig archiviert. Es kommt selten vor, dass Daten freiwillig Datenzentren überlassen werden. Im institutionellen Rahmen gibt es seit einigen Jahren eine Aufbewahrungspflicht (DFG 1998). Ebenso sind viele wissenschaftliche Projekte bzw. Programme mit entsprechenden Auflagen versehen. Absprachen in solchen Kontexten erleichtern die Datenerfassung, können das Problem jedoch nicht vollständig beseitigen.

Als relativ effektiv hat sich Datenmanagement als finanzierter Bestandteil wissenschaftlicher Projekte erwiesen (<http://www.pangaea.de/projects/>). Diese Art der Projektförderung ist zunehmend in Verbundprojekten von BMBF und DFG zu beobachten; eine allgemeine Regelung steht jedoch noch aus. Auf der EU Ebene ist Datenmanagement seit den ersten Förderprogrammen wichtiges Kriterium für die Bewertung von Projekten (MAST Data Management Code). Projekte wie z.B. CARBOOCEAN (<http://www.carboocean.org/>), die eine verbesserte Quantifizierung von Kohlenstoffbilanzen im marinen Bereich zum Ziel haben, ist die möglichst vollständige Erfassung qualitätsgesicherter Daten notwendige Bedingung für den Erfolg des Projekts. Allgemein bauen komplexe und/oder großskalige Ansätze in der Global Change Forschung auf die Daten und Ergebnisse vieler kleiner Projekte und sind oft nur so realisierbar.

Die PANGAEA® Gruppe betreibt Projektdatenmanagement seit mehr als 10 Jahren. Dies ist die wichtigste Quelle für neu zu archivierende Daten, wesentlich bedingt durch die Nähe mit den Wissenschaftlern. Zudem trägt Projektdatenmanagement auch erheblich zur Finanzierung des Betriebes von PANGAEA® bei. Dies schafft Kapazitäten, die es der Gruppe ermöglichen auch nicht finanzierte Projekte wie z.B. die globale abschließende Erfassung und Publikation von Daten aus dem IGBP Project Joint Global Ocean Flux Studies (JGOFS) durchzuführen (Sieger et al 2005).

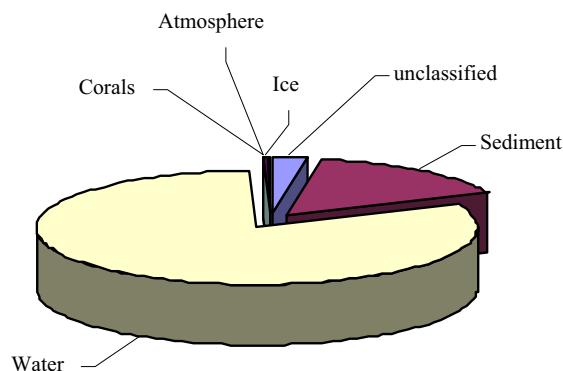
Unverzichtbarer Bestandteil des Datenmanagements ist die Qualitätssicherung. Kernaussage ist dabei, dass es weniger auf die Qualität der Daten ankommt als vielmehr, dass diese durch den Nachnutzer der Daten einschätzbar wird. Wichtig ist also die vollständige Erfassung von beschreibenden Informationen, den Metadaten, und die Einhaltung entsprechender Standards (s. nächster Abschnitt). Minimal zu beantwortende Fragen sind: Wer hat was wo wann und wie gemessen oder beobachtet. Darüber hinaus wird bei PANGAEA® standardmäßig die Validität der eingesetzten Methoden geprüft und ob z.B. die Präzision der Daten mit der eingesetzten Methodik korrespondiert. Ausreißer werden ermittelt und entsprechend markiert. Für die eigentliche Qualität der Daten steht der Datenproduzent mit seinem Namen und/oder seiner Institution.

Die Bearbeitung und Archivierung von Daten variiert im Allgemeinen mit den Datenzentren und Datentypen. Es gibt praktisch weder einheitliche Datenmodelle noch allgemein nutzbare redaktionelle Systeme. Üblich ist die Nutzung relationaler Datenbanken im Serverbereich, was insbesondere für die Metadaten eine gewisse Konsistenz gewährleistet.

Daten für ca. 30000

Parameter, z.B.:

Sediment- und Eis Profile
Seismische Profile
Atmosphärenprofile
Ozeanprofile
Mineralverteilungen
Geologische Karten
Bilder und Filme
Zeitreihen
Plankton und Fisch



Datensätze: 551 012

Datenobjekte: 1 823 824 468

Abbildung 1 Inhalte von PANGAEA®. Stand 9/2007

Das WDC-MARE mit dem Informationssystem PANGAEA® archiviert seit 1992 georeferenzierte Daten primär aus der Meeres- und Umweltforschung (Diepenbroek et al 2002). Derzeit sind ca. 550.000 Datensätze mit knapp zwei Milliarden Mess- und Observationsdatenpunkten verfügbar. Sie sind verknüpft mit ca. 30.000 unterschiedlichen Datentypen (Parameter, Variable), ca. 6000 Autoren (Datenurheber), knapp 6000 klassischen Publikationen und mehr als 300.000 Probenlokationen. Der jährliche Zuwachs beträgt durchschnittlich mehr als 20% des Gesamtbestands.

In PANGAEA® werden Daten und Metadaten systematisch über ein redaktionelles System erfasst und bearbeitet. Das System steigert erheblich die Effizienz bei der Pflege der Daten und hilft, Fehler zu vermeiden. Bei kleinen Datenzentren mit relativ spezialisierten Dateninhalten ist ein redaktionelles System meist noch entbehrlich. PANGAEA® war jedoch von Beginn an als System gedacht um alle Arten von geowissenschaftlichen und biologischen Daten zu archivieren.

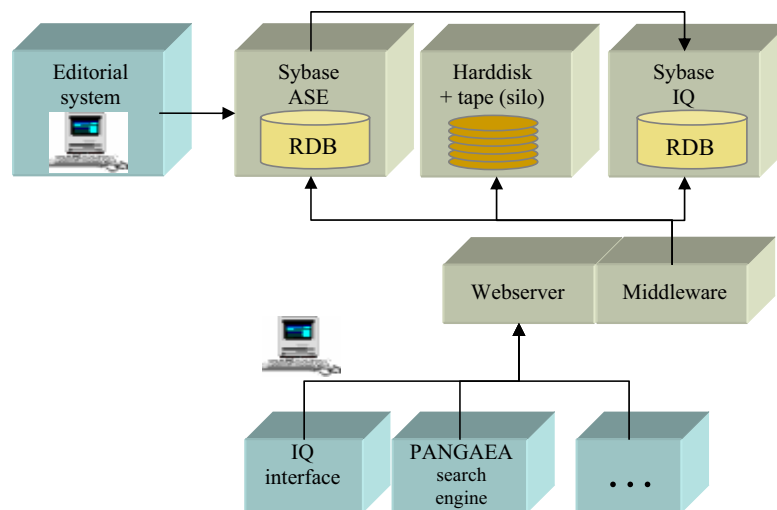


Abbildung 2 Technischer Aufbau von PANGAEA®

Für die Archivierung von Daten und Metadaten wird ein Relationales Datenbank Managementsystem (RDBMS - Sybase) genutzt. Backups werden an physikalisch unterschiedlichen und gesicherten Orten abgelegt. Damit ist der Datenbestand weitestgehend vor Verlust geschützt. Abbildung 2 zeigt vereinfacht den technischen Aufbau von PANGAEA®. Massendaten, wie z.B. Geophysik, und binäre Objekte, wie z.B. Bilder werden auf einem Festplattenarray abgelegt und migrieren ggf. in ein Bandsilo. Sybase IQ ist ein Datenwarenhause, in dem sämtliche numerischen und textuellen Datenwerte gespiegelt werden. Das Datenwarenhause ermöglicht performante retrievals von beliebigen raumzeitlichen Ausschnitten. Die Metadaten werden als Kompilat an das Suchergebnis angehängt.

Datenpublikation

Zusammen mit weiteren WDC in Deutschland hat die PANGAEA® Gruppe innerhalb des DFG-Projekts „Zitierfähigkeit und Publikation wissenschaftlicher Primärdaten“ in den letzten 3 Jahren das Konzept zur Publikation wissenschaftlicher Daten entwickelt und prototypisch implementiert (Schindler et al 2005, Klump et al 2006). Gemeinsam wurden allgemeine Anforderungen an diesen neuen Publikationstyp untersucht:

- Der formale Aufbau der Publikation, d.h. welche beschreibenden Elemente sind notwendig und welche optional, wie müssen diese gestaltet werden, welche Datenformate und Standards sind sinnvoll?
- Die Granularität von publizierbaren Datensätzen
- Die Notwendigkeit „peer review“ ähnlicher Verfahren zur Qualitätssicherung
- Die Anforderungen an die Datenarchive bezüglich Langzeitverfügbarkeit und dauerhafter eindeutiger Referenzierbarkeit von publizierten Datensätzen mittels persistenten Identifizierern wie dem ‚Digital Object Identifier‘ (DOI). Für die Zertifizierung von Datenzentren wurde neben dem eigenen Erfahrungshintergrund auch wesentlich auf das OAIS Referenzmodell (NASA 2002) und die Ergebnisse des BMBF Projekts NESTOR (<http://www.langzeitarchivierung.de/>) zurückgegriffen.

Die Ergebnisse wurden in den beteiligten Einrichtungen als Richtschnur genutzt, um einerseits die Strukturen der Datenarchive entsprechend anzupassen, andererseits für neue Datensätze und die Pflege der vorhandenen Datensätze redaktionelle Organisationsschemata zu entwickeln. Derartige Schemata wurden in allen Einrichtungen prototypisch verwirklicht, sind jedoch – nach Einrichtung verschieden – erst mehr oder weniger vollständig in die technische Umgebung und den wissenschaftlichen Arbeitsprozess integriert worden.

Hier ist insbesondere das oben beschriebene Problem der Granularität zu nennen. In der bisherigen Arbeit haben sich die beteiligten Einrichtungen auf ein einfaches Modell geeinigt, indem individuelle publizierbare Datenentitäten im wissenschaftlichen Archiv zu zitierfähigen Datenentitäten zusammengefasst werden. Zitierfähige Datenentitäten repräsentieren die Schnittstelle zwischen Datenarchiv und wissenschaftlicher Literatur. Einerseits können Datensätze als publizierbare Entität direkt zitierfähig sein (wie Beispiele aus der Erdsystemforschung und aus dem medizinischen Bereich zeigen), andererseits kann es sein, dass man Datensätze – abhängig von der Nutzung – auf unterschiedlichen Ebenen zitieren möchte. So treten z.B. Fälle auf, bei denen man auf eine komplette Datensammlung verweisen möchte und andere Fälle, bei denen man auch einzelne Datensätze aus einer Sammlung zitieren möchte. Hier wird es notwendig sein, unterschiedliche, von den Datentypen abhängige Modelle zu entwickeln, um eine flexiblere Nutzung und Zitierweise zu ermöglichen.

Als weiteres Problem kommt hinzu, dass für die großen heterogenen Datenarchive (wie WDC-MARE, GFZ) ein beträchtlicher initialer Aufwand nötig ist, um die in mehr als 10 Jahren gewachsenen Bestände entsprechend zu bearbeiten und bereits registrierte Datensätze als zitierfähig auszuweisen oder zu zitierfähigen Einheiten zusammenzufassen. Jeder zu publizierende Datensatz erfordert Rücksprache mit den Datenurhebern oder weiteren Wissenschaftlern aus dem jeweiligen Forschungsumfeld und ggf. manuelle Korrekturen der Metadaten. In allen Einrichtungen wurden erste Ansätze eines „peer reviews“ erprobt.

Es zeigt sich, dass die verschiedenen Datenwelten einen sehr unterschiedlichen Aufwand beinhalten. Im WDC-C sind eine überschaubare Anzahl an relativ

homogenen aber extrem hochvolumigen Simulationsdatensätzen archiviert, WDC-MARE und GFZ stellen dagegen eine große Anzahl an äußerst heterogenen Observations- und Messdaten zur Verfügung, die im GFZ zudem noch über mehrere Datenbanken verteilt sind.

Die bisher nur in Ansätzen bewältigte Aufarbeitung der „Altbestände“ im Sinne des neuen Review- und Publikationsprozesses und die noch wenig etablierten Arbeitsabläufe führten dazu, dass zwar nahezu alle Datensätze in der DOI registry an der TIB erfasst sind, bislang jedoch weniger als 1000 von den Autoren zertifizierte und damit zitierfähige Einträge in TIBORDER, dem Online-Katalog der TIB, sichtbar sind. Nichtsdestotrotz sind alle Datensätze qualitätsgeprüft, vollständig mit Metadaten annotiert, direkt über die DOI zugreifbar und unterliegen denselben Anforderungen an die Langzeitarchivierung wie die von den Autoren zertifizierten und zitierfähigen Einträge. Künftig wird es jedoch nicht mehr zulässig sein, Datensätze ohne eine eindeutige Zuordnung zu einer zitierfähigen Datenentität zu registrieren. Die entsprechende Aufarbeitung der bereits registrierten und neu hinzukommenden Datensätze wird als langfristige Eigenleistung weitergeführt.

Auf diesem Hintergrund waren zu Beginn des Datenpublikationsprojektes vor einigen Jahren die vordringlichsten Aufgaben, (1) die Überarbeitung der Metadatenstrukturen und -bestände und (2) die technische Anpassung des vorhandenen Systems zur redaktionellen Bearbeitung von Daten und Metadaten.

Nach einer zügigen Anpassung der Metadatenstrukturen wurden eine Reihe von Routinen geschrieben, mit deren Hilfe ein großer Teil der für die Datenpublikation relevanten Metadaten harmonisiert wurden. So wurden damit z.B. weitgehend Doppeleinträge von Autoren beseitigt und die Schreibweise vereinheitlicht. Die Korrektur von fraglichen und fehlenden Einträgen, sowie z.B. auch die Vergabe und Korrektur von Titeln mussten händisch nachgeführt werden. Der Prozess wurde 2006 nahezu abgeschlossen.

Für den Import neuer Daten und die Metadatenbearbeitung nutzt PANGAEA® ein in mehr als 10 Jahren gewachsenes, global nutzbares System (client/server), welches die manuellen Arbeitsschritte auf ein Minimum reduziert. Die Anpassung dieses Systems an ein redaktionelles System zur Datenpublikation ist ein laufender iterativer Prozess, an dem sowohl die Systemmanager als auch die mit PANGAEA® arbeitenden Datenkuratoren intensiv beteiligt sind. So war es neben einer Vielzahl von Anpassungen auch notwendig, einen zeitlichen Ablauf einzuarbeiten. Neu importierte Datensätze werden danach nicht sofort registriert, sondern sind in einem Zeitraum von 28 Tagen noch änderbar bzw. vollständig ersetzbar. Nach Ablauf dieser Zeit werden sie automatisch registriert, ggf. als zitierfähig ausgewiesen und sind dann bis auf einige beschreibende Elemente statisch. Die Registrierung wurde flexibel in die bestehende Infrastruktur integriert (Abb. 2).

Insgesamt ist die für ein Publikationssystem notwendige Umstellung weitgehend abgeschlossen, so dass der redaktionelle Arbeitsaufwand – abgesehen von der vermehrten Kommunikation – insgesamt etwa gleich geblieben ist; ein nicht unwesentlicher Aspekt bezogen auf die laufenden Betriebskosten. Unvollständig ist

jedoch noch die Ausweisung von Datensätzen als zitierfähige Datensätze, primär wegen der überwiegend noch fehlenden Zertifizierung der Datenpublikationen durch die Autoren aber auch wegen des noch nicht ausgearbeiteten Reviewverfahrens und des mehrfach angesprochenen Granularitätsproblems. Mit Blick auf die Akzeptanz von Datenpublikationen und die vom WDC-MARE zu erwartende Menge zitierfähiger Datensätze erfordert das weitere Vorgehen jedoch Zeit und Sensibilität. Beispiele zitierfähiger Datensätze sind:

<http://doi.pangaea.de/10.1594/PANGAEA.472287>

<http://doi.pangaea.de/10.1594/PANGAEA.472492>

<http://doi.pangaea.de/10.1594/PANGAEA.472492>

Standards, Vernetzung und Portale

Die Vernetzung von Datenproduzenten, -archiven, und -konsumenten – allgemein als Geodateninfrastrukturen bezeichnet – ist eine notwendige Vorbedingung für komplexe oder großskalige Datenkompilationen. *Global Spatial Data Infrastructures* (GSDI - <http://www.gsdi.org/>) ist eine Vision, die in vielerlei Hinsicht erheblich zur Effizienz bei der Zusammenstellung und Auswertung von wissenschaftlichen Daten beitragen würde. Mit der Zustimmung von mehr als 60 Ländern zum 10-Jahres Implementierungsplans eines *Global Earth Observing System of Systems* (GEOSS) der *Group on Earth Observations* (GEO - <http://www.earthobservations.org/>) wurde 2005 erstmals auf der ministeriellen Ebene ein wirksamer Rahmen geschaffen, der die Vernetzung und Bereitstellung von wissenschaftlichen Daten fördern soll. Nichtsdestotrotz ist GEOSS aufgrund fehlender Ressourcen auf existierende Kapazitäten und Aktivitäten angewiesen. Während des diesjährigen Treffens der WDC Direktoren wurde daher eine Initiative zur Vernetzung der WDC beschlossen. Dies wird von den WDC nicht nur als sinnvoller Beitrag zu GEOSS gesehen sondern auch als erster Schritt zur Schaffung eines globalen Verbunds von Datenbibliotheken. Die Vernetzungsinitiative wird von der PANGAEA® Gruppe koordiniert.

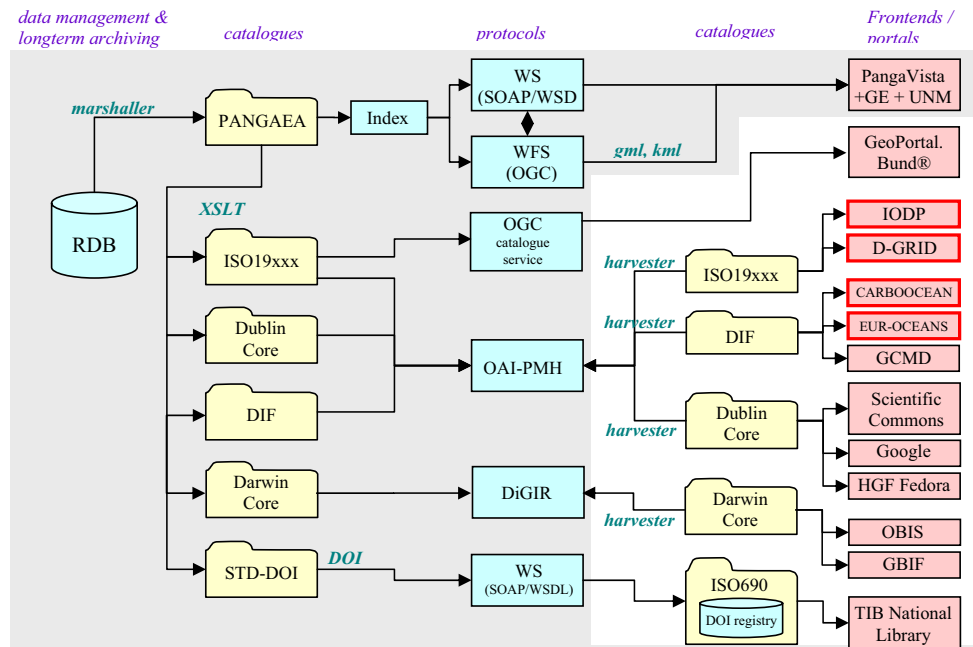


Abbildung 3 Metadaten-Infrastruktur von PANGAEA®. Grau hinterlegte Teile gehören vollständig zur internen Infrastruktur von PANGAEA®. Rot umrandet sind Portale, die weitestgehend von der PANGAEA® Gruppe implementiert wurden.

Die Gruppe hat in den letzten 5 Jahren systematisch an der Vernetzungsfähigkeit des Systems gearbeitet und stellt mittlerweile eine Vielzahl unterschiedlicher, allgemein nutzbarer Dienste zur Verfügung. Die Implementierung internationaler Informations- und Geodatenstandards spielt dabei eine wesentliche Rolle. Wichtige Instanzen sind hierbei ISO, OGC und W3C. Auf der Metadatenebene werden für jeden Datensatz die in der RDB gespeicherten Informationen über eine ‚marshalling‘ Routine zunächst in einem proprietärem Metadatenschema gesammelt und dann per XSLT in die diversen Inhaltsstandards übertragen. Dazu gehören z.B. ISO19115 als mittlerweile wichtigster Standard und das ‚Directory Interchange Format‘ (DIF - <http://gcmd.nasa.gov/User/difguide/>). Wichtige Protokolle sind der OGC Catalogue Service (<http://www.opengeospatial.org/standards/cat>) und das Open Archives Initiatives Protocol for Metadata harvesting (OAI-PMH - http://de.wikipedia.org/wiki/Open_Archives_Initiative). Letzteres ist relativ einfach zu implementieren und findet breite Verwendung nicht nur in der Bibliothekswelt, sondern zunehmend auch bei Datenzentren. Einen Überblick gibt Abbildung 3.

So hat die PANGAEA® Gruppe eine Reihe von community- und projektspezifischen Portalen implementiert. Das Portalframework ist generisch und basiert auf den Komponenten harvester, Indizierer mit Suchmaschine (Lucene - <http://lucene.apache.org/java/docs/>) und API (Schindler et al 2007,

<http://www.panfmp.org/>). Beispiele sind das Portal für das International Ocean Drilling Program (IODP - <http://sedis.wdc-mare.org/>) und für die EU Projekte EUR-OCEANS (<http://www.eur-oceans.eu/integration/wp2.2/>) und CARBOOCEAN (<http://dataportal.carboocean.org/>)

Da informationstechnische Entwicklungen einer erheblichen Dynamik unterliegen, hat sich die PANGAEA® Gruppe bewusst darauf eingestellt, eine Vielzahl von Standards bedienen zu können.

Z.Zt. beschränkt sich der standardisierte Informationsaustausch jedoch noch fast ausschließlich auf Metadaten. Dies liegt vor allem an dem Fehlen praktikabler und allgemein akzeptierter Datenaustauschformate. Lediglich für das Deutsche Community GRID C3 (<http://www.c3grid.de/>) werden von PANGAEA® auch Daten zur Verfügung gestellt.

Im Bereich des Datenaustausches sind GRID-Architekturen üblich, die derzeit jedoch noch sehr auf spezielle Datentypen und Dienste fokussiert sind. Für allgemeine und leicht zu implementierende Architekturen besteht noch Entwicklungsbedarf. So ist ein spezielles Problem die Verwendung einheitlicher Begriffe in den Metadaten zur Steuerung von Applikationen. Hier fehlen ebenfalls internationale Standards. Über die Europäischen Initiative INSPIRE (<http://de.wikipedia.org/wiki/INSPIRE>) besteht die Hoffnung, dass solche Vokabularien aufgebaut werden. Dies ist jedoch als langfristige Aufgabe zu sehen.

Fazit

Mit seinen langfristig angelegten und gesicherten Archivstrukturen, dem hocheffizienten redaktionellen System und der weitgehenden Interoperabilität mit Datenportalen und anderen Datenzentren hat sich PANGAEA® als exemplarisches Verlags- und Bibliothekssystem für wissenschaftliche Daten entwickelt.

Der im deutschen WDC Konsortium entwickelte und innerhalb von PANGAEA® realisierte Ansatz zur Publikation wissenschaftlicher Daten geht weit über die z.B. von der Human Genom Community praktizierte Kopplung von Datenarchivierung und wissenschaftlichem Aufsatz hinaus. Es ermöglicht Wissenschaftlern Datensätze eigenständig zu publizieren. Jede Datenpublikation wird mit einem aussagekräftigen Zitat und einem persistenten Identifizierer (DOI) versehen und erlaubt somit eine verlässliche Referenzierung. Die Zitierfähigkeit ist ein starkes Motiv für Datenproduzenten, Daten zu veröffentlichen und wird somit langfristig zu einer Verbesserung von Datenqualität und –verfügbarkeit führen.

Das Konzept stößt bei den Datenproduzenten auf breite Resonanz, dennoch wird erwartet, dass es noch Jahre brauchen wird, bevor diese neue Form der Publikation in der Wissenschaft allgemein akzeptiert wird.

Die in den Deutschen WDC entstehenden Referenzsysteme bedürfen einer Extrapolation in das globale Umfeld. Mit der Vernetzung der ICSU WDC ist ein erster Schritt für die Entstehung eines globalen Bibliotheksverbunds für wissenschaftliche Daten getan. Ein solcher Verbund wäre transdisziplinär und bietet den Vorteil, dass

sämtliche Daten im Sinne des offenen Zugangs ohne Restriktionen zur Verfügung gestellt werden.

Um einerseits Langzeitsicherung wissenschaftlicher Daten zu gewährleisten, andererseits die Arbeit der Datenzentren weiter zu entwickeln in Richtung auf Standards sowohl für die Bearbeitung, Archivierung und Publikation als auch die Dateninfrastruktur, d.h. Interoperabilität von Datenzentren, fehlt jedoch bislang ein gesicherter Rahmen. Die Revision des ICSU WDC Systems kann dabei nur eine konzeptionelle Vorgabe leisten. Ein Langzeitbetrieb erfordert die Absicherung durch nationale und internationale Verträge. Bislang bestehen für die WDC keine justiziablen Verpflichtungen. Dies gilt auch für PANGAEA®.

Literatur

Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen (2003),
http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf

DFG (1998) Empfehlungen der Kommission "Selbstkontrolle in der Wissenschaft",
http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf

Diepenbroek M, Grobe H, Reinke M, Schindler U, Schlitzer R, Sieger R, Wefer G (2002) PANGAEA — an information system for environmental sciences. *Computers & Geosciences* 28: 1201–1210

ESF (2000) Good scientific practice in research and scholarship,
http://www.esf.org/typo3conf/ext/naw_securedl/secure.php?u=0&file=fileadmin/be_user/CEO_Unit/Science_Policy/ESPB10.pdf&t=1182214053&hash=911cf3e6d783883eb9b83ee634c36d9a

Klump J, Bertelmann R, Brase J, Diepenbroek M, Grobe G, Höck H, Lautenschlager M, Schindler U, Sens I, Wächter J (2006) Data publication in the open access initiative, *Data Science Journal*, Vol. 5 (2006) pp.79-83,
http://www.jstage.jst.go.jp/article/dsj/5/0/5_79/_article

NASA Consultative Committee for Space Data Systems (2002) Reference Model for an Open Archival Information System (OAIS) -
<http://public.ccsds.org/publications/archive/650x0b1.pdf>

OECD (2004) Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique,
http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

Schindler U, Brase J, Diepenbroek M (2005) Webservices infrastructure for the registration of scientific primary data. In: Rauber A et al. (Eds.): ECDL 2005, LNCS 3652, Springer-Verlag Berlin Heidelberg: 128–138

Schindler U, Diepenbroek, M (2007) Generic Framework for Metadata Portals. Computers & Geosciences, submitted.

Sieger R, Grobe H, Diepenbroek M, Schindler U & Schlitzer R (Eds) (2005) International Collection of JGOFS - Volume 2: Integrated Data Sets (1989-2003), WDC-MARE Reports 0003 (2005), ISSN 1611 – 6577, http://www.wdc-mare.org/reports/wdc-mare_0003.pdf

Zitierfähige Datensätze: Primärdaten-Management durch DOIs

Jan Brase, Jens Klump

Zusammenfassung

Der Zugang zu wissenschaftlichen Primärdaten ist eine grundlegende Voraussetzung für die Forschungsarbeit vor allem in den Naturwissenschaften. Im wissenschaftlichen Bereich besteht zwar grundsätzlich Bereitschaft, Daten für eine interdisziplinäre Nutzung zur Verfügung zu stellen, aber es ist zurzeit unüblich, dass die erforderliche Mehrarbeit für Aufbereitung, Kontextdokumentation und Qualitätssicherung im Wissenschaftsbetrieb anerkannt wird.

Projektdateien sind breit über Forschungsinstitute verstreut und werden von Wissenschaftlern erhoben und meist selbst verwaltet. Aufgrund der fehlenden Anerkennung der mit der Aufbereitung verbundenen Arbeit sind Projektdateien häufig schlecht dokumentiert und somit schwer zugänglich sowie nicht langfristig gesichert. Große Datenbestände bleiben ungenutzt, da sie nur einen kleinen Kreis von Wissenschaftlern bekannt und zugänglich sind. Viele Primärdaten bleiben ungenutztes Rohmaterial und sind häufig nach wenigen Jahren verloren.

Beispielhaft für den Bereich der Geowissenschaften werden seit Ende 2004 von der *Technische Informationsbibliothek* (TIB) Primärdatensätze registriert. Die Datensätze selber verbleiben bei den lokalen Datenzentren. Die TIB vergibt für jeden Datensatz einen DOI Namen als dauerhaften Identifizierer, wodurch der Datensatz weltweit über jeden Webbrowser auf einfachste Weise zugänglich wird. Zusätzlich werden für Sammlungen oder Auswertungen von Primärdatensätzen die Metadatenbeschreibungen in den Katalog der TIB aufgenommen. Diese Beschreibungen enthalten alle Angaben, die nach ISO 690-2 zur Zitierung elektronischer Medien verlangt werden.

Der Datensatz wird somit eine eigene zitierfähige Einheit. Mittlerweile wurden über dieses System über 450.000 Datensätze mit einer DOI Namen versehen und zitierfähig gemacht.

Abstract

Especially in nature science, access to scientific primary data is fundamental for research. In the scientific community there is a basic willingness to publish primary data for an inter-disciplinary use, but there is no acknowledgement however for the extra amount of work for editing, documentation or quality assurance.

Project data are widely spread over research institutions and mostly collected and managed by single researchers. The lack of acknowledgement for "good" data preparation leads to poor documented and almost inaccessible data sets with no persistency. Large amounts of data remain unused or known only to a small group of

scientists and therefore are unused raw material that disappears after a small number of years.

Starting with the field of earth-science the German National Library of Science and Technology (TIB) has started to register scientific data sets. The data itself remains at the local data centers. The TIB awards a DOI name for every data set as a persistent identifier, making the data accessible through every web browser worldwide in an easy way. Collections or analysis of data are furthermore included into the library of the TIB with a metadata description including all elements necessary for the citing of electronic media as defined in ISO 690-2 making the datasets therefore a citable entity.

The system was used to register over 450.000 datasets with a DOI name so far.

Hintergrund

„The rapid advances in digital technologies and networks over the past two decades have radically altered and improved the ways that data can be produced, disseminated, managed and used, both in science and in all other spheres of human endeavour. [...] This has created unprecedented opportunities for accelerating research and creating wealth based on the exploitation of data as such. [...] New software tools help to interpret and transform the raw data into unlimited configurations of information and knowledge. And the most important and pervasive research tool of all, the Internet, has collapsed the space and time in which data and information can be shared and made available, leading to entirely new and promising modes of research collaboration and production.”

Uhlir, Paul F. (2003) The Role of Scientific and Technical Data and Information in the Public Domain, National Academic Press, Washington DC (Uhlir, 2003)

Der Zugang zu wissenschaftlichen Primärdaten ist eine grundlegende Voraussetzung für die Forschungsarbeit vor allem in den Naturwissenschaften. Deshalb ist es notwendig, bestehende und zum Teil auch neu aufkommende Einschränkungen bei der Datenverfügbarkeit zu vermindern.

Im wissenschaftlichen Bereich besteht zwar grundsätzlich Bereitschaft, Daten für eine interdisziplinäre Nutzung zur Verfügung zu stellen, aber es ist zur Zeit unüblich, dass die erforderliche Mehrarbeit für Aufbereitung, Kontextdokumentation und Qualitätssicherung im Wissenschaftsbetrieb anerkannt wird. Die klassische Form der Verbreitung wissenschaftlicher Ergebnisse ist ihre Veröffentlichung in Fachzeitschriften, normalerweise ohne Veröffentlichung der zugrunde liegenden Daten. Diese klassische Publikation wird im "Citation Index" erfasst. Dieser Index wird zur Leistungsbewertung von Wissenschaftlern herangezogen. Datenveröffentlichungen werden darin bisher nicht berücksichtigt.

Projektdateien sind breit über Forschungsinstitute verstreut und werden von Wissenschaftlern erhoben und meist selbst verwaltet. Aufgrund der fehlenden Anerkennung der mit der Aufbereitung verbundenen Arbeit sind Projektdateien häufig schlecht dokumentiert und somit schwer zugänglich sowie nicht langfristig gesichert.

Große Datenbestände bleiben ungenutzt, da sie nur einen kleinen Kreis von Wissenschaftlern bekannt und zugänglich sind. Viele Primärdaten bleiben ungenutztes Rohmaterial und sind häufig nach wenigen Jahren verloren.

Die Diskussion um die Fälschung wissenschaftlicher Ergebnisse führte zur Empfehlungen in den ‚Regeln guter wissenschaftlicher Praxis‘ durch die DFG. Die Regeln beinhalten auch Richtlinien für den Datenzugang. Primärdaten einer Veröffentlichung müssen mindestens 10 Jahre gespeichert und zugänglich sein, um eine Prüfung der Ergebnisse zu ermöglichen. Zwar werden diese Vorschriften im Regelfall eingehalten, aufgrund der damit verbundenen Zeitbelastung werden Daten aber normalerweise nur in Rohform archiviert und nicht in ihrer Feinstruktur aufgearbeitet, dokumentiert und allgemein verfügbar erschlossen.

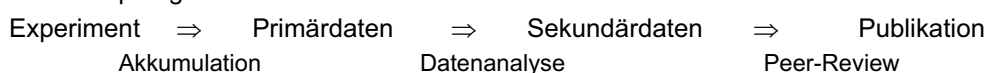
Die wichtigsten Ziele eines neuen Umgangs mit Primärdaten sind also langfristige und allgemein zugängliche Verfügbarkeit und Speicherung. Durchsetzbar ist dies am besten über eine persönliche Motivation der Wissenschaftler. Dies ließe sich durch die Möglichkeit der Primärdatenpublikation fördern:

Daten sollen nicht mehr ausschließlich Teil einer wissenschaftlichen Veröffentlichung sein, sondern eine eigenständige Identität besitzen. Damit würden Primärdaten, ähnlich wie Zeitschriftenartikel zitierbar.

Mit der anerkannten Datenpublikation erhält ein Autor also eine zitierfähige Veröffentlichung. Zeitschriftenartikel, welche die Daten verwenden, verweisen auf die Datenpublikation. Umgekehrt kann auch von den publizierten Daten auf Artikel in Zeitschriften verwiesen werden, die den Datensatz verwenden. Die Publikation von Daten kann also sinngemäß in das bestehende System von wissenschaftlichen Veröffentlichungen und deren Zitierbarkeit eingebunden werden.

Primärdatenregistrierung

Traditionell sind Primärdaten eingebettet in einen singulären Forschungsprozess, ausgeführt von einer definierten Gruppe von Forschern, geprägt von einer linearen Wertschöpfungskette:



Durch die Möglichkeiten der neuen Technologien und des Internets können einzelne Bestandteile des Forschungszyklus in separate Aktivitäten aufgeteilt werden (Daten-Sammlung, Daten-Auswertung, Daten-Speicherung, usw.) die von verschiedenen Einrichtungen oder Forschungsgruppen durchgeführt werden können. Die Einführung eines begleitenden Archivs und die Referenzierung einzelner Wissenschaftlicher Inhalte durch persistente Identifikatoren wie einen DOI Namen schafft die Möglichkeit anstelle eines linearen Forschungsansatzes, den Wissenschaftlerarbeitsplatz einzubinden in einen idealen Zyklus der Information und des Wissens (Abb. 1), in dem durch Zentrale Datenarchive als Datenmanager Mehrwerte geschaffen werden können und so für alle Datennutzer, aber auch für die Datenautoren selber ein neuer Zugang zu Wissen gestaltet wird.

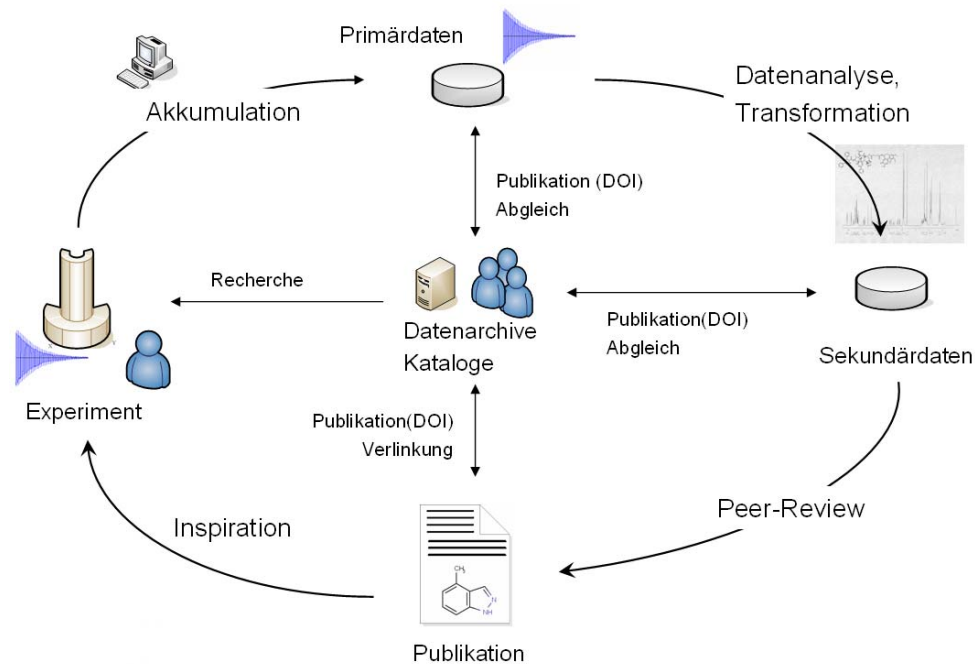


Abbildung 1: Ein idealer Zyklus der Information und des Wissens

Die TIB als DOI Registrierungsagentur

Der Digital Object Identifier (DOI) wurde 1997 eingeführt, um Einheiten geistigen Eigentums in einer interoperativen digitalen Umgebung eindeutig zu identifizieren, zu beschreiben und zu verwalten. Verwaltet wird das DOI-System durch die 1998 gegründete International DOI Foundation (IDF, <http://www.doi.org/>).

Der DOI Name ist ein dauerhafter persistenter Identifizierer, der zur Zitierung und Verlinkung von elektronischen Ressourcen (Texte, aber Primärdaten oder andere Inhalte) verwendet wird. Über den DOI Namen sind einer Ressource aktuelle und strukturierte Metadaten zugeordnet.

Ein DOI Name unterscheidet sich von anderen, gewöhnlich im Internet verwendeten Verweissystemen wie der URL, weil er dauerhaft mit der Ressource als Entität verknüpft ist und nicht lediglich mit dem Ort, an dem die Ressource platziert ist.

Der DOI Name identifiziert eine Entität direkt und unmittelbar, also nicht eine Eigenschaft des Objekts (eine Adresse ist lediglich eine Eigenschaft des Objekts, die verändert werden und dann ggf. nicht mehr zur Identifikation des Objekts herangezogen werden kann).

Das IDF System besteht aus der „International DOI Foundation“ selbst, der eine Reihe von Registrierungsagenturen („Registration Agencies “) zugeordnet sind . Für die Aufgaben einer RA können sich beliebige kommerzielle oder nicht kommerzielle Organisationen bewerben, die ein definiertes Interesse einer Gemeinschaft vorweisen können, digitale Objekte zu referenzieren.

Der DFG-Ausschuss „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ hat 2004 ein Projekt gestartet, um den Zugang zu wissenschaftlichen Primärdaten zu verbessern. Aus diesem Projekt heraus ist die TIB seit Mai 2005 weltweit erste DOI-Registrierungsagentur für wissenschaftliche Daten. Beispielhaft für den Bereich der Geowissenschaften werden Primärdatensätze registriert. Die Datensätze selber verbleiben bei den lokalen Datenzentren und die TIB vergibt für jeden Datensatz einen DOI Namen als dauerhaften Identifizierer, wodurch der Datensatz weltweit über jeden Webbrowser auf einfachste Weise zugänglich wird.

Der Datensatz wird somit eine eigene zitierfähige Einheit. Mittlerweile wurden über dieses System über 450.000 Datensätze mit einer DOI versehen und zitierfähig gemacht. Die Metadatenbeschreibungen der Datensätze werden zentral an der TIB gespeichert. Diese Beschreibungen enthalten alle Angaben, die nach ISO 690-2 zur Zitierung elektronischer Medien verlangt werden.

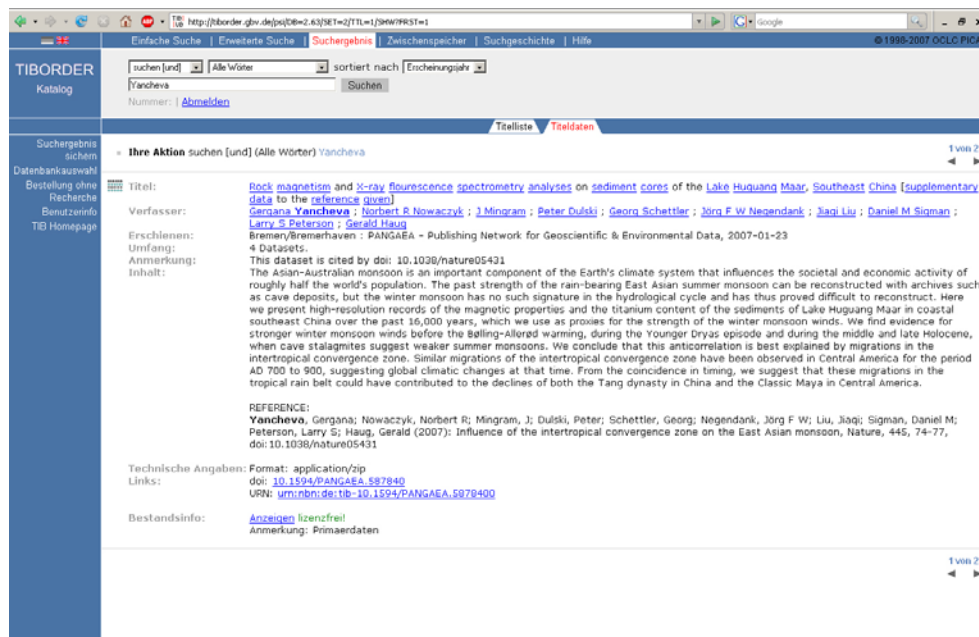


Abbildung 2: Anzeige eines Primärdatensatzes im Online-Katalog der TIB Hannover

Zusätzlich werden Sammlungen oder Auswertungen von Primärdatensätzen auch in den Katalog der TIB aufgenommen. Die Anzeige eines Primärdatensatzes im Katalog der TIB sehen sie in Abbildung 2.

Die DOI Registrierung erfolgt bei der TIB immer in Kooperation mit lokalen Datenzentren als sog. „Publikationsagenten“, Einrichtungen die weiterhin für Qualitätssicherung und die Pflege und Speicherung der Inhalte, sowie die Metadatenerzeugung zuständig sind. Die Datensätze selber verbleiben bei diesen

lokalen Datenzentren, die TIB speichert die Metadaten und macht alle registrierten Inhalte über eine Datenbank suchbar. (Brase, 2004; Lautenschlager et al., 2005)

Für die Registrierung von Datensätzen wurde an der TIB ein Webservice eingerichtet (SOAP/WSDL). Komplementär wurden bei den Publikationsagenten entsprechende Clienten eingerichtet, die sowohl eine automatisierte als auch manuelle Registrierung ermöglichen. In allen Datenzentren sind die SOAP Clienten vollständig in die Archivierungsumgebung integriert, so dass zusätzlicher Arbeitsaufwand für die Registrierung entfällt. Mithilfe dieser Infrastruktur sind bisher problemlos mehrere Hunderttausend DOI Namen registriert worden. Das System baut auf Seiten der TIB auf dem XML-basierten Publishing-Framework COCOON von Apache auf. Dazu wurde COCOON um eine integrierte Webservice-Schnittstelle erweitert, wodurch die Anbindung von weiterer Software überflüssig wird. Die modulare Struktur des Systems erlaubt es das System auf einfache Weise auf alle weiteren Inhalte, die mit DOI Namen registriert werden, anzupassen.

GeoForschungsZentrum Potsdam als Publikationsagent

Seit 2005 ist das GeoForschungsZentrum Potsdam Publikationsagent für Primärdaten aus dem Bereich Erd- und Umweltwissenschaften.

Das GFZ wurde am 1. Januar 1992 als Stiftung des öffentlichen Rechts gegründet. Es umfasst, weltweit erstmals, alle Disziplinen der Wissenschaften der festen Erde von der Geodäsie über die Geophysik, Geologie und Mineralogie bis zur Geochemie in einem multidisziplinären Forschungsverbund.

Anders als die anderen Agenturen bedient das GFZ Potsdam nicht ein einzelnes Repository, sondern mehrerer von einander unabhängige Repositories mit unterschiedlichen Datenbeständen und jeweils eigenen Arbeitsabläufen.

Um die Vergabe von DOIs zentral zu verwalten wurde das System DOIDB entwickelt, das die Registrierung von DOIs aus den Datenrepositories des GFZ Potsdam bündelt.

- Das System DOIDB besteht aus folgenden Komponenten (Abb. 3):
- Datenbank zur Protokollierung aller Vorgänge im Zusammenhang mit Datenpublikationen
- Administrationsoberfläche der DOIDB
- DOIDB Webservice zur Annahme von Vorgängen im Zusammenhang mit Datenpublikationen aus Datenrepositories des GFZ Potsdam
- Registrierung von Datenpublikationen über DOIDB
- Client-Servlet für die Kommunikation mit der TIB Hannover

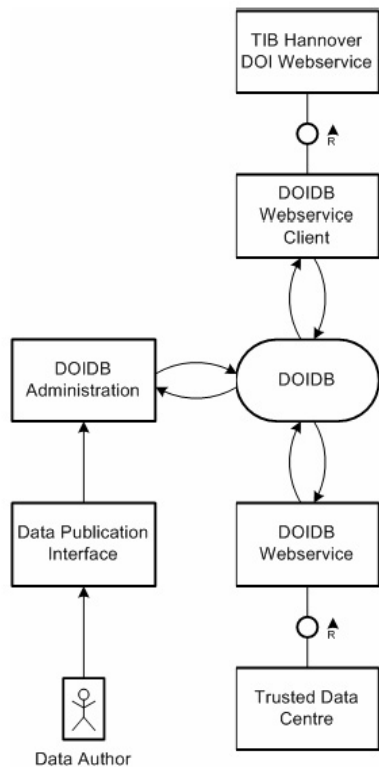


Abbildung 3: Schematischer Aufbau des Systems DOIDB

In Analogie zum Webservice an der TIB (Hinzmann, 2005) bietet der DOIDB Webservice fünf Methodenaufrufe für die Registrierung und Verwaltung von DOI Namen, sowie zwei Testmethoden, mit denen die Funktionsfähigkeit der Kommunikation zwischen dem DOIDB Webservice und dem aufrufenden Klienten einerseits und zwischen der DOIDB und dem Webservice der TIB Hannover andererseits geprüft werden kann. Die Kommunikation zwischen DOIDB Webservice und Klienten geschieht über HTTPS und SOAP.

Auch wenn es hier eine grundlegende Übereinstimmung gibt, zwischen dem DOIDB Webservice und dem STD-DOI Webservice, so erfordern die Gegebenheiten am GFZ Potsdam auch Unterschiede zum Webservice der TIB Hannover.

Im Webservice der DOIDB wird unterschieden zwischen Datenpublikation durch vertrauenswürdige Datenproduzenten (Trusted Data Centre) und anderen Datenproduzenten (Data Author). Die von einem Trusted Data Centre angemeldeten Datenveröffentlichungen werden in der DOIDB registriert und dann automatisch an den Webservice der TIB weitergereicht. Datenpublikationen von andern Datenproduzenten, die normalerweise nur in kleinen Stückzahlen angemeldet werden, müssen vom Administrator der DOIDB manuell frei geschaltet werden.

Status

Die DOI-Registrierung von Primärdaten ermöglicht eine elegante Verlinkung zwischen einem Wissenschaftlichen Artikel und den im Artikel analysierten Primärdaten. Artikel und Datensatz sind durch die DOI in gleicher Weise eigenständig zitierbar.

So wird beispielsweise der Datensatz:

G.Yancheva, . R Nowaczyk et al (2007)

Rock magnetism and X-ray fluorescence spectrometry analyses on sediment cores of the Lake Huguang Maar, Southeast China, PANGAEA

doi:10.1594/PANGAEA.587840

in folgendem Artikel zitiert.

G. Yancheva, N. R. Nowaczyk et al (2007)

Influence of the intertropical convergence zone on the East Asian monsoon

Nature 445, 74-77

doi:10.1038/nature05431

Mittlerweile hat die TIB ihr Angebot auch auf andere Inhaltsformen ausgeweitet. Als Beispiele seien hier genannt:

- doi:10.1594/EURORAD/CASE.1113 in Kooperation mit dem European Congress for Radiology (ECR) wurden über 6.500 medizinische Fallstudien registriert.
- doi:10.2312/EGPGV/EGPGV06/027-034 in Kooperation mit der European Association for Computer Graphics (Eurographics) wurden über 300 Artikel (Graue Literatur) registriert.
- doi:10.1594/ecrystals.chem.soton.ac.uk/145 Gemeinsam mit dem Projekt eBank des UK Office for Library Networking wurden erstmals DOI Namen für Kristallstrukturen vergeben.
- doi:10.2314/CERN-THESIS-2007-001 in Kooperation mit dem CERN werden DOI Namen für Berichte und Dissertationen vergeben
- doi:10.2314/511535090 Seit Sommer 2007 vergibt die TIB auch DOI Namen für BMBF Forschungsberichte.

Literatur

Brase, J., 2004. Using Digital Library Techniques - Registration of Scientific Primary Data. Lecture Notes in Computer Science, 3232: 488-494.

Hinzmann, J., 2005. Erweiterung des XML Web Development Frameworks Cocoon zu einem Webservice am Beispiel der Registrierung von Primärdaten. B.Sc. Thesis, Universität Hannover, Hannover, Germany, 82 pp.

Klump, J. 2006. Registrierung von Datenpublikationen im GFZ Potsdam über DOI DB Technical Report, GFZ Potsdam

Lautenschlager, M., Diepenbroek, M., Grobe, H., Klump, J. and Paliouras, E., 2005. World Data Center Cluster "Earth System Research" - An Approach for a Common Data Infrastructure in Geosciences. EOS, Transactions, American Geophysical Union, 86(52, Fall Meeting Suppl.): Abstract IN43C-02.

Uhlir, Paul F. ,2003 The Role of Scientific and Technical Data and Information in the Public Domain, National Academic Press, Washington DC

Die Systemarchitektur von PROBADO: Der allgemeine Zugriff auf Repositorien mit nicht-textuellen Inhalten

Harald Krottmaier

Abstract:

Zentrale Fachbibliotheken und Fachinformationszentren haben bislang die Informationsversorgung mit Textdokumenten, vor allem Zeitschriftenartikel, für ihre Kunden aus Forschung, Lehre und Wirtschaft sicherstellen können. Exemplarisch sei hier auf die klassischen Bibliotheksdienstleistungen wie Dokumentlieferdienste, virtuelle Fachbibliotheken, aber auch auf digitale Bibliotheken verwiesen. Die Erfüllung dieses Versorgungsauftrages befindet sich allerdings durch neue mediale Formate und Multimedia-Objekte (z.B. Musik, Architekturmodelle und E-Learning-Material) im Umbruch, die in der Praxis immer öfter von den Benutzern benötigt und angefordert werden und daher zur umfassenden Informationsversorgung über ein Bibliotheksportal dazugehören, das sowohl die Anforderungen einer text-basierten wie auch audio-visuellen Suche integriert.

Die Herausforderung des von der DFG geförderten Projektes PROBADO liegt im Umgang mit diesen neuen Dokumenten und einer neuen Generation von Metadaten. Jene nicht-textuellen digitalen Objekte benötigen als inhaltliche Basis einer Digitalen Bibliothek selbstverständlich andere Methoden und Werkzeuge als jene, die derzeit in der klassischen bibliothekarischen Prozesskette eingesetzt werden. Die zu entwickelnde Systemarchitektur muss einfach in die bestehende Infrastruktur der Bibliotheken eingebunden und nachhaltig genutzt werden können. Die neuen Werkzeuge zum Retrieval, zur Darstellung usw. müssen so einfach zu bedienen sein, wie die derzeit eingesetzten Werkzeuge für textbasierte Dokumente.

Abstract:

In the past most libraries supplied their customers with textual documents. However, more and more documents are not just available in textual format. Music as well as 3D-models and E-Learning-material are often supplements to textual documents and users want to access them. These types of media require new types of search interfaces and these interfaces must be provided to users.

In PROBADO we want to handle these types of documents like textual documents are handled today in the well-known and established workflows. New tools and methods must be developed for librarians as well as for repository suppliers and customers of the libraries, i.e. the readers. These tools should be as easy to use as current tools for textual documents.

Einführung

Eines der wichtigen Ziele in PROBADO ist es, die bereits in der Forschung vorliegenden und stabilen Konzepte zur inhaltsbasierten Erschließung von nicht-textuellen Dokumenten innerhalb eines Dienstes für Bibliotheken und Repository-Betreiber verfügbar zu machen.

Benutzer sind es gewohnt, von einer zentralen Stelle bzw. Suchmaske aus Systeme abzufragen und ihr Informationsbedürfnis zu stillen. Als Beispiel sei hier auf herkömmliche Suchmaschinen im Internet (Google, Yahoo, Vivisimo) verwiesen. Bei derartigen Suchmaschinen werden unterschiedlichste Datenquellen indiziert und die Dokumentbestände über eine Suchmaske zugänglich gemacht. Die Visualisierung der Suchergebnisse erfolgt uniform, sodass dem Benutzer die Darstellung unterschiedlicher Datenquellen geboten wird, die eine Differenzierung nach Herkunft des Resultats nicht ermöglicht. Die Reihenfolge eines Suchergebnisses ist demnach entscheidend für den Zugriff der Benutzer auf die jeweilig gefundene Quelle. Ganze Beratungsunternehmen beschäftigen sich mit sog. *Search Engine Optimization*, damit Resultate möglichst weit oben bei einer Ergebnisliste erscheinen. Die Suchmaschinen indizieren hauptsächlich fast ausschließlich text-basierte Dokumente.

Ein weiterer Nachteil, der sich bei den klassischen Suchmaschinen im Internet ergibt, ist die fehlende Qualitätskontrolle der Inhalte. Klassische Suchmaschinen indizieren meist ohne Einschränkungen große Bereiche des Internets. Differenzierungen --- seien sie nun themen- oder bereichsspezifisch --- sind oft nicht gegeben und damit für den Benutzer nicht umzusetzen. Für Bibliothekare sind Internetsuchmaschinen bekannter weise (nicht nur...) aufgrund der Metadatenproblematik inakzeptable Werkzeuge.

Suchmaschinen für nicht-textuelle Dokumente sind im Internet noch nicht verbreitet. Die Erschließung der Dokumente erfolgt oft nicht automatisch. Meist verlässt sich das System auf herkömmliche Benutzer. Die Dokumente werden über unterschiedliche Mechanismen mit Metadaten versehen. Beispielsweise kann bei Youtube, einer populären Suchmaschine für Videos, über Folksonomy Mechanismen (sog. „Kollaborative Kategorisierung“) eine „Beschlagwortung“ der Dokumente durchgeführt werden. Freie (also im Gültigkeitsbereich nicht eingeschränkte), beschreibende textuelle Metadaten werden hier den Dokumenten zugeordnet. Für Bibliothekare ergeben sich auch hier wieder viele Fragen: Wer garantiert die Qualität der Metadaten? Was passiert mit ähnlichen Begriffen? u.v.m.

Dieser kurze Einstieg in das Thema von Suchmaschinen und der sehr kurze Abriss über die Probleme, die aktuell noch herrschen, zeigt, dass die Technologie und Architektur, wie sie derzeit bei Internetapplikationen verwendet wird, nicht ohne weiteres in einen Bibliotheksbetrieb übernommen werden kann.

Neben der Benutzerschnittstelle ergeben sich für Bibliothekare auch im Bereich der Akquise Fragen. Viele der nicht-textuellen Dokumente, z.T. Massenmaterialien, sind auf Institutsservern bzw. dezentralen Repositorien gespeichert und sollen dauerhaft zentral zugreifbar, nutzbar und zitierbar gemacht werden. Dabei sollen diese Daten

physisch *nicht* in einem zentralen Repository vorhanden sein. Damit die Dokumente schnell gefunden werden können, wird man Teilaspekte der Architektur, wie sie bei Internetsuchmaschinen verwendet werden, verwenden. D.h. ein zentraler Index über den Dokumentbestand muss erstellt werden. Das Dokument selbst soll aber nur über das Ursprungsrepository zur Verfügung gestellt werden.

Anforderungen

In der Einführung wurde gezeigt, dass herkömmliche Internetsuchsysteme für den Bibliotheksbereich nicht ohne Adaptionen übernommen werden können. Anpassungen müssen durchgeführt werden, um den Ansprüchen der Bibliothekare, der Repository-Betreiber und der Benutzer gerecht zu werden. Jeder dieser drei Gruppen hat unterschiedliche Anforderungen an das System die berücksichtigt werden sollen.

Bibliothekare: Diese Anwendergruppe will dem Benutzer einen einfach zu bedienenden Suchservice anbieten. Ausgewählte und hochqualitative Datenquellen sollen durchsucht werden können. Das Ergebnis soll klar und übersichtlich mit Kontextinformation dargestellt werden. Für die Akquise von neuen Dokumenten ergeben sich ebenso wichtige Anforderungen: Die Akquise soll ohne großen personellen Aufwand möglich sein und die akquirierten Dokumente müssen stabil sein, d.h. für lange Zeit verfügbar gemacht werden. Detaillierte und austauschbare Metadaten müssen zu den Dokumenten (semi-)automatisch erstellt werden, sollen aber auch von einer eingeschränkten Benutzergruppe gewartet werden können.

Repository-Betreiber: Der Inhalt, der in den Repositorien gespeichert ist, soll möglichst vielen Benutzern einfach zugänglich gemacht werden. Die eventuell vorhandenen außergewöhnlichen Suchmasken sollen dem Benutzer einfach die Erstellung der Suchanfrage ermöglichen. Bei z.B. Bilderdatenbanken kann man sich sog. „query-by-example“-Schnittstellen vorstellen, bei dem ähnliche Bilder über eine spezielle Benutzerschnittstelle gezeichnet werden können. Das Repository kann dann nach ähnlichen Bildern suchen. Für ein Repository mit Musik als Rohdaten muss eine derartige Benutzerschnittstelle vollkommen anders umgesetzt werden. Hier sollen beispielsweise Noten einfach eingegeben und dargestellt werden können (z.B. über eine Klavierwalze) oder es können weitere Schnittstellen (z.B. über „vorsingen/-summen“) umgesetzt werden. Für die Akquise von Dokumenten soll der bestehende Arbeitsablauf eingehalten werden können, die vorhandenen Metadaten zu Dokumenten sollen den Benutzer weitergereicht werden können. Eine Anreicherung durch weitere Metadaten soll einfach möglich sein. Die Anbindung an ein Gesamtsystem muss ohne zu großen Aufwand erfolgen können. Die Dokumente sollen immer über das Repository selbst abgeholt werden. Damit können bestehende Verrechnungssysteme weiterverwendet werden und die Dokumente verbleiben zur Auslieferung an den Benutzer im Repository.

Benutzer: Der Benutzer will alle Vorteile, die eine Bibliothek bietet, nutzen. Die Vorteile erstrecken sich von den hochqualitativen Dokumentsammlungen hin zur bereichseingeschränkten Suche. Bei virtuellen Bibliotheken bzw. Bibliotheksverbünden ist es für den Benutzer belanglos, wo die Dokumente gespeichert werden. Der Zugriff soll komfortabel und schnell sein.

Man sieht an diesem Überblick, dass sehr unterschiedliche Anforderungen an ein Gesamtsystem gestellt werden. Im folgenden Abschnitt wird der aktuelle Entwurf von PROBADO dargestellt und diskutiert.

Systemarchitektur

Abb. 1 zeigt den Aufbau der Systemarchitektur von PROBADO. Die Aufteilung erfolgt in drei Ebenen: Dem Frontend-Layer (1), dem Core-Layer (2), und einem Repository-Wrapper (3).

Der Entwurf dieser Architektur wurde im Rahmen von Workshops mit den Projektpartnern diskutiert und erarbeitet. Die wesentlichsten Ziele, die bei der Umsetzung dieser Architektur zu erreichen sind, sind Erweiterbarkeit und Skalierbarkeit. Für das Projekt ergeben sich während der Entwicklungsphase drei Dokumenttypschwerpunkte: 3D- und Musik-Dokumente, sowie E-Learning Material. Diese Dokumenttypen haben wenig gemeinsam und sind in ihrer Abgeschlossenheit sehr speziell.

Kommunikation zwischen den einzelnen Ebenen ausgewählt. Ein weiterer Parameter (Plattformunabhängigkeit) begründet die Wahl der Programmiersprache auf der Serverseite. Hier wird Java verwendet. Bereits seit Anfang 2002 sind für Java Webservice-Pakete verfügbar, die die Entwicklung von Webservices vereinfachen.

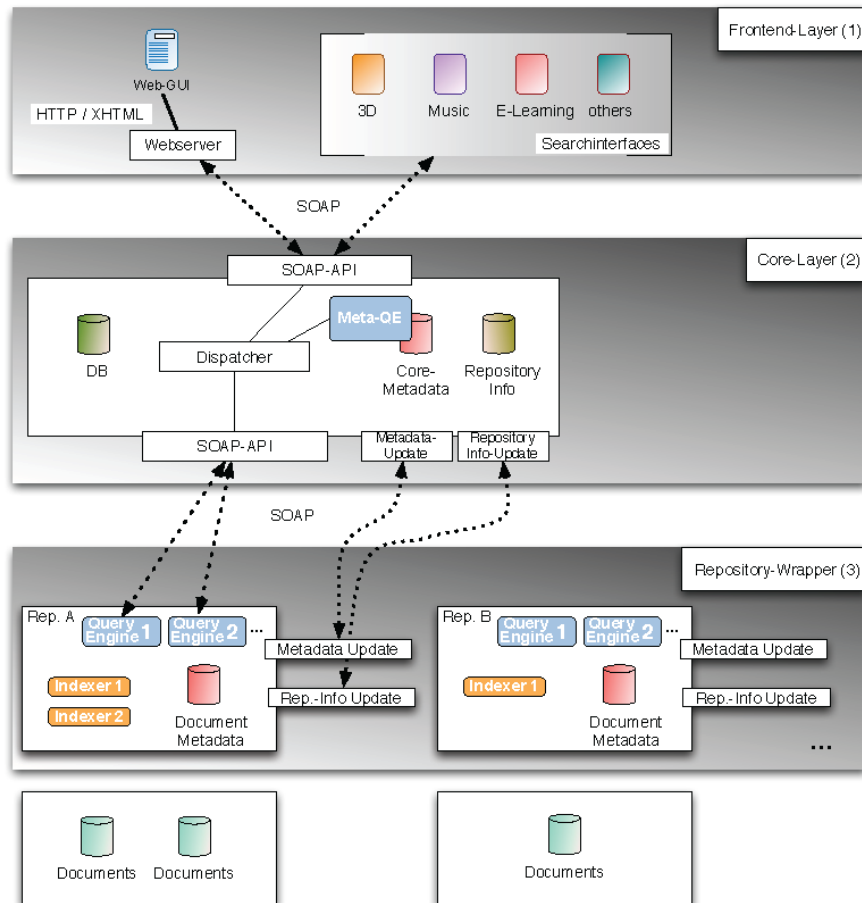


Abb. 1: Systemarchitektur von PROBADO

Die Benutzerschnittstelle (**Frontend-Layer (1)**) ist in zwei Bereiche aufgeteilt: Ein Bereich, der über einen einfachen Webbrowser umgesetzt wird und ein weiterer, der mit dokumententypspezifischen Programmen (plug-ins für den Browser und/oder separate Client-Programme) arbeitet. Beide Bereiche sollen die Erstellung von Suchanfragen ermöglichen und die Suchergebnisse mit Kontextinformation darstellen können.

Wenn lediglich ein Webbrowser für die Formulierung von Suchanfragen zur Verfügung steht, können nur „einfache“ Suchanfragen formuliert werden. Einschränkungen in Themenbereiche sind möglich und werden über drop-down Listen dargestellt. Da ein Webbrowser nur einfache Textfelder für die Formulierung der Anfrage zur Verfügung stellen kann, ist die Suchmöglichkeit nur nach Texten möglich. Mit Hilfe von JavaScript-Programmen und DHTML-Elementen kann man auch im Webbrowser mit dynamischen Elementen arbeiten und komfortable

Benutzerschnittstellen erstellen. Die Verwendung von AJAX-Technologien erlauben sogar noch intuitivere und sich anpassende Benutzerschnittstellen.

Im zweiten Bereich der obersten Systemschicht ergeben sich weitere Möglichkeiten: Programme für spezielle Dokumenttypen können die bereits genannten „query-by-example“ Anfragen formulieren. Um auch hier möglichst vielen Benutzern die Verwendung des Systems zu ermöglichen, wird auf Plattformunabhängigkeit Rücksicht genommen.

Die im Frontend-Layer formulierten Anfragen werden über SOAP-Nachrichten an den **Core-Layer (2)** gesendet. Die SOAP-Nachricht, die von den Benutzerschnittstellen abgeschickt werden, werden auf dieser Ebene interpretiert, teilweise an andere Systemelemente weitergeschickt und manipuliert. Für die Benutzerschnittstelle ist nur ein definierter Satz an Nachrichten gültig. Notwendige Erweiterungen werden von den Systementwicklern in Absprache mit den Entwicklern der Benutzerschnittstellen umgesetzt.

Neben der Bearbeitung der eingehenden SOAP-Nachrichten, werden in dieser Ebene die Core-Metadaten (ein mit allen Partnern abgestimmter Satz an dokumenttypübergreifenden Metadaten) und diverse administrative Daten (Benutzerinformationen, Repository-Information etc.) gespeichert, sowie weitere Dienste (z.B. Empfehlungs- und Annotationssysteme) zentral an das System angebunden. Damit sollen sämtliche Dienste sofort für alle Dokumenttypen zugänglich sein.

Diese Ebene ist das Herzstück von PROBADO. Sie ist u.a. auch für die Weiterleitung von SOAP-Nachrichten an die untere Schicht verantwortlich. Das in Abb. 1 eingezeichnete Modul „Dispatcher“ ist für die Weiterleitung an das jeweils angebundene Repository zuständig. Um eine uniforme Darstellung von Suchergebnissen zu ermöglichen, werden die Resultate in dieser Schicht aufbereitet, mit der notwendigen Kontextinformation versehen und Hyperlinks zu den Dokumenten selbst aber auch den Voransichten erstellt. Die Formatierung der Resultate erfolgt aber selbstverständlich in der Ebene über den Core-Layer (z.B. über XSLT für die Darstellung im Webbrowser).

Die letzte Schicht der Systemarchitektur --- **der Repository-Wrapper (3)** --- bindet existierende Repositorien an das System an. Auch hier werden, wie bei der Kommunikation zwischen Ebene 1 und 2, definierte SOAP-Nachrichten ausgetauscht.

Die existierenden Repositorien werden um weitere Metadaten angereichert. Diese werden über dokumenttypspezifische Indexer erstellt und können über Benutzerschnittstellen angepasst werden. Es hat sich bei Diskussionen gezeigt, dass sich hier kein allgemeiner, für alle Dokumenttypen gültiger Mechanismus umsetzen lässt. Die Verantwortlichen für die jeweiligen Dokumenttypen können hier selbst entscheiden, wie (d.h. beispielsweise in welcher Programmiersprache) die Module zur Metadatengenerierung implementiert werden. Wichtig für den Core-Layer ist lediglich, dass die SOAP-Nachrichten die ausgeschildet werden, auch von dem Repository-Wrapper verstanden werden.

Zwei weitere wichtige Schnittstellen sind bei der Anbindung von Repositorien noch zu berücksichtigen:

Metadatenupdate: Dem Core-Layer muss es möglich sein, nach neuen Dokumenten im Repository zu fragen. Hier ist --- ähnlich wie bei anderen kooperativen, lose-gekoppelten Systemen (wie beispielsweise Systeme, die das *Open Archive Initiative Protocol for Metadata-Harvesting* unterstützen) --- ein bestimmtes Protokoll zu implementieren. Auch hier wird auf die bewährte Webservice-Technologie gesetzt. Sind im Repository tatsächlich neue Dokument vorhanden, müssen die Core-Metadaten der Dokumente in die Datenbank des Core-Layers eingepflegt werden.

Repository-Information: Der Core-Layer muss automatisch über die aktiven Repositories informiert sein. Nur so kann das Dispatcher-Modul entscheiden, welche Repositories für Anfragen von den Benutzern aktuell zur Verfügung stehen. Für die erstmalige Anbindung an das System muss der Administrator von PROBADO die Verbindung herstellen. Im Sinne eines kooperativen Dienstes, sollen die einzelnen Teile, die miteinander kommunizieren, von einem Administrator geprüft werden. Zugangscodes sowie diverse administrative Informationen (beispielsweise wie oft sind im Repository neue Dokumente zu erwarten) werden hier festgelegt. Im Wesentlichen werden die einzelnen Schichten über definierte SOAP-Nachrichten angesprochen. Die Nachrichten, die beispielsweise für die einzelnen Resultate verschickt werden, sind definiert und müssen von den Komponenten implementiert werden.

Durch diese Architektur lassen sich weitgehend alle existierenden Repositorien anbinden und für den Dienst PROBADO nutzen.

Zusammenfassung

In dieser Arbeit wurde anfangs die Situation mit herkömmlichen Suchmaschinen und den aktuellen Problemen dargestellt. Danach wurden die Anforderungen der unterschiedlichen Benutzergruppen notiert. Hierbei wurde klar, dass die Anforderungen sehr unterschiedlich sind.

Mit der vorgeschlagenen Systemarchitektur wird versucht, die unterschiedlichen Anforderungen abzudecken. Die Kommunikation der einzelnen Schichten erfolgt über bereits etablierte Web-Technologien (Webservices, SOAP) und garantiert Plattformunabhängigkeit sowie Flexibilität in der Umsetzung und Wahl der Entwicklungswerkzeuge.

Danksagung

Die Ergebnisse dieser Ausarbeitung wurden in vielen Workshops mit den Projektpartnern von PROBADO erarbeitet. Das Projekt ist auf 5 Jahre Projektlaufzeit ausgelegt und wurde im Februar 2006 gestartet. Die Koordination des Projekts erfolgt von der Gruppe um Prof. Fellner in Graz. Für den Bereich 3D-Dokumente zeichnen Universität Bonn, Institut für Informatik II (Prof. Klein), Technische

Informationsbibliothek Hannover (Dir. Rosemann) und TU Graz (Prof. Fellner) verantwortlich, der Bereich Musik wird von der Universität Bonn, Institut für Informatik III (Prof. Clausen) sowie der Bayerischen Staatsbibliothek, München (Gen.-Dir. Griebel) umgesetzt, der Bereich E-Learning wird von OFFIS – Institut für Informatik, Oldenburg (Prof. Appelrath) verwirklicht.

PROBADO wird von der Deutschen Forschungsgemeinschaft (554975 (1) Oldenburg, BIB48 OLoF 01-02) gefördert.

Literatur

[Google] <http://www.google.com>

[Yahoo] <http://www.yahoo.com>

[Vivisimo] <http://www.vivisimo.com>

[Youtube] <http://www.youtube.com>

[Java] <http://java.sun.com>

[AJAX] Jesse James Garrett: „Ajax: A New Approach to Web Applications“. 2005. <http://adaptivepath.com/publications/essays/archives/000385.php>

[OAI] <http://www.openarchives.org>

[SOAP] <http://www.w3.org/2000/xp/Group/>

Open Access Publishing in der Hochenergiephysik: Das SCOAP³ Projekt

Annette Holtkamp

Abstract

Die Hochenergiephysik (HEP) spielt durch ihre Preprint-Kultur und das arXiv-Repositorium seit langem eine Vorreiterrolle in der Open Access(OA)-Bewegung. Dieser Artikel beschreibt ein Projekt, die Mehrheit der HEP-Zeitschriftenartikel unter Bewahrung des bewährten Peer Review-Systems frei zugänglich zu machen. Charakteristika der existierenden Publikationslandschaft und das vorgeschlagene zukünftige Modell werden diskutiert.

High energy physics (HEP) has for a long time pioneered Open Access principles through its preprint culture and the arXiv repository. This paper describes a project to convert the majority of HEP journal literature to Open Access (OA) while maintaining the well-established peer review system. Characteristics of the existing publishing landscape and the proposed future model are discussed.

1. Einleitung

Ziel der Open Access-Bewegung ist es, jedermann jederzeit freien Zugang zu den Ergebnissen wissenschaftlicher Forschung zu gewähren (MPG, 2003). Die OA-Diskussion wird im Wesentlichen durch zwei Faktoren vorangetrieben:

- Die sog. Zeitschriftenkrise, in der sich auf Grund stetig steigender Subskriptionskosten Bibliotheken gezwungen sehen, immer mehr Abonnements zu kündigen. Den Wissenschaftlern wird so der Zugang zu wichtigen Forschungsergebnissen immer stärker beschnitten.
- Das wachsende Bewusstsein, dass die Ergebnisse öffentlich geförderter Forschung allgemein zugänglich sein sollten.

Die Gemeinschaft der Hochenergiephysiker verwirklicht seit vielen Jahren OA-Prinzipien. Dies drückt sich vor allem in der lebhaften Preprint-Kultur aus, die schon lange vor der Erfindung des WWW begann und dann Anfang der 90er Jahre ihre elektronische Entsprechung in Form des arXiv-Repositoriums arXiv.org fand. Preprints stellen heute das wichtigste Medium wissenschaftlichen Informationsaustausches in der Hochenergiephysik dar und sind zu ca. 90% in Repositorien frei zugänglich.

Es herrscht in der Gemeinschaft jedoch weitgehend Konsens, dass Repositorien Zeitschriften nicht ersetzen können. Besonders im Kontext der Begutachtung wissenschaftlicher Leistungen einzelner Wissenschaftler oder Forschungsinstitute werden traditionelle Zeitschriften auch im Internetzeitalter u.a. wegen ihrer Qualitätskontrolle wissenschaftlicher Veröffentlichungen (Peer-Review) nach wie vor als unverzichtbar angesehen.

Die hier vorgestellte Initiative SCOAP³ (Sponsoring Consortium for Open Access Publishing in Particle Physics) zielt darauf ab, in Zusammenarbeit mit Verlagen freien Zugang zur HEP-Zeitschriftenliteratur unter Beibehaltung des traditionellen Peer Review-Systems zu etablieren. Im Folgenden werden die Besonderheiten der HEP-Publikationslandschaft und sich daraus ergebende Möglichkeiten für einen Übergang zu Open Access, das SCOAP³-Modell und die bisher erfolgten Schritte beschrieben.

2. Die HEP-Publikationslandschaft

Das hervorstechendste Merkmal der HEP-Publikationslandschaft ist ihre Preprint-Kultur. Um dem Zeitschriftenpublikationsprozess inhärente Zeitverzögerungen zu umgehen, verlassen sich Teilchenphysiker seit Jahrzehnten auf den Austausch von Preprints als primäres Informationsmedium. Der zunächst postalische Versand von Papierkopien wurde Anfang der 90er Jahre durch den Aufbau einer elektronischen Infrastruktur, des heute als arXiv.org bekannten Repositoriums, ersetzt. Heutzutage sind ca. 90% aller Preprints in diesem oder anderen Repositorien über das Internet frei zugänglich. Dieser unbeschränkte und zeitnahe elektronische Zugang zu allen wesentlichen Forschungsergebnissen revolutionierte die Geschwindigkeit und Effizienz wissenschaftlicher Kommunikation und sorgte dafür, dass die etablierten Zeitschriftenverlage ihre ursprüngliche Funktion der Kommunikation wissenschaftlicher Resultate für den Bereich der Hochenergiephysik weitgehend verloren haben.

Dennoch herrscht unter Teilchenphysikern weitgehend Einigkeit, dass wissenschaftliche Zeitschriften weiterhin wichtige Funktionen erfüllen.

- Als vorrangig wird die Aufgabe der Qualitätssicherung in Gestalt des Peer Reviews angesehen.
- Publikationen in referierten Zeitschriften dienen als Bewertungsbasis bei Stellenbesetzungen und Karriereentscheidungen.
- Zeitschriftenveröffentlichungen gelten Förderorganisationen als Maß für die Qualität und Produktivität von Forschungsgruppen und –instituten.
- Langzeitarchivierung wird von Repositorien in ihrer gegenwärtigen Form nicht hinreichend garantiert. Zudem stellen Repositorien meist nicht die endgültige publizierte Version eines Artikels zur Verfügung.

Trotz laufender Diskussionen über neue Qualitätsmetriken und alternative Peer Review-Systeme gelten Zeitschriften nach wie vor als unverzichtbar, und die weitestmögliche Verbreitung qualitätskontrollierter Zeitschriftenartikel ist von vorrangiger Bedeutung für die wissenschaftliche Gemeinschaft. Aufgrund der Spirale ständig steigender Subskriptionspreise bei sinkender Zahl der Abonnements wird der Zugang zu referierter Forschungsliteratur jedoch zunehmend erschwert und einige wichtige Zeitschriftentitel sehen sich in ihrer Überlebensfähigkeit bedroht. Unter der Voraussetzung, dass geeignete Finanzierungsmodelle gefunden werden können, stellen dem OA-Prinzip gehorchende Zeitschriften einen Ausweg aus der Krise dar.

Zusätzlich zu der schon existierenden starken Verankerung des OA-Gedankens gibt es weitere Faktoren, die die Hochenergiephysik für einen vollständigen Übergang zu Open Access prädestiniert erscheinen lassen. Dies ist zum einen die weitgehende Überschneidung von Leser- und Autorenbasis, die eine Ablösung des bisherigen Subskriptionsmodells durch neue Finanzierungsformen als leichter realisierbar erscheinen lässt als in manch anderen Forschungszweigen.

Ein weiterer wichtiger Faktor ist die Überschaubarkeit der Zahl der HEP-Publikationen. Im Jahr 2005 wurden ca. 8500 HEP-Artikel in ca. 150 Zeitschriften publiziert (S. Bianco et al., 2007). In Zeitschriften veröffentlichte Vorträge sind in dieser Zahl nicht eingeschlossen, da diese in der Regel nicht dem gleichen Peer Review-Verfahren unterliegen wie reguläre Artikel und deshalb zunächst außerhalb des von SCOAP³ gesteckten Rahmens fallen. Diese im Vergleich zu manch anderen Disziplinen relativ niedrige Anzahl von Publikationen lässt das benötigte Finanzvolumen von SCOAP³ überschaubar erscheinen.

Auffällig ist weiterhin eine starke Konzentration auf wenige Zeitschriften (siehe Abb. 1). 2005 erschienen ca. 62% aller Originalarbeiten in nur sieben Zeitschriften, herausgegeben von vier Verlagen (S. Bianco et al., 2007). Die Kooperation mit nur einigen wenigen Verlagen würde also den Übergang zu Open Access für den Großteil der HEP-Literatur ermöglichen.

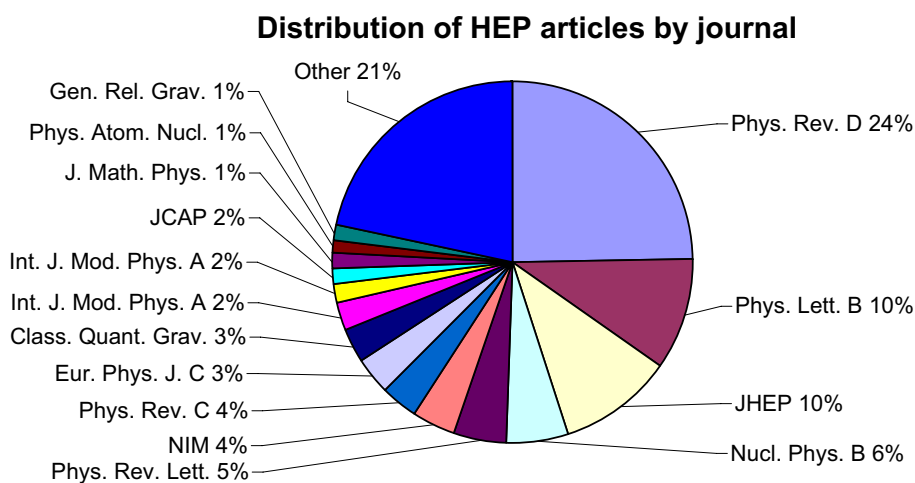


Abbildung 1: Anteil einzelner Zeitschriften an den HEP Veröffentlichungen des Jahres 2005

Fünf der in Abb. 1 genannten Zeitschriften publizieren überwiegend HEP-Artikel:

- Physical Review D (American Physical Society),
- Physics Letters B (Elsevier),
- Journal of High Energy Physics (SISSA/IOP),
- Nuclear Physics B (Elsevier),
- European Physical Journal C (Springer).

SCOAP³ zielt darauf ab, diese Zeitschriften komplett nach OA zu konvertieren.

Eine weitere wichtige Zeitschrift, *Physical Review Letters* (American Physical Society), ist eine Breitbandzeitschrift mit nur ca. 10% HEP-Anteil. SCOAP³ strebt an, diesen Anteil auf der Basis einzelner Artikel freizukaufen. Dies gilt auch für eine andere grundlegende Zeitschrift: *Nuclear Instruments and Methods in Physics Research A* (Elsevier), deren Artikel zu ca. 25% in den Bereich der Hochenergiephysik fallen.

Diese Zeitschriften, die gegenwärtig den Löwenanteil der HEP-Literatur veröffentlichen, stehen naturgemäß im Zentrum des Interesses von SCOAP³. Um jedoch einen dynamischen Markt mit Wettbewerb und Wahlmöglichkeiten zu fördern, ist SCOAP³ im Rahmen seiner Finanzmöglichkeiten offen für alle existierenden oder neu entstehenden hochwertigen Zeitschriften, die HEP-Artikel veröffentlichen.

3. Das SCOAP³-Modell

SCOAP³ verfolgt zwei wesentliche Ziele:

- freien ungehinderten Zugang zur gesamten HEP-Forschungsliteratur in ihrer endgültigen referierten Version zu gewährleisten,
- die Kostensteigerung des HEP-Zeitschriftenmarktes z.B. durch Wettbewerbsförderung zu dämpfen.

SCOAP³ zielt darauf ab, in Zusammenarbeit mit wissenschaftlichen Verlagen ein lebensfähiges Geschäftsmodell für OA-Publikationen zu etablieren, das so transparent wie möglich für HEP-Autoren ist, ihnen weiterhin eine Auswahl an Publikationsmöglichkeiten bietet und gleichzeitig den bisherigen Qualitätsstandard sichert.

Darüber hinaus strebt SCOAP³ eine neue e-science Infrastruktur für die Hochenergiephysik an:

- Repositorium von referierten Zeitschriftenartikeln mit Volltext- und data mining-Anwendungen,
- System zur Messung der wissenschaftlichen Produktion einzelner Länder, das die Basis für die Kostenverteilung innerhalb des Konsortiums bilden wird,
- umfassender, frei zugänglicher Zitationsindex für HEP-Publikationen.

Da bei einer OA-Zeitschrift die leserbasierten Subskriptionsgebühren entfallen, müssen diese durch alternative, autorenbasierte Finanzquellen ersetzt werden. Im Wesentlichen dominieren zwei Modelle die Diskussion: Autorengebühren pro Artikel und Sponsorenmodelle.

Artikelgebühren, die vom Autor oder seinem Institut zu entrichten sind, mögen zwar als fairste Methode der Kostenverteilung erscheinen, erschweren aber den Übergang zu Open Access, da sie das Forschungsgeldbudget einzelner Gruppen zusätzlich belasten und zu Ungunsten kleiner publikationsfreudiger Theoriegruppen gehen, die den Großteil der HEP-Literatur produzieren. SCOAP³ wird deshalb das Sponsorenmodell praktizieren.

In dem hier vorgeschlagenen Modell werden Subskriptionsgebühren, die heute von vielen einzelnen Institutionen und Bibliothekskonsortien entrichtet werden, durch Blockzahlungen von SCOAP³, dem dann einzigen finanziellen Partner der Verlage, ersetzt. Das SCOAP³-Konsortium wird sich im Wesentlichen aus Förderorganisationen, Forschungsinstituten und Bibliotheken zusammensetzen. Finanzierung und Lenkung werden so weit wie möglich dem Vorbild der großen HEP-Forschungskollaborationen folgen. Der finanzielle Beitrag jedes Landes wird proportional zur Zahl seiner wissenschaftlichen Publikationen bestimmt, wobei jeder Partner seinen Anteil größtenteils durch Umlenkung der bestehenden Subskriptionsbudgets finanzieren wird.

Um die Größenordnung der auf die einzelnen Länder entfallenden Beiträge zu ermitteln, wurden alle HEP-Artikel untersucht (J. Krause et al., 2007), die in den Jahren 2005 und 2006 in den in Kap. 2 genannten fünf Kern- und zwei Breitbandzeitschriften publiziert wurden. Für jeden Artikel wurden die Autoren je einem Land zugeordnet und anteilig gezählt. Die Ergebnisse dieser Studie sind in Abb. 2 zusammengefasst.

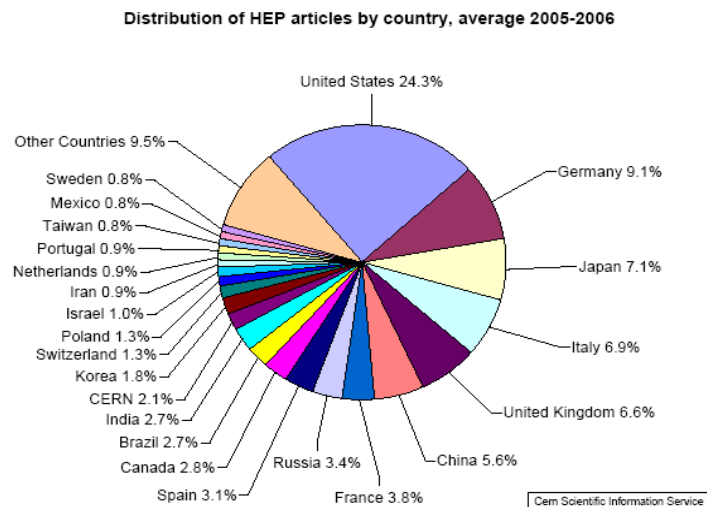


Abbildung 2: Beitrag einzelner Länder zur HEP-Literatur in den fünf Kern- und zwei Breitbandzeitschriften (siehe Kap. 2) in den Jahren 2005/2006. Koautorenschaft wurde pro rata berücksichtigt, indem Artikel-Bruchteile einzelnen Ländern aufgrund der Institutszugehörigkeit der Autoren zugeordnet wurden. Länder mit Beiträgen unter 0.8% wurden in der Kategorie „Other countries“ zusammengefasst. (J. Krause et al., 2007)

Die pro rata-Anteile der teilnehmenden Länder werden in jedem Jahr neu ermittelt. Um für die Publikationen finanzschwacher Länder, von denen momentan kein Beitrag zum Konsortium erwartet werden kann, aufzukommen, wird ein Anteil von höchstens 10% des SCOAP³-Budgets vorgesehen.

Der Finanzbedarf einer OA-Zeitschrift² wird im Wesentlichen durch die Kosten des Peer Review-Verfahrens, der Redaktionsarbeit, der elektronischen Publikation und der Archivierung bestimmt. Die von den meisten Verlagen genannten Zahlen bewegen sich zwischen 1000 und 2000€. Berücksichtigt man, dass Zeitschriften gewissen Standards bezüglich Qualität und HEP-Anteil genügen sollten, ergibt sich auf Grundlage der Zahlen aus Kap. 2 ein notwendiges jährliches Budget zur Umwandlung der HEP-Zeitschriftenliteratur nach Open Access von maximal 10 Mio €, wovon auf Deutschland ca. 1 Mio € entfallen würde.

Als Vertragspartner von SCOAP³ kommen Zeitschriften in Frage, die den folgenden Kriterien bezüglich ihrer Infrastruktur, ihrer Politik sowie technischer und finanzieller Aspekte genügen:

- deutlicher Anteil an HEP-Artikeln (mindestens 10%),
- strenges unabhängiges Peer Review-System,
- gute Reputation,
- unmittelbarer und permanenter freier Zugang über das Internet zum Volltext und den zugehörigen Metadaten,
- Recht von SCOAP³, den Volltext mit den entsprechenden Metadaten in ein anderes Repositorium zu übertragen,
- Recht jedes Lesers, Abbildungen, Tabellen und numerische Daten zu extrahieren und wiederzuverwenden,
- Copyright-Vereinbarung, die dem Autor erlaubt, Kopien seines Artikels im Web zu deponieren oder dessen Inhalt in späteren Publikationen zu verwenden,
- verlagsunabhängige Langzeitarchivierung,
- Lieferung der Metadaten aller OA-Artikel an crossref.org,
- Herausnahme einer vollständig konvertierten Zeitschrift aus Lizenzpaketen und entsprechende Reduzierung des Paketpreises,
- Reduzierung des Subskriptionspreises für nur teilweise konvertierte Zeitschriften,
- Rückzahlung bereits geleisteter Subskriptionszahlungen im Falle langfristiger Verträge,
- Zugang zu Zeitschriftenarchiven im gleichen Umfang wie bei bestehenden Subskriptionen.

4. Erste Schritte und Zeitplan

Im Dezember 2005 wurde in Genf anlässlich eines OA-Workshops am CERN eine Task Force, bestehend aus Vertretern von Forschungsinstituten, Verlagen und Förderorganisationen, eingesetzt mit der Aufgabe, Geschäftsmodelle für Open Access-Zeitschriften in der Teilchenphysik zu untersuchen. Als Ergebnis ihrer Arbeit

schlug diese Task Force in ihrem Bericht (R. Voss et al., 2006) im Juni 2006 ein „Sponsoring Consortium for Open Access Publishing in Particle Physics“ vor, einer zentralen Agentur, die Verleger im wesentlichen für den Aufwand des Peer Review-Verfahrens entschädigt und das „reader pays“-Modell effektiv durch eine autorensseitige Finanzierung ersetzt. Dieser Vorschlag wurde im November 2006 von Repräsentanten führender Europäischer Organisationen aufgegriffen, die eine Arbeitsgruppe einsetzten, um einen detaillierten Vorschlag zur Errichtung eines solchen Konsortiums zu erarbeiten. Die Ergebnisse (S. Bianco et al., 2007) dieser Arbeitsgruppe wurden im Mai 2007 der Öffentlichkeit zugänglich gemacht.

Unter der wachsenden Zahl von Organisationen, die die Gründung von SCOAP³ unterstützen, befinden sich die zwei größten Europäischen HEP-Laboratorien CERN und DESY, die Max-Planck-Gesellschaft, Förderorganisationen wie IN2P3 und CNRS (Frankreich), INFN (Italien), LIP (Portugal) und Bibliothekskonsortien wie COUPERIN (Frankreich) und JISC (UK). Der nächste Schritt auf dem Weg zur Realisierung des Konsortiums wird im Laufe des Sommers 2007 die Annahme des vorgeschlagenen Modells durch die Unterzeichnung eines „Expression of Intent“ sein – zunächst durch die Organisationen, die die SCOAP³-Arbeitsgruppe eingesetzt haben, und anschließend durch möglichst viele weitere interessierte Partner. Auf deutscher Seite wurde ein solcher Brief vom Deutschen Elektronen-Synchrotron (DESY) in Hamburg für ein Konsortium von Helmholtz-Instituten, von der Max-Planck-Gesellschaft (MPG) und für die deutschen Universitäten von der Technischen Informationsbibliothek (TIB) Hannover, die zudem einen Förderantrag an die Deutsche Forschungsgemeinschaft (DFG) stellen wird, unterzeichnet. Diese drei Organisationen werden gemeinsam für die Sicherstellung der Finanzierung des deutschen Anteils sorgen. Der Erfolg von SCOAP³ wird jedoch von einer weltweiten Beteiligung abhängen. Deshalb ist ein Schwerpunkt der gegenwärtigen Arbeit, innerhalb der nächsten Monate außereuropäische Partner zu gewinnen.

Sobald hinreichend viele Organisationen ihr Interesse an SCOAP³ bekundet haben, erfolgen als weitere Schritte die formelle Gründung des Konsortiums durch die Unterzeichnung eines „Memorandum of Understanding“ sowie eine öffentliche Ausschreibung, auf die sich alle Verleger bestehender oder neu zu gründender Zeitschriften bewerben können, die die von SCOAP³ formulierten Kriterien erfüllen. Informelle Diskussionen mit den wichtigsten Verlagen in der Hochenergiephysik haben ein starkes Interesse auf Verlagsseite an der Aufnahme von Verhandlungen auf der Grundlage der dargelegten Prinzipien deutlich werden lassen. Auf der Basis der Ergebnisse der Ausschreibung werden das Gesamtbudget und die pro rata-Anteile der einzelnen Länder festgelegt werden. Die formelle Gründung des Konsortiums wird erfolgen durch Unterzeichnung eines „Memorandum of Understanding“, das die Lenkung des Konsortiums regeln und sich dabei am Beispiel der großen experimentellen HEP-Kollaborationen orientieren wird, und Zustimmung zu der errechneten Kostenverteilung für die Startphase von SCOAP³. Anschließend sollen Verträge mit den einzelnen Verlagen abgeschlossen werden, so dass im Sommer 2008 zum Zeitpunkt des Starts des Large Hadron Collider LHC am CERN,

der zu einer Fülle von Neuveröffentlichungen Anlass geben wird, Open Access in der HEP-Literatur hoffentlich umfassend realisiert sein wird.

5. Ausblick

Das SCOAP³-Projekt bietet die kosteneffektive Chance, die Zeitschriftenliteratur eines ganzen Forschungszweiges komplett nach Open Access zu überführen. Ein globales Netzwerk von Forschungsinstituten, Förderorganisationen und Bibliotheken als starker Verhandlungspartner gegenüber den Verlagen ermöglicht einen Ausweg aus der Zeitschriftenkrise. Aufgrund ihrer Preprint-Kultur, der langen Tradition weltweiter Experimente-Kollaborationen und des überschaubaren Umfangs der Zeitschriftenliteratur erscheint die Hochenergiephysik als besonders geeignetes Feld, einen solchen Meilenstein in der Geschichte wissenschaftlichen Publizierens zu realisieren und praktische Erfahrungen zu sammeln, die eine Ausdehnung des SCOAP³-Modells auf andere Disziplinen, besonders auf benachbarte Bereiche wie die Kern- und Astrophysik ermöglichen.

Literatur

MPG (2003): Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

S. Bianco et al. (2007): Report of the SCOAP3 Working Party, CERN-OPEN-2007-015, <http://cern.ch/oa/Scoap3WPReport.pdf>

J. Krause et al. (2007): J. Krause, C.M. Lindqvist and S. Mele, Quantitative Study of the Geographical Distribution of the Authorship of High-Energy Physics Journals, CERN-OPEN-2007-014, <http://cdsweb.cern.ch/record/1033099>

R. Voss et al. (2006): Report of the Task Force on Open Access Publishing in Particle Physics, http://library.cern.ch/OATaskForce_public.pdf

Metadaten-Management mit MyCoRe

Wiebke Oeltjen

Zusammenfassung

Metadaten sind essentiell für den wissenschaftlichen Austausch und die Wissenschaftskommunikation, ebenso wie der freie Zugriff (Open Access) zu den Publikationen. Aus diesem Grund sollten Publikationen und ihre Metadaten auf institutionellen oder disziplinären Dokumentenservern, in so genannten Repositorien und Online-Archiven auffindbar, durchsuchbar und austauschbar sein. Die Verwendung von Standards ist dafür eine Voraussetzung. Mit der Open Source Software MyCoRe können Repositorien, Dokumentenservern und Archiven entwickelt werden. Diese MyCoRe-Anwendungen ermöglichen die Verwaltung und den Austausch von Metadaten über Webschnittstellen.

Abstract

Metadata are essential for scientific exchange and communication, as well as open access to publications is. To provide these conditions, publications have to be findable in institutional and disciplinary document servers, in so-called repositories and online archives. The use of standards is one precondition for this purpose. The open source software MyCoRe is used as a framework for the development of repositories, document servers and archives. With MyCoRe applications a web based metadata management and exchange of metadata is enabled.

Metadaten im wissenschaftlichen Kontext

„Metadaten sind Daten und gleichzeitig Daten über Daten“. So lautet die Definition von Tim Berners-Lee, der feststellt, dass Metadaten die ersten Daten im World Wide Web waren, nämlich Informationen über Informationen (Berners-Lee 1999, p. 181). Darüber hinaus definiert Berner-Lee Metadaten als strukturierte Daten über digitale und nicht-digitale Ressourcen, die maschinenlesbar sind (Berners-Lee 1997). Dieser Definition folgend soll es hier um Metadaten gehen, die im Zusammenhang mit wissenschaftlicher Literatur und anderem wissenschaftlich relevanten Material digital erfasst und verwaltet werden. Metadaten können nach (Hurley et al. 1999) grob unterschieden werden in

- beschreibende Metadaten, die den Inhalt beschreiben und die Ressource durchsuchbar und erschließbar machen, z.B. der Titel eines Artikels oder der Autorenname,
- strukturelle Metadaten, die Aussagen über den Kontext machen, in den die Ressource eingebunden ist, z.B. die Seitenzahlen eines Artikels, die ISBN eines Buches oder eine Klassifikation und schließlich

- administrative Metadaten, die Auskunft über die Art der Ressource geben, etwa Format, Typ oder Größe des beschriebenen Objektes, die aber auch die Zugriffsrechte wiedergeben, um nur einige Beispiele zu nennen.

Um die Wissenschaftskommunikation zu unterstützen, müssen Metadaten von wissenschaftlichen Publikationen und digitalen Ressourcen verfügbar und zugreifbar sein. Durch den Einsatz webbasierter Anwendungen kann solch ein Zugang zu Metadaten und den wissenschaftlichen Publikationen gewährleistet werden. Voraussetzung dafür ist einerseits der freie Zugang (*Open Access*¹) zu Repositorien², die die Daten bereithalten. Andererseits müssen die Metadaten ausgetauscht werden. Erst die Interoperabilität der Metadaten macht einen breiten Zugriff auf die Daten möglich. Dies setzt wiederum die Verwendung von Standards voraus. All diese Anforderungen werden von der im Folgenden vorgestellten *Open Source Software MyCoRe* unterstützt.

MyCoRe

„MyCoRe“ ist eine *Open Source Software*, die in einem Projekt³ an mehreren deutschen Hochschulen entwickelt und eingesetzt wird. Das Projekt stellt eine Softwarebasis (*Framework*) für die Entwicklung von Repositorien bereit. Das Akronym MyCoRe steht für „**My Content Repository**“, womit die MyCoRe Software benannt wird, die den Kern zur Entwicklung eigener Anwendungen und Repositorien bildet. Dieser Kern kann an die eigenen Bedürfnisse angepasst und erweitert werden (daher das "My" für die lokale Adaption). Für alle Grundfunktionen eines Dokumenten- und Publikationsservers bietet MyCoRe Software-Komponenten (Bausteinen) an, so dass eigene Anwendungen weitgehend durch Konfiguration realisiert werden können. MyCoRe-Anwendungen gibt es für folgende Einsatzbereiche; es wird jeweils exemplarisch eine Beispielanwendung genannt:

- **Archive**

Die Papyri-Sammlung⁴ der Universitäten Leipzig, Jena und Halle, die eine umfangreiche Papyrus-Sammlung katalogisiert und digitalisiert online bereitstellt.

- **Zeitschriftenportale**

Das Journals@Urmel⁵ an der Universitätsbibliothek Jena mit einem großen Bestand an Online-Publikationen.

- **Bibliothekssysteme**

DuEPublico⁶ an der Universität Duisburg Essen, das neben der Universitätsbibliographie online auch Dissertationen, Open Access E-Publikationen,

¹ Siehe die Informationsplattform open-access.net unter <http://open-access.net/>, die freien Zugang zu wissenschaftlichen Informationen propagiert.

² Eine große Anzahl an Open Access Repositorien ist in dem Verzeichnis „OpenDOAR“ unter <http://www.opendoar.org/> erfasst.

³ Siehe die Homepage des MyCoRe-Projektes unter <http://www.mycore.de/>.

⁴ Papyri unter <http://papyri.dl.uni-leipzig.de/>

⁵ Journals@Urmel unter <http://zs.thulb.uni-jena.de/content/below/index.xml>

⁶ DuEPublico unter <http://duepublico.uni-duisburg-essen.de/>

Semesterapparate, Lehr- und Lernmaterial bereithält. Dieses System ist das Ursprungssystem von MyCoRe, es basiert auf Miless (Lützenkirchen 2002).

- **Dokumentenserver**

DOL⁷, der Dokumentenserver der Universität Leipzig ermöglicht eine Suche nach digitaler Dokumenten und multimedialen Objekten.

- **Lexikon**

LexM⁸, das Lexikon verfolgter Musiker und Musikerinnen der NS-Zeit an der Universität Hamburg

- **Verlag**

Hamburg UP⁹, der Verlag an der Bibliothek der Universität Hamburg.

Eine Anwendung – DocPortal genannt – dient als Beispielapplikation für einen Dokumentenserver z.B. zur Verwaltung von Dissertationen mit ihren Metadaten oder anderen Publikationen. Weitere Anwendungen sind auf der MyCoRe-Web-seite www.mycore.de aufgelistet. Darüber hinaus sind weitere Projekte in Entwicklung, zum Teil mit sehr großen Datenbeständen. Alle MyCoRe-Anwendungen haben ein eigens an die speziellen Anforderungen angepasstes Metadatenmodelle. MyCoRe stellt einen Satz an vordefinierten Basisdatentypen für die Metadaten bereit, so dass Textfelder, Datumswerte, Zahlen, Wahr-/Falschwerte, Verweise und Kategorien beschrieben werden können. Eigene Datentypen können zusätzlich über Java-Schnittstellen implementiert werden. Darüber hinaus können Objekttypen wie "Dokument" oder "Person" als Menge von Feldern wiederholbarer Basisdatentypen definiert werden. Über Querverweise und Beziehungen zwischen Objekttypen lassen sich Verbindungen zwischen den Metadaten herstellen. Komplexe Metadatenstrukturen lassen sich über Hierarchien von Objekttypen abbilden, die mit Vererbung von Metadaten an untergeordnete Typen, z. B. Hierarchie Zeitschrift, Heft, Artikel realisiert werden. Bei der Entwicklung einer MyCoRe-Anwendung kann auf die jeweiligen Anforderungen individuell eingegangen werden. Verwendet wird Java- und XML/XSL-Technologien.

Metadaten-Management

Zur Verwaltung von Metadaten gehört die Eingabe der Daten, deren Verarbeitung und die Ausgabe. Je nach Art der Anwendung sind die Eingabeschnittstellen unterschiedlich und anpassbar. Für die Erfassung einzelner Datensätze stellt MyCoRe individuell anpassbare Formulare bereit. Eine Validierung der Daten ermöglicht die Prüfung der Eingabe auf Vollständigkeit und bedingt auf Richtigkeit, etwa bei Datumsangaben. Auch Plausibilitätsprüfungen können vorgenommen werden, so muss z.B. eine E-Mail-Adresse immer ein @-Zeichen enthalten. Mittels Datei-Upload können zu den Metadaten z.B. Dokumente, Bilder oder andere Medien

⁷ DOL unter <http://dol.dl.uni-leipzig.de/index.xml>

⁸ LexM unter <http://www.lexm.uni-hamburg.de/>

⁹ Hamburg UP unter <http://hup.rrz.uni-hamburg.de/>

hochgeladen werden. Sollen große Datenmengen importiert werden, so steht dafür ein Kommandozeilen-Interface (CLI) zur Verfügung.

Ein wichtiger Schritt bei der Verarbeitung von Metadaten ist deren Überprüfung und Begutachtung. In MyCoRe ist ein Workflow implementiert, der diesem Anspruch Rechnung tragen will. Benutzende, die mit den entsprechenden Rechten versehen sind und die sich über die Webschnittstelle angemeldet haben, können die Metadaten weiterbearbeiten, überprüfen, speichern oder löschen. Auch könne sie anderen Personen die Rechte zur Weiterbearbeitung der Metadaten oder auch einzelner Objekte geben bzw. nehmen.

Sind die Metadaten erfasst und veröffentlicht, dann können sie über die Webschnittstelle recherchiert werden. Die Suchfunktion in MyCoRe umfasst eine einfache oder auch kombinierte Suche in beschreibenden Metadaten, in Volltexten oder in extrahierten Daten. Ebenso ist die Suche in Volltexten von PDF-, Word-, OpenOffice-, HTML- und XML-Dateien möglich. Des Weiteren kann in extrahierten Metadaten verschiedener Medien gesucht werden, z. B. ID3-Tags von MP3-Dateien, EXIF-Daten in JPEG-Dateien, Dateityp und Dateigröße. Eine qualifizierte Suche in XML-Dokumenten, z. B. in der manifest.xml Datei eines SCORM-Paketes¹⁰ oder einer Excel-Tabelle im XML-Format ermöglicht das Suchen in den Metadaten der Metadaten. Gängige Verfahren zur Begrenzung des Suchraumes sind die Phrasensuche und Suche mit Platzhaltern (*,?), Links-/Mitte-/Rechts-Trunkierung, ebenso wie Boolesche Operatoren (UND, ODER, NICHT) oder die Suche mit Vergleichsoperatoren (=, <, >, <=, >=, like, contains). Darüber hinaus sind die Suchmasken frei konfigurierbar und so an die Anforderungen der jeweiligen Anwendung anpassbar. Mit der MyCoRe-Abfragesprache MCRQL (MyCoRe Query Language) können eigene Suchanfragen formuliert werden. Abschließend sei noch erwähnt, dass eine verteilte Suche über mehrere Server via WebServices möglich ist.

Interoperabilität von Metadaten

Der Schlüssel zum Austausch von Metadaten ist die Standardisierung sowohl auf technischer als auch auf semantischer Ebene. Technisch sind Schnittstellen die Voraussetzung für Interoperabilität. MyCoRe bietet zu diesem Zweck verschiedene Systemschnittstellen an. Da MyCoRe die Metadaten auf der Basis der *Extensible Markup Language* (XML) verarbeitet, ist eine einheitliche Syntax gegeben, die grundsätzlich einen Austausch über gängige Protokolle (wie z.B. das *Hypertext Transfer Protocol* HTTP) erlaubt. Darüber wird das OAI-Protokoll¹¹ der *Open Archive Initiative* (OAI) (Lagoze et al. 2002) von MyCoRe unterstützt. Es ermöglicht dem so genannten *Harvester* Metadaten aus den angeschlossenen Repositorien auf einen zentralen Server zu laden und für eine Suche zur Verfügung zu stellen. Die Repositorien der *Dataprovider* liefern so über das OAI-Protokoll Metadaten in einem

¹⁰ SCORM (*Sharable Content Object Reference Model*) ist ein Referenz-Modell für austauschbare elektronische Lerninhalte der Advanced Distributed Learning Initiative.

¹¹ Siehe *The Open Archives Initiative Protocol for Metadata Harvesting* unter <http://www.openarchives.org/OAI/openarchivesprotocol.html>

standardisierten Metadatenformat an die Datenbank des *Serviceprovider*. Die Repositorien, die das OAI-Protokoll zum Austausch komplexer Metadatenschemata nutzen, entsprechen damit der Empfehlung der Deutsche Initiative für Netzwerkinformation (DINI) zum elektronischen Publizieren an Hochschulen, die detailliert auf die inhaltliche Gestaltung der OAI-Schnittstelle eingeht (DINI 2005). Des Weiteren bietet MyCoRe eine Z39.50¹² Server Schnittstelle. Z39.50 ist ein internationaler Standard für ein Netzwerkprotokoll zum Austausch bibliographischer Daten im Bibliotheksbereich (ANSI/NISO Z39.50-2003), wie sie in MyCoRe-Anwendungen gespeichert werden können. Eine weitere Schnittstelle ist in MyCoRe durch *Web Services* (SOAP) gegeben. Web Services können implementiert werden für die Suche in MyCoRe-Repositorien. Damit die Objekte mit den Metadaten in den Repositorien über Suchmaschinen, wie z.B. Google gefunden werden, kann in den MyCoRe-Anwendungen ein Sitemapprotokoll angelegt werden, das der Suchmaschine die URLs zu den Publikationen bereitstellt.

Semantische Interoperabilität geht über die technische weit hinaus und bringt entsprechende Probleme mit sich. Wenn Metadaten ausgetauscht werden, dann muss es auf beiden Seiten ein Verständnis von deren Bedeutung geben. Day bemerkt dazu: „there is then a need to consider the greater problem of semantic interoperability, e.g. dealing with differences in terminology and meaning across domains. This can be very problematic.“ (Day 2005: 17). Eine einfache Lösung ist ein Mapping von Metadaten oder auch eine Transformation. Wenn Standards eingesetzt werden, um Metadaten zu beschreiben, ist eine Abbildung eines Metadatenschemas auf ein anderes relativ einfach. Ein sehr weit verbreitetes Datenmodell ist das Dublin Core¹³ Metadaten-Schema zur Beschreibung von Dokumenten. Dieses Metadatenmodell wird z.B. in der MyCoRe-Beispielanwendung DocPortal eingesetzt. Weitere Ansätze, die Interoperabilität von Metadaten zu unterstützen sind XMetaDiss¹⁴, das eine xml-basierte Struktur zur Beschreibung von Online-Hochschulschriften bietet und Epicur¹⁵, ein Projekt zur Vereinheitlichung von Persistenten Identifiern (PI).

Klassifikation von Metadaten

Die Austauschbarkeit von Metadaten wird erhöht, wenn nicht nur Standards für die Metadatenformate verwendet werden, sondern auch einheitliche Kategorien für die Daten vergeben werden. Es gibt eine Reihe von etablierten Klassifikationen, wie z.B. die Dewey-Dezimalklassifikation (DDC), das Physik und Astronomie Klassifikationsschema PACS, die „Mathematical Subject Classification“ MSC oder die Sachgruppen der Deutschen Nationalbibliothek, die in MyCoRe-Anwendungen

¹² Siehe Z39.50 Resource Page der National Information Standards Organization, http://www.niso.org/standards/resources/Z3950_Resources.html

¹³ Dublin Core Metadata Initiative unter <http://www.dublincore.org/>

¹⁴ XMetaDiss, Format des Metadatensatzes der DNB für Online-Hochschulschriften unter <http://www.d-nb.de/standards/xmetadiss/xmetadiss.htm>

¹⁵ EPICUR unter <http://www.ddb.de/wir/projekte/epicur.htm>

eingebunden werden können, um die Kategorisierung von Metadaten zu ermöglichen. Auch Sprachen und Länder sind als Kategorien in MyCoRe verfügbar. Darüber hinaus können Objekte (Dokumente oder andere Medien) einer oder mehreren Kategorien in beliebigen Klassifikationen zugeordnet werden. Über die Webschnittstelle kann in der Klassifikationsstruktur navigiert und mit einer gezielten Suche verknüpft werden. Anpassungen der Klassifikationen sind über einen Online-Klassifikationseditor interaktiv möglich. Auch können Klassifikationen über XML-Dateien importiert oder exportiert werden.

Ausblick

Metadaten-Management mit MyCoRe ist für eine breite Palette von Anwendungsgebieten realisiert. Mit dem MyCoRe-Software-Kern können weitere Anwendungen entwickelt werden, die die Anforderungen an institutionellen Repositorien erfüllen und den freien Zugang zu wissenschaftlichen Publikationen unterstützen. Da die Software in einem Open Source Projekt entwickelt wird, ist eine Mitarbeit möglich und es kann individuell Einfluss genommen werden auf die Weiterentwicklung von MyCoRe. Ziel ist es, den freien Zugang zu Publikationen zu ermöglichen und die wissenschaftlichen Ressourcen über Metadaten erschließbar zu machen. Mit MyCoRe-Anwendungen ist dies möglich.

Literatur

ANSI/NISO Z39.50 – 2003: Information Retrieval: Application Service Definition & Protocol Specification. Bethesda, Maryland, USA: NISO Press NISO, 2002, from http://www.niso.org/standards/standard_detail.cfm?std_id=465

Berners-Lee, Tim (1997): Metadata Architecture. Documents, Metadata, and Links. January 1997, last edit date: 2000-09-21, <http://www.w3.org/DesignIssues/Metadata.html>

Berners-Lee, Tim (1999): Weaving the Web. New York, HarperCollins, 1999

Day, M. (2005): "Metadata", DCC Digital Curation Manual, S.Ross, M.Day (eds), November 2005, retrieved 8th August 2007, from <http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata/>

DINI Arbeitsgruppe Elektronisches Publizieren (2005): Elektronisches Publizieren an Hochschulen. Inhaltliche Gestaltung der OAI-Schnittstelle. Empfehlungen, DINI Schriften 2-de, Version 2.0, September 2005, <http://nbn-resolving.de/urn:nbn:de:kobv:11-10049220>

Hurley, B. J., Price-Wilkin, J., Proffitt, M., Besser, H. (1999). The Making of America II. Testbed Project: a digital library service model, Washington, D.C.: Council on Library and Information Resources. Retrieved January 30,2006, from <http://www.clir.org/pubs/abstract/pub87abst.html>

Lagoze, C., Van de Sompel, H., Nelson, M., & Warner, S. (2002). The Open Archives Initiative Protocol for Metadata Harvesting, v. 2.0, 2002-06-14, <http://www.openarchives.org/OAI/openarchivesprotocol.html>

Lützenkirchen, F. (2002): MyCoRe - Ein Open-Source-System zum Aufbau digitaler Bibliotheken. Datenbank Spektrum, 2(4), November 2002, 23-27, http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-11198/MyCoRe_DB-Spektrum.pdf

Alle Web-Seiten wurden am 8.8.2007 geprüft.

Open-Access-Zeitschriften als neue Form wissenschaftlicher Kommunikation: Vorbehalte und Vorschläge für Maßnahmen zur Akzeptanzsteigerung

Karin Weishaupt

Abstract

Gegen das Open-Access-Publizieren bestehen diverse Vorbehalte, die zum Teil schwer zu entkräften sind. Die Bedürfnisse der potenziellen Autorinnen und Autoren müssen fachspezifisch systematisch ermittelt werden, damit daraus Maßnahmen zur Akzeptanzsteigerung abgeleitet werden können und die Vorteile vor den Nachteilen überwiegen.

Many authors are rather reluctant against open access and it is difficult to counter part of their inhibitions. Their needs must be examined systematically making distinctions between the different disciplines so that measures can be deduced to increase acceptance of open access. The aim is that advantages will prevail to disadvantages.

Ausgangslage

Als die Deutsche Forschungsgemeinschaft im Spätherbst 2004 ihre „Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access“ (DFG 2005) durchführte, ging es noch zum großen Teil um die Frage der Bekanntheit von Möglichkeiten des Open-Access-Publizierens. Inzwischen sollte sich das Bild geändert haben: Das Thema hat Eingang selbst in populäre Computerzeitschriften wie die „c't“ (Sietmann 2006) und den „Technology review“ (Heuer 2007) und wird vor allem im Zusammenhang mit der Reform des Urheberrechts öffentlich heftig diskutiert (Dambeck 2007).

Aber noch immer ist die Umsetzung der Open-Access-Idee längst noch nicht so weit gediehen, wie es aufgrund der technischen Entwicklung möglich wäre. Klagen darüber finden sich immer wieder, hier beispielhaft eine Aussage zum Hochschulschriftenserver der Humboldt-Universität zu Berlin: „Der Blick in die Praxis zeigt, dass der Transformationsprozess noch ganz am Anfang steht. ‚An der HUB werden pro Jahr etwa 2200 Artikel verfasst, weniger als ein Prozent davon liegen auf unserem Institutional Repository‘, berichtet der Direktor des Rechenzentrums der Humboldt-Universität und Vorstand der Deutschen Initiative für Netzwerk-Information (DINI), Peter Schirmbacher; ‚Wir müssen die Autoren gewinnen, dass sie Open Access praktizieren.‘“ (Sietmann 2006, S. 192)

Das Resümee der Tagung „Digitales Publizieren in den Geisteswissenschaften“ der Europäischen Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen am 30./31.3.2007 in Bad Neuenahr-Ahrweiler klingt ähnlich: „Digitales Publizieren bestimmt dabei noch keineswegs den universitären Alltag. Im Gegenteil:

Professorin Dr. Gudrun Gersmann (Historisches Seminar der Universität zu Köln), ihrerseits Herausgeberin elektronischer Zeitschriften, stellte fest, dass es im Wesentlichen kleine Kreise sind, die die Angebote intensiv nutzen. Viele Wissenschaftler begegnen den Neuerungen gar mit Skepsis. So zählen in vielen Berufungsverfahren elektronische Publikationen weniger als solche in traditionellen Fachzeitschriften – und dies, obwohl es nach Meinung der Fachleute inzwischen zahlreiche e-Journals gibt, die hinsichtlich der Qualitätsstandards mit gedruckten Publikationsorganen ohne weiteres mithalten können.“ (Europäische Akademie 2007a)

Ganz krass beschreibt die Informationsplattform Open Access den aktuellen Stand: „In den Sozialwissenschaften spielt Open Access bislang keine zentrale Rolle.“¹

Andererseits nimmt die Zahl der Open-Access-Publikationen stetig zu. Mit Stand vom 11.7.2007 verzeichnet das von der Universität Lund erstellte „Directory of Open Access Journals“ 2750 Zeitschriften, 30 davon sind in den letzten 30 Tagen neu eingefügt worden; 139933 Zeitschriftenaufsätze sind verzeichnet – eine durchaus beträchtliche Anzahl². Sowohl die Klagen über mangelnde Akzeptanz als auch die Erfolgsmeldungen über stetig wachsende Zahlen von Open-Access-Publikationen sind also sehr differenziert zu betrachten.

Vorbehalte gegen Open Access seitens der Autorinnen und Autoren

Die Studie „Open Access & Scientific Publishing“ der Universität München und der University of Arkansas at Little Rock, deren Ergebnisse Anfang 2007 publiziert worden sind, untersucht u. a. die Haltung von Wissenschaftler/inne/n aus den Bereichen Wirtschaftsinformatik, deutsche Literatur und Medizin gegenüber Open Access. Die grundsätzliche Einstellung ist durchweg sehr positiv. Auf einer Fünfer-Skala von „very positive“ bis „very negative“ lagen 91% der Antworten bei der besten und zweitbesten Bewertung. Die Frage nach der Nutzung vermittelte bereits ein anderes Bild: Circa zwei Drittel der Befragten hatte bereits auf Open-Access-Publikationen zugegriffen, und nur ein Drittel hatte selbst in Open-Access-Medien veröffentlicht. Die Frage, wer beabsichtigte, in den nächsten sechs Monaten im Open-Access-Bereich zu publizieren, erbrachte die ungünstigsten Werte: Je nach Wissenschaftsbereich hielten das nur 12 - 17% der Befragten für sehr wahrscheinlich: „There is a gap between the positive attitude towards Open Access and the low level of use and future intention to use OpenAccess Media.“

(Hess / Wigand / Mann / von Walter 2007, S. 5-8)



(Abb. 1, Quelle: Hess / Wigand / Mann / von Walter 2007, S. 8)

¹ http://open-access.net/de/oa_in_verschiedenen_faechern/sozialwissenschaften/, Stand: 11.7.2007

² <http://www.doaj.org>, Stand: 11.7.2007

Peter Schirmbacher sieht im Übergang vom Publizieren in gedruckten zu elektronischen Dokumenten einen gravierenden Wandel in der wissenschaftlichen Kommunikation: „Ein solch gravierender Umbruch ist nicht leicht zu bewerkstelligen, denn nahezu jede Veränderung birgt Gefahren in sich, die nicht nur Vorteile für alle Beteiligten, sondern auch nicht sofort überschaubare Auswirkungen haben kann. Kritik an Open Access kommt von etablierten Wissenschaftsverlagen, aber auch von Autoren, die um Einnahmen aus Autorenverträgen fürchten. Vor allem gerät das System der Bewertung der wissenschaftlichen Leistungen, das bisher von den Verlagen organisiert wurde, ins Schlingern, wenn nun jeder seine Erkenntnisse ins Netz stellt und die Auszeichnung, eine Veröffentlichung in einem so genannten renommierten wissenschaftlichen Journal zu platzieren, nichts mehr gelten soll.“ (Schirmbacher 2007a, S. 23f.)

Hier werden bereits einige der am häufigsten genannten Vorbehalte angerissen, die die Ausbreitung von Open-Access-Publikationen behindern; einer davon betrifft die Frage der Kosten und der Geschäftsmodelle. Auch Open-Access-Zeitschriften und andere Publikationen werden nicht ohne Kosten produziert. Das, was im herkömmlichen Produktionsprozess die Verlage leisten, muss nun zum Teil von den Autor/inn/en selbst und – im Falle von Zeitschriften - von den Redaktionen geleistet werden. In vielen Fällen – insbesondere im Bereich der Medizin – werden Open-Access-Zeitschriften von Fachgesellschaften getragen, das heißt, die Redaktionsarbeit wird über diese finanziert. Alternativ sind die Zeitschriftenredaktionen bei Hochschulbibliotheken und Forschungseinrichtungen angesiedelt und müssen oft neben ihren sonstigen Aufgaben die Herausgebertätigkeiten abwickeln. Ein verbreitetes Finanzierungsmodell besteht darin, dass die Autor/inn/en die Kosten für den Produktionsprozess tragen – es kann sich dabei um einen vierstelligen Betrag pro Aufsatz handeln! Anstelle von Honoraren kommen also oft auf die Autor/inn/en erhebliche Kosten zu.

Ein weiterer Punkt, der in Befragungen und in der Literatur immer wieder auftaucht, ist das Problem der Langfristverfügbarkeit. Dieses hat mehrere Dimensionen. Es geht erstens um die Frage der Speichermedien: Wie kann man sicher stellen, dass die Datenträger nach längeren Zeiträumen noch funktionsfähig sind, oder welche Verfahren muss man einsetzen, um die Dateien immer wieder neu auf zukunftsträglichen Medien zu speichern? Diese Frage betrifft den Komplex der Hardware-Probleme. Zweitens geht es um die Frage der Dateiformate: Wie kann man Dokumente so abspeichern, dass sie nach längeren Zeiträumen noch lesbar sind? Hier geht es um Dateiformate und Software zum Lesen der Dateien. Drittens geht es um das bekannte Problem der Flüchtigkeit von Internet-Adressen: Wie kann gewährleistet werden, dass Dateien zwar physikalisch noch lesbar sind, aber wegen eines „error 404 – file not found“ nicht mehr auffindbar sind? Für dieses Problem gibt es zwar mehrere Lösungsansätze in Form von dauerhaften Adressierungen – DOI, URN, PURL – (siehe dazu Kunze 2005); aber ob diese bekannt sind und Vertrauen erwecken, erscheint fraglich.

Eine Liste von Vorbehalten gegen Open Access findet sich auf der Open-Access-Plattform, die Anfang Mai 2007 ans Netz gegangen ist:

- „Qualitätsvorbehalte
- Fragen der Auffindbarkeit
- Langzeitarchivierung von Dokumenten
- Vorbehalte hinsichtlich der Authentizität der Dokumente
- Rechtliche Vorbehalte
- Finanzierbarkeit des Author-Pays-Modells
- Vorbehalte hinsichtlich Publikationsmittelverteilung und Interessenkonflikten
- Zeitaufwand für Wissenschaftler und Wissenschaftlerinnen
- Open Access als Herausforderung für Non-Profit-Verlage
- Keine Auswahl der Zielgruppe von Publikationen möglich“³

An einigen dieser Punkte, die objektiv Probleme darstellen, wird zurzeit gearbeitet wie zum Beispiel den genannten Problemen der Langfristverfügbarkeit; auch Maßnahmen zur Sicherung der Authentizität der Dokumente sind möglich und werden umgesetzt. Schwieriger wird es bei dem Komplex von Vorbehalten, der eher psychologisch zu begründen ist.

Dienen Open-Access-Publikationen der Reputationssteigerung? Diese Frage wird häufig gestellt. Als Gegenmaßnahme ist denkbar, dass Institutionen der Forschungsförderung gezielt darauf hinweisen, dass in Projektanträgen solche Veröffentlichungen denselben Stellenwert haben wie gedruckte, und dies in die Praxis umsetzen. Bei Berufsverfahren und Stellenbesetzungen müssten Hochschulen und andere wissenschaftliche Einrichtungen ebenso damit umgehen. Viele Open-Access-Zeitschriften sind noch sehr jung und können daher nicht dasselbe Renommee erworben haben wie ältere etablierte Zeitschriften. Hier wäre es hilfreich, wenn Print-Zeitschriften, die bisher gedruckt oder als Hybrid-Zeitschriften in kommerziellen Verlagen erschienen sind, auf das Open-Access-Modell umstellen würden, damit ihr guter Ruf in die Open-Access-Welt übertragen wird.

Es gibt aber noch „weichere“ psychologische Faktoren, denen noch schwerer zu begegnen ist. Das Sprichwort: „Was nichts kostet, taugt nichts“ lässt sich mit Sicherheit auf die Frage der Akzeptanz des Open-Access-Publizierens anwenden, auch wenn kaum jemand ein solch banales Gegenargument offen zugeben würde.

Wer kann nicht nachvollziehen, dass es nach einer längeren mühsamen Gedankenarbeit ein echtes Erfolgserlebnis darstellt, ein gut aufgemachtes Buch oder ein Zeitschriftenheft in ansprechendem Layout in der Hand zu halten, in dem der eigene Name und der eigene Text gedruckt vorliegen? Beim Open-Access-Publizieren entfällt dieser Autorenstolz weitgehend – eine Datei zum Herunterladen hat nicht dieselbe Wirkung wie ein Druckerzeugnis mit optisch ansprechender Aufmachung. Für den „haptischen Wert“ von Büchern und Zeitschriftenheften, der vor

³ http://open-access.net/de/allgemeines/gruende_und_vorbehalte/vorbehalte_gegen_oa/#c562
Stand: 10.7.2007

allem von Geisteswissenschaftler/innen betont wird und auf einer Tagung zum Thema „Digitales Publizieren in den Geisteswissenschaften“ im März 2007 (Europäische Akademie 2007b) wiederholt angesprochen wurde, gibt es keine Entsprechung beim elektronischen Publizieren.

Fachübergreifende Maßnahmen zur Steigerung der Akzeptanz von Open Access

Aus der Erkenntnis, dass einige der beim Open-Access-Publizieren bestehenden Probleme nur schwer und andere gar nicht lösbar sind, muss nicht folgen, dass die Bemühungen um Open Access zum Scheitern verurteilt sind und diese Chance, allen Menschen freien Zugang zu Information und Wissen und insbesondere wissenschaftlichen Erkenntnissen zu gewähren⁴, nicht weiter verfolgt werden sollte. Abhilfe kann ein lösungsorientierter Ansatz schaffen.

Der Weg der Konzentration auf Lösungsansätze anstelle von Problemanalysen lässt sich schon auf Ludwig Wittgenstein zurückführen. Dieser schreibt in seinem „Tractatus logico-philosophicus“: „Die Tatsachen gehören alle zur Aufgabe, nicht zur Lösung“ (Nr. 6.4321). Der Philosoph weist mit diesen Satz darauf hin, dass „Aufgaben“ im Sinne von Problemen, die sich uns stellen, und Lösungen von gänzlich unterschiedlicher Art sein können und nicht zwangsläufig voneinander ableitbar sind. Die Problem- und Ursachenanalyse ist keine notwendige Bedingung für das Konstruieren von Lösungen (Sparrer 2001, S. 31).

Das Prinzip, nicht ein Problemsystem zu fokussieren, sondern ein Lösungssystem zu entwickeln, steht im Mittelpunkt des Ansatzes von Steve de Shazer, der zusammen mit seiner Frau Insoo Kim Berg ab 1982 den so genannten „lösungsorientierten Therapieansatz“ entwickelt und mit seinen Vorstellungen die Beratungs- und Therapiemethodik in den letzten Jahren wesentlich beeinflusst hat. „Problem talk creates problems, solution talk creates solutions“ (zitiert nach von Schlippe / Schweitzer, 2000, S. 35) lautet das Grundprinzip, auf dessen Basis ein spezielle Gesprächs- und Fragetechnik entwickelt worden ist.

Das lösungsfokussierte Arbeiten, das sich in der Psychotherapie als außerordentlich erfolgreich erwiesen hat, hat inzwischen darüber seinen Weg nicht nur in die Pädagogik, sondern vor allem auch ins Management gefunden. Nachdem sich die Anwendung ursprünglich auf die Bereiche Change Management und Coaching konzentriert hatte, ist mittlerweile diese Idee auch in Teilgebieten der Personalentwicklung (Personalauswahl, Management Development), aber auch im Strategischen Management, Verkauf und Unternehmensführung gut eingeführt. (Rothenwänder 2005)

Bezüglich Open Access bedeutet Lösungsfokussierung, da anzusetzen, wo es bereits funktioniert, und die Vorteile dieser Publikationsform so deutlich herauszuarbeiten, bekannt zu machen und nach Möglichkeit zu verstärken, dass sie vor den Nachteilen überwiegen. Damit sollen potenzielle Autor/inn/en motiviert

⁴ So begründet die Deutsche UNESCO-Kommission ihr Engagement für Open Access, siehe dazu <http://www.unesco.de/openaccess.html?&L=0>, Stand:17.7.2007

werden, sich dieser Publikationsform trotz eventuell bestehender Bedenken zu bedienen.

Aus der umfangreichen Open-Access-Literatur lässt sich eine Reihe von Vorteilen von Open-Access-Publikationen ableiten, die mit Sicherheit noch um diverse Punkte verlängert werden kann:

- Der Prozess des elektronischen Publizierens geht im Normalfall schneller als die Produktion von Printveröffentlichungen, sodass aktuelle Forschungsergebnisse zügiger der Fachwelt mitgeteilt werden können.
- Aufsätze in elektronischen Zeitschriften können einzeln unmittelbar nach der Fertigstellung veröffentlicht werden, ohne dass die Erstellung eines ganzen Heftes abgewartet werden muss – damit wird eine weitere Beschleunigung erreicht.
- Elektronische Dokumente können dynamisch angelegt werden: Der/Die Autor/in kann selbst Aktualisierungen und Ergänzungen vornehmen, es können interaktive Elemente einbezogen werden, die eine Diskussion mit der Fachwelt in die Veröffentlichung integrieren.
- Die Einbeziehung multimedialer Elemente ist möglich, also Bildmaterial in guter Qualität, Filmsequenzen, Tondokumente etc.
- Beim Zitieren können Links auf die verwendete Literatur gesetzt werden, sodass jeder Text eine Brücke zu weiterer relevanter Literatur darstellen kann.
- Aufgrund der schnellen und leichten Verfügbarkeit werden elektronische Dokumente signifikant häufiger zitiert als Texte, die lediglich gedruckt vorliegen.
- Während bei gedruckten Veröffentlichungen die Rechte meistens an den Verlag abgetreten werden, bleiben die Urheberrechte von Open-Access-Publikationen normalerweise beim/bei der Autor/in (siehe z.B. die Lizenz der Digital Peer Publishing Initiative des Landes NRW), sodass der Text nach Belieben weiter verwendet werden kann.

Die DFG stellte im Rahmen ihrer Open-Access-Studie die Frage: „Welche Maßnahmen zur Unterstützung halten Sie für sinnvoll?“ Die Bewertung auf einer Skala von 1 bis 5 durch die Befragten ergab folgende Prozentzahlen bei den Antworten „sehr sinnvoll“ und „sinnvoll“ (zusammengefasst):

- „Anreize für Publikationsorgane, ihre Beiträge im Internet für den entgeltfreien Zugriff anzubieten: 86,2 %
- Förderung sollte auf Konkurrenzfähigkeit mit konventionellen Zeitschriften abzielen: 75,0 %
- Einrichtung von entgeltfreien, zentralen, disziplinspezifischen Archiven im Internet: 73,2 %
- Förderung der Diskussion um Open Access in Universitäten und Forschungsinstitutionen: 73,1 %
- Wissenschaftler sollen aufgefordert werden, ihre eigenen Arbeiten auch im Internet bereitzustellen: 71,5 %

- Verbesserung der Beratung und Information im Bereich des Open Access für Wissenschaftler: 70,6 %
- Autorenverträge sollten entgeltfreie Veröffentlichung im Internet zulassen: 68,4 %
- Verbesserung des Angebots an Schulungskursen und Publikationstechniken für Wissenschaftler: 31,6 %“ (Mugabushaka / Fournier 2005, S. 5)

Diese Maßnahmen, die sicher noch relevant sind, wurden allerdings nicht von den Wissenschaftler/innen vorgeschlagen, sondern von der DFG vorgegeben und von den Autor/innen lediglich in ihrer Wichtigkeit bewertet.

Bei seinem Vortrag auf dem Bibliothekskongress in Leipzig im März 2007 weist Peter Schirmbacher auf eine oft einseitige Perspektive bei der Suche nach Lösungen hin: „Bei der Gradwanderung zwischen perfektem System und zufriedenem Autor / Nutzer steht häufig das System im Mittelpunkt.“ (Schirmbacher 2007b, Folie 6) Es erscheint sehr wichtig, dem entgegenzusteuern und potenzielle Autor/innen zu befragen, was sie benötigen, um sich zum Open-Access-Publizieren bereit zu erklären.

Einige Maßnahmen zur verbesserten Umsetzung des Open-Access-Gedankens insbesondere bei Zeitschriften könnten sein:

- Es reicht nicht aus, mit einer elektronischen Veröffentlichung lediglich die gedruckte Version eines Textes nachzubilden, sondern die oben aufgelisteten Vorteile des digitalen Publizierens müssen so weit wie eben möglich ausgenutzt werden.
- Eine gute Qualitätskontrolle und –sicherung muss gewährleistet sein, damit auch Open-Access-Zeitschriften eine Chance haben, sich in der *scientific community* zu etablieren. und damit auch Aufsätze in diesen Zeitschriften zur Reputationssteigerung beitragen.
- Damit sich ein Aufsatz aus einer Open-Access-Zeitschrift mit Begutachtungsverfahren von einer beliebigen Download-Datei aus dem Internet abhebt, sollte er mit einem sorgfältig gestalteten Titelblatt versehen sein. Dieses muss außer dem Aufsatztitel und der Verfasserangabe enthalten: die genaue Quellenangabe (Titel und Jahrgangs- sowie Heft-Nummer der Zeitschrift), dauerhafte Internet-Adresse, Hinweis auf die Lizenz, Hinweis auf das Begutachtungsverfahren. Letzteres ist zu empfehlen, da bei vielen Zeitschriften nur schwer ersichtlich ist, ob sie ein Begutachtungsverfahren einsetzen oder nicht; die Notwendigkeit der Recherche nach diesem Qualitätsmerkmal lässt sich durch entsprechende Hinweise bei der Zeitschrift insgesamt, aber auch bei jedem Aufsatz vermeiden.
- Es muss sicher gestellt sein, dass die Aufsätze langfristig verfügbar durch die Abspeicherung in relevanten Repositorien und auffindbar durch Verwendung von dauerhaften Adressen sind.
- Die aktuelle Auffindbarkeit muss gefördert werden, indem die Aufsätze in allen relevanten Datenbanken und Suchmaschinen verzeichnet werden.

- Die Abruf- und Zitierhäufigkeit der Aufsätze aus einer Zeitschrift muss für zukünftige Autor/inn/en derselben Zeitschrift transparent gemacht werden.
- Für die Autor/inn/en von Zeitschriftenaufsätzen darf nicht mehr Aufwand für Layout und andere Teile des Produktionsprozesses bei Open-Access-Zeitschriften entstehen als bei Print-Zeitschriften.

Fachspezifische Unterschiede bei Bedürfnissen und Maßnahmen

Johannes Fournier weist darauf hin, dass Förderungsmaßnahmen die unterschiedlichen Gegebenheiten und Bedürfnisse der einzelnen Fachgebiete berücksichtigen müssen: „Verschiedene Fachgebiete pflegen unterschiedliche Publikationspraktiken. Diesen Unterschieden ist auch in Bezug auf Open Access gerecht zu werden: Für den Biologen ist der schnellstmögliche Zugriff auf einen Zeitschriftenartikel, für den Historiker die komfortabel zugreifbare elektronische Edition, für den Geowissenschaftler die Verifizierbarkeit einer Publikation durch Zugriff auf zugrunde liegende Klimadaten wichtig. In der Diskussion mit den unterschiedlichen Fachgebieten muss daher besser definiert werden, wo und in welcher Form Open Access von besonderer Relevanz ist.“ (Fournier 2007, S. 83)

Die bisher vorliegenden Untersuchungen zur Akzeptanz von Open Access geben kein ausreichendes Material bezüglich der Frage her, welche Bedürfnisse zu diesem Thema in den einzelnen Fachgebieten bestehen. Das Mitte 2007 erschienene UNESCO-Handbuch (Deutsche UNESCO-Kommission 2007) liefert aber eine Fülle von interessanten Einzelaussagen.

In den Natur- und Lebenswissenschaften spielen ökonomische Fragen und insbesondere die viel zitierte Zeitschriftenkrise eine große Rolle. Diesen Aspekt sowie den der Transparenz und Nachvollziehbarkeit von naturwissenschaftlichen Forschungen betont Norbert Lossau: „In den Natur und Lebenswissenschaften ist die Fachzeitschrift das relevante Medium; als Nutzer stoßen Natur- und Lebenswissenschaftler aber zunehmend an Grenzen. Die kostenlose Bereitstellung über durch Bibliotheken lizenzierte Abonnements ist nicht mehr garantiert, da die Abonnements aufgrund von Preissteigerungen und sinkenden Bibliotheksetats stark zurückgefahren werden. Durch die zurückgehenden Subskriptionen sinkt gleichzeitig die Zahl der Leser und damit die Sichtbarkeit der eigenen Forschungsergebnisse in der *Community*.

Neuland ist der in der Berliner Erklärung⁵ geforderte Zugang auch zu Rohmaterialien wie den Primärdaten, da diese von den Wissenschaftlern im allgemeinen sorgsam gehütet und unter Verschluss gehalten werden. Befürworter argumentieren, Open Access könne durch die Verbreitung dieser Daten zur guten wissenschaftlichen Praxis beitragen, was angesichts einiger spektakulärer Betrugsfälle in den letzten Jahren für die Natur- und Lebenswissenschaften zunehmend relevant sein könnte.“ (Lossau 2007, S. 20)

⁵ <http://www.dipp.nrw.de/lizenzen/dppl/>, Stand: 17.7.2007

Wolfgang Voges äußert sich dagegen bedenklich: „Die Bereitschaft von Naturwissenschaftlern, freien Zugang zu Forschungsergebnissen zu gewähren, ist aufgrund der großen Vielfalt der Disziplinen und des sehr unterschiedlichen Umgangs mit wissenschaftlichem Wissen nicht einheitlich. [...] In etlichen Fachbereichen spielen z.B. kommerzielle Erwägungen bzgl. Erfindungen und Patentrechten eine so große Rolle, dass hier – weder jetzt noch in Zukunft – nicht mit kostenlosem Zugang zu Forschungsergebnissen und Daten gerechnet werden kann.“ (Voges 2007, S. 76)

Bezüglich der Geisteswissenschaften sagt Lossau: „Das gedruckte Buch ist für Historiker, Philosophen, Literaturwissenschaftler und Linguisten, Archäologen, Musikwissenschaftler, Juristen und Wirtschaftswissenschaftler nach wie vor das primäre Medium zur Verbreitung der eigenen Forschungsergebnisse. Online-Medien werden in diesen Disziplinen vorwiegend zur Recherche genutzt. Der Zugang ist prinzipiell gesichert, wenn nicht über die Bibliothek vor Ort, dann mit Zeitverzug über Fernleihe oder Dokumentenlieferung aus anderen Bibliotheken. Statt des konkreten Zugangs wird in den Geistes- und Gesellschaftswissenschaften das Potenzial des elektronischen Publizierens in den Vordergrund gerückt, beispielsweise eine erhöhte Publikationsgeschwindigkeit, die Zusammenführung verschiedener Medien (Text, Bild, Sprache, Film u. a.) und die Entwicklung neuer ‚Textsorten‘ (Hypertext). Dabei ist weniger an die Ablösung des Printmediums als an eine sinnvolle Ergänzung gedacht. Open Access selbst findet als Grundprinzip der wissenschaftlichen Infrastruktur zwar Zustimmung, ist aber kein primäres Ziel.“ (Lossau 2007, S. 20)

Gudrun Gersmann stellt überwiegend heraus, warum sich das Open-Access-Publizieren in den Geisteswissenschaften noch nicht recht durchgesetzt hat: „Obwohl der Umgang mit elektronischen Publikationen heute einen selbstverständlichen Bestandteil der Lehr- und Forschungstätigkeit des Historikers oder Literaturwissenschaftlers darstellt, ist die Debatte über Open Access in den Geisteswissenschaften bisher erst auf geringe Resonanz gestoßen. Anders als in den STM-Fächern (*Science, Technology, Medicine*) wissen nur wenige Fachvertreter im Detail, was die Forderung nach Open Access bedeutet, geschweige denn, dass sie dem Ruf nach Archivierung ihrer Texte auf Hochschulschriftenservern oder in fachspezifischen Repositorien Folge leisten würden. Die Gründe für die momentan noch überwiegend passive Haltung gegenüber elektronischen Publikationsformen liegen auf verschiedenen Ebenen: Vielfach herrscht schlicht und einfach Unkenntnis in Bezug auf den Entstehungsprozess und die Stabilität digitaler Publikationen. Weit verbreitet ist offenbar die – falsche – Vorstellung, elektronisches Publizieren erfordere ein so hohes Maß an eigener EDV-Kompetenz, dass der Laie dazu gar nicht in der Lage sei. Ebenso häufig wird die Vermutung geäußert, elektronische Publikationen seien per se flüchtig und peripher, weil die Langzeitarchivierung nicht garantiert werden könne. Damit verknüpft hört man häufig auch das Argument, dass die großen und wichtigen Fachbeiträge selbstverständlich nur in Büchern präsentiert werden könnten: ‚Sollen wir etwa Hegel am Bildschirm lesen? Unmöglich!‘ Dass bei solchen Aussagen zwei Betrachtungsebenen – die Frage nach der Materialität des

Mediums und die Frage nach der Qualität wissenschaftlicher Texte – miteinander vermengt werden, die gar nichts miteinander zu tun haben, spielt keine Rolle. Als letztes sei schließlich die Angst genannt, durch die Partizipation an der ‚Spaßkultur‘ des Netzes die eigene Seriosität unter Fachkollegen einzubüßen.“ (Gersmann 2007, S. 78)

Für die Geistes- und Sozialwissenschaften wurden auf der bereits erwähnten Tagung zum digitalen Publizieren im März 2007 wichtige Aspekte herausgehoben: „Dr. Günter Mey und Dr. Katja Muck von der Freien Universität Berlin zeigten die Potentiale netzbasierten wissenschaftlichen Arbeitens am Beispiel des von ihnen betriebenen Portals ‚qualitative-research.net‘ auf. Zwei Teilnehmer von der Universität Hamburg, Dr. Stefan Gradmann von der dortigen Bibliothek und der Literaturwissenschaftler Professor Dr. Jan C. Meister, unterstrichen den möglichen Mehrwert des digitalen Texts gegenüber der gedruckten Alternative, wenn intelligente Software – ‚heuristische Maschinen‘ – den wissenschaftlichen Interpreten bei der Analyse der Texte unterstützen.“ (Europäische Akademie 2007b)

Eigenes Vorhaben

Wie diese Statements zeigen, stellt sich die Situation in den verschiedenen Fachgebieten ganz unterschiedlich dar.

In einem Projekt des Instituts Arbeit und Technik in Gelsenkirchen sollen die fachspezifischen Bedürfnisse und vor allem die daraus abzuleitenden Maßnahmen systematisch herausgearbeitet werden. Nach dem Prinzip der Lösungsfokussierung ist geplant, existierende Open-Access-Zeitschriften verschiedener Fachgebiete auf ihre Erfolgsfaktoren hin zu untersuchen. Nach Vorgesprächen mit den Herausgeber- und Redaktionsteams sollen die Autor/inn/en verschiedener Open-Access-Zeitschriften befragt werden, was sie zum Publizieren in diesem Medium bewogen hat, wie sie mit dem Erfolg zufrieden sind, welche Vorteile und welche Verbesserungsmöglichkeiten sie sehen. An diese erste Online-Befragung, in der jede Idee zur Förderung des Open-Access-Gedankens willkommen ist, sollen die Ergebnisse in Anlehnung an die Delphi-Methode noch einmal allen Autor/inn/en vorgelegt werden und auf ihre Relevanz hin geprüft werden. Damit soll ausgeschlossen werden, dass Meinungen, die von Einzelpersonen geäußert, aber von niemandem geteilt werden, ein zu großes Gewicht bekommen.

Die Ergebnisse dieser Untersuchung werden den Verantwortlichen für Open-Access-Zeitschriften und Repositorien, aber auch Autor/inn/en zur Verfügung gestellt – selbstverständlich als Open-Access-Publikation –, um damit den Weg zu freiem Zugang zu Information und Wissen weiter zu ebnen.

Literatur

Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen, 2003:

Berlin:Max-Planck-Gesellschaft, 22.10.2003.

http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf

Dambeck, H., 2007: Deutschen Forschern blüht Zukunft mit Fax und Papierkopien.

In: Spiegel Online, 06. Juli 2007, 3 S.

<http://www.spiegel.de/wissenschaft/mensch/0,1518,492954,00.html>

Deutsche Forschungsgemeinschaft, 2005: Publikationsstrategien im Wandel?

Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access. Weinheim: Wiley-VCH-Verl. ISBN 978-3-527-31535-2.

http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/download/oa_ber_dt.pdf

Deutsche UNESCO Kommission, 2007: Open Access – Chancen und

Herausforderungen: ein Handbuch. Bonn.

http://www.unesco.de/fileadmin/medien/Dokumente/Kommunikation/Handbuch_Open_Access.pdf

Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer

Entwicklungen, 2007a: Frühjahrstagung „Digitales Publizieren in den

Geisteswissenschaften“: Pressemitteilung, 2. April 2007. Bad-Neuenahr-Ahrweiler.

[http://www.ea-](http://www.ea-aw.de/susanis/global/download.php/news/pdf/id/37/FT07_Presse_Danach.pdf)

[aw.de/susanis/global/download.php/news/pdf/id/37/FT07_Presse_Danach.pdf](http://www.ea-aw.de/susanis/global/download.php/news/pdf/id/37/FT07_Presse_Danach.pdf)

Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer

Entwicklungen, 2007b: Frühjahrstagung 2007: Digitales Publizieren in den

Geisteswissenschaften: 30.-31.3.2007, SETA-Hotel, Bad Neuenahr-Ahrweiler.

http://www.ea-aw.de/pages/veranstaltungen/frueh_herbsttagung_2007frueh.php

Fournier, J., 2007: Open Access in der Deutschen Forschungsgemeinschaft. In:

Deutsche UNESCO Kommission: Open Access – Chancen und Herausforderungen: ein Handbuch. Bonn, S. 82-83

Gersmann, G., 2007: Open Access in den Geisteswissenschaften. In: Deutsche

UNESCO Kommission: Open Access – Chancen und Herausforderungen: ein Handbuch, S. 78-79

Hess, T. / Wigand, R. T. / Mann, F. / Walter, B. von, 2007: Open access & science publishing: results of a study on researchers' acceptance and use of open access publishing. München: Univ. Management reports of the Institute of Information Systems and New Media in cooperation with: University of Arkansas at Little Rock, USA, Department of Information Science, no. 1/07, S. 5-8. http://openaccess-study.com/Hess_Wigand_Mann_Walter_2007_Open_Access_Management_Report.pdf

Heuer, S., 2007: Mash-ups für Professoren. In: Technology review 07, 5 S. <http://www.heise.de/tr/artikel/92138>

Kunze, S., 2005: Zitierbarkeit: eine wichtige Anforderung an digitale Publikationen. In: cms-journal, H. 27, S. 51-54. <http://edoc.hu-berlin.de/cmsj/27/kunze-sabine-51/PDF/kunze.pdf>

Lossau, N., 2007: Der Begriff „Open Access“. In: Deutsche UNESCO Kommission: Open Access – Chancen und Herausforderungen: ein Handbuch. Bonn, S. 18-25

Mugabushaka, A.-M. / Fournier, J., 2005: Offener Zugang zu wissenschaftlichem Wissen: Erfahrungen und Ansichten DFG-geförderter Wissenschaftler; Ergebnisse einer empirischen Untersuchung des Publikations- und Rezeptionsverhaltens von DFG-geförderten Wissenschaftlern unter besonderer Berücksichtigung von Open Access. Deutsche Forschungsgemeinschaft: Bonn. DFG-Infobrief, Nr. 1/2005. http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/ib/download/ib01_2005.pdf

Rothenwänder, K.-P., 2005: Wo wirkt das Lösungsfokussierte Management? In: TRAiNiNG Nr. 1/Feb. 2005. <http://www.gfp.at/index.php?id=868>

Schirmbacher, P., 2007a: Open Access – ein historischer Abriss. In: Deutsche UNESCO Kommission: Open Access – Chancen und Herausforderungen: ein Handbuch. Bonn, S. 22-25

Schirmbacher, P., 2007b: Open Access – Informations- und Repository-Netzwerk in Deutschland: Vortrag auf dem BID-Kongress, Leipzig, 19.03.2007. <https://zope.cms.hu-berlin.de/schirmbacher/publ07/oa-bidkongress-070319/>

Schlippe, A. von / Schweitzer, J., 2000: Lehrbuch der systemischen Theorie und Beratung. 7. Aufl. Göttingen: Vandenhoeck & Ruprecht

Sietmann, R., 2006: Über die Ketten der Wissensgesellschaft: der Kulturkampf über den Zugang zu wissenschaftlichen Veröffentlichungen verschärft sich. In: c't, Nr. 12, S. 190 ff. <http://www.heise.de/ct/06/12/190/>

Sparrer, Insa, 2001: Wunder, Lösungen und System: lösungsfokussierte systemische Strukturaufstellung. Heidelberg: Auer

Voges, W., 2007: Open Access in den Naturwissenschaften. In: Deutsche UNESCO Kommission: Open Access – Chancen und Herausforderungen: ein Handbuch. Bonn, S. 76-78

Wittgenstein, L., 1921: Tractatus logico-philosophicus

Wissenschaftsindikatoren: Neue Aspekte und Entwicklungen

Bibliometrie zwischen Forschung und Dienstleistung

Wolfgang Glänzel, Koenraad Debackere

Zusammenfassung

Bibliometrie verbindet eigenständige Forschung und weitgefächerte Dienstleistungen. Beide Aktivitäten haben eine lange Tradition. Ursprüngliche Anwendungen waren neue Informationsdienste; die Forschungsevaluation wurde die zweite Wurzel der modernen Bibliometrie. Die Kenntnis dieses Hintergrunds ist notwendig, um die gegenwärtige Situation des Fachgebietes zu verstehen: Viele Methoden wurden für die wissenschaftliche Information entwickelt, werden aber auch in der Evaluierung angewendet. Trotzdem sind die meisten Indikatoren bei entsprechender Anwendung für die Erfüllung dieser Aufgabe geeignet.

Die Autoren geben einen Überblick über Standardindikatoren, deren Einsatz in der Forschungsevaluation und andere „Dimensionen“ in der Evaluation, wie die Rolle der Aggregationsebenen und der strukturbasierten Methoden. Schließlich werden drei neue Forschungsthemen und deren mögliche Anwendungen in evaluativen Studien besprochen.

Abstract

Bibliometrics connects research and services. Both areas of activity have a quite long tradition. Originally, bibliometrics was focussed on presenting new information services; research evaluation has become a second root of present-day bibliometrics. Knowing this background is important for understanding the present situation of the field: Many methods have been developed for science information purposes, but are applied in an evaluative context. Nonetheless, advanced indicators are fit for this task if they are properly applied.

An overview of advanced standard indicators, their fields of applications and of other important “dimensions” of evaluation such as the role of the aggregation level and structural analysis is given. Finally, three “hot topics” in bibliometric research and their possible application in evaluation studies are discussed.

1. Einleitung

Aus heutiger Sicht stellt sich die Bibliometrie als eine komplexe Disziplin dar, die eine Brücke zwischen spezifischer Forschung und weitgefächerten Dienstleistungen schlägt. Um ihre noch immer etwas kontroverse und mitunter einseitig interpretierte Rolle und Funktion der Bibliometrie begreifen zu können, sollen zunächst kurz Hintergrund und Wurzeln dieser Disziplin beleuchtet werden.

Zwei wichtige und charakteristische Begleiterscheinungen des Wandlungsprozesses der Wissenschaft zu einer zunehmend „Großen Wissenschaft“ (Weinberg, 1961, Price, 1963) sind die sich verändernden Kommunikationsprozesse mit gravierenden

Folgen für Möglichkeit und Praxis des Auffindens relevanter Information und die grundlegend veränderten Organisationsstrukturen der Wissenschaft mit neuen Strategien zur Verteilung der Forschungsmittel. Aus dieser Entwicklung ergab sich die Notwendigkeit, zum einen Methoden zu entwickeln, um die rasch wachsende Informationsflut zu bewältigen und zum anderen Instrumente zur Erhöhung der Effizienz der Verteilung von Fördermitteln zur Verfügung zu stellen. Es galt also, wissenschaftliche Kommunikation und Informationsströme zu quantifizieren und zu messen mit dem Ziel, Wissen zielgerichtet und effizient verfügbar zu machen sowie Methoden zu entwickeln, um Effizienz und Leistungsfähigkeit der Forschung messbar zu machen. Es bestand somit eine wissenschaftshistorische Notwendigkeit zum Entstehen einer Metrik der Wissenschaft. Historisch gesehen hat die Bibliometrie diese zwei Wurzeln, die ihr Profil und ihre Anwendungen maßgeblich bestimmen. Zunächst dominierte der kommunikationsbasierte Hintergrund. Der Grundgedanke früher Bibliometrie war, Wissenschaftlern beim Auffinden einschlägiger Literatur zu helfen, Forschern Übersicht über Trends und neueste Entwicklungen in ihrem Fachgebiet zu geben, Studenten bei der Wahl einer akademischen Laufbahn zu beraten und Bibliotheken Empfehlungen zum Erwerb relevanter Zeitschriften zu geben. Die ersten Anwendungen, oder wenn man will, die ersten Dienstleistungen hatten auch genau das zum Ziel, nämlich die Ausweitung und Verbesserung der Wissenschaftsinformation für Forscher und Anwender aber auch für solche Einrichtungen, die wiederum selbst informationsbasierte Dienste anboten. Die erste systematische Zitationsstudie von Gross und Gross (1927) wandte sich an kleine Hochschulbibliotheken zur Optimierung der Neuzugänge ihres Zeitschriftenbestands. Auch die Studie von Bradford (1934) über die Verteilung von Artikeln zu einem Thema über verschiedene Zeitschriften ist im Zusammenhang der Wissenschaftsinformation und Analyse von Kommunikationsmustern zu sehen. Gleichzeitig wird hier aber deutlich, dass Bibliometrie seit je den Anspruch hatte, mehr als Dienstleistung zu sein. Diese Studie und Alfred Lotkas (1926) Untersuchung über die Publikationsaktivität von Autoren gelten als erste bedeutende theoretische Ansätze in der Bibliometrie. Es ging beiden Autoren darum, Gesetzmäßigkeiten der dokumentierten wissenschaftlichen Kommunikation zu erfassen und mathematisch zu beschreiben. Auch spätere, komplexere Methoden wie die von Henry Small entwickelte Kozitationsanalyse (z.B. Small, 1973, Small & Garfield, 1985) stand noch ganz im Dienste der wissenschaftlichen Information. Eugene Garfield betrachtete den auf Kozitationen basierenden "Atlas of Science" als neue Form der Übersichtsliteratur (z.B. Garfield, 1981), die auch geeignet ist, um Studenten bei der Wahl einer Karriere in der Wissenschaft zu helfen. Gleichzeitig wurde auch die evaluative Kraft einzelner bibliometrischer Instrumente früh erkannt. Während der von Garfield und Sher (1963) entwickelte *Impact Factor* und deren Vorgänger (z.B. Brodman, 1944) dazu dienten, relevante Zeitschriften zu identifizieren und der IF dazu benutzt wurde, die Auswahl von Zeitschriften für den *Science Citation Index* zu erleichtern, erkannte Garfield bald den Nutzen seines Zeitschriftenindikators als Leistungsmaß für wissenschaftliche Fachzeitschriften. Damit war der Weg für die evaluative

Anwendung von bibliometrischen Indikatoren geebnet. Dieser Weg, der auch von wissenschaftssoziologischer Forschung methodisch geprägt wurde, führte schließlich bis zu den heute üblichen Anwendungen und Dienstleistungen; so hat die mertonische Interpretation der Zitation als Teil des Belohnungssystem in der Wissenschaft (Merton, 1973) zu teilweise extremen Auffassungen wie die durch Sevinc (2004) beschriebene Interpretation des Impact Factors als eine gemeinschaftliche Währung der wissenschaftlichen Qualität. Eine ähnliche Perspektivverschiebung erfuhr die *strukturelle Bibliometrie* (vgl. Glänzel & Debakkere, 2003): Die zunächst als Erweiterung der Review-Literatur konzipierte Kartographie der Wissenschaft wurde zunehmend in den Dienst der Forschungsevaluation und Wissenschaftspolitik gestellt (Noyons, 2001). Der Schwerpunkt der Bibliometrischen Anwendungen verlagerte sich also allmählich von der Unterstützung von Bibliotheken und Informationsdiensten in die Richtung von Benchmarking und Evaluation im Dienste der Wissenschaftspolitik und Forschungsorganisation (s. Glänzel, 2006a). Neben dem wissenschaftssoziologischen Forschungshintergrund war auch die oben beschriebene „zweite Wurzel“ der Bibliometrie und die damit verbundenen objektiven Erfordernisse entscheidende Triebkraft dieser Entwicklung, nämlich Wissenschaftspolitik und Forschungsorganisation über Leistungs- und Wettbewerbsfähigkeit von nationalen Wissenschaftssystemen, über regionale und institutionelle Forschungsleistung zu informieren und das aufwendige, von Experten ausgeführte Beurteilungssystem durch quantitative Kennziffern zu ergänzen. Wesentliche Impulse kamen hierzu in den 70iger Jahren mit der Wissenschaftsindikator-Bewegung aus den Vereinigten Staaten. Trotz der genannten Perspektiv-Verschiebung, die nicht ohne Folgen für die bibliometrische Forschung und die wissenschaftliche Gemeinschaft blieb (Glänzel, 2006a), ist der wissenschaftsinformationsbezogene Bereich noch stets ein wesentlicher Bestandteil bibliometrischer Forschung und Anwendung. Allerdings ist dieser Bereich wegen geringerer Publizität weniger spektakulär und wegen seines unmittelbaren Nutzens auch weniger kontrovers. Im folgenden wollen wir uns aber auf den zweiten Bereich, auf die evaluative Bibliometrie konzentrieren und die Standardinstrumente und neuesten Entwicklungen der gegenwärtigen bibliometrischen Forschung vor dem Hintergrund des oben gesagten abreißen.

2. Bibliometrische Indikatoren: Fit für die Zukunft?

2.1 Konvergenz der Indikatorsysteme

Bibliometrische Indikatoren, genauer gesagt Bibliometrische Wissenschaftsindikatoren sind Outputindikatoren. Sie setzen den Gebrauch wohldokumentierter und reproduzierbarer bibliographische Datenquellen voraus. Eigentlich kann fast jede Bibliographie als Grundlage einer bibliometrischen Untersuchung dienen, doch sind es die großen bibliographische Datenbanken, die sich aus den genannten Gründen als Standardquellen durchgesetzt haben (Glänzel & Debakkere, 2005). Obwohl dem Web bereits eine ergänzende Rolle als Input für

bibliometrische Studien zukommt, ist mangelnde Dokumentation, unklare Auswahlkriterien und unvollständige Literaturabdeckung das Web als mögliche Alternative noch nicht geeignet.

Aufgaben, Datenquellen und vorgegebene Kriterien bestimmen weitgehend Methodik und Anwendungsmöglichkeiten. So ist es auch in der Bibliometrie: Um dem Kriterium der Wissenschaftlichkeit gerecht zu werden, sollten auch bibliometrische Ergebnisse unter identischen Bedingungen reproduzierbar sein (Glänzel, 1996). Die nahezu unüberschaubare Vielfalt oft inkompatibler Indikatoren in den 90iger Jahren veranlasste die bibliometrische Gemeinschaft zur Organisation eines Workshops über bibliometrische Standards (Glänzel et al., 1996). Die Entwicklung und konsequente Anwendung solcher Standards bedeutet natürlich nicht, dass dieselben einmal erarbeiteten Methoden und Indikatoren von allen Bibliometrikern und Anwendern benutzt werden müssen; aber es bedeutet, dass mit ähnlichen Methoden gewonnene Ergebnisse kompatibel und vergleichbar sein müssen. Natürlich lässt sich in den letzten Jahren auch eine gewisse Konvergenz beobachten. Die europäischen Bibliometrieforschungszentren in Budapest (ISSRU), Leuven (SOOI) und Leiden (CWTS) haben in den letzten 25 Jahren eigene komplexe Indikatorsysteme entwickelt, die für die Anwendung auf allen Aggregationsebenen geeignet sind (vgl. Braun et al, 1985, Braun & Glänzel, 1990, Moed et al., 1995, van Raan et al., 2003, Glänzel & Thijs, 2004b). Ein kurzer Blick auf die wesentlichen Komponenten dieser Systeme zeigt, dass die meisten Indikatoren praktisch übereinstimmen (s. Tabelle 1). Dies belegt wiederum, dass gleiche Aufgaben und identische Datenquellen zur Entwicklung von ähnlichen Methoden führen.

Tabelle 1 Beispiel konformer bibliometrischer Indikatoren verschiedener Institute

CITATION-BASED INDICATOR	ISSRU/SOOI	CWTS
Number of papers	p	P
Number of citations received	c	C
Share of (un-)cited papers	f_0 [cited: $1-f_0$]	%Pnc
Share of author self-citations	S	% SELF CIT
Mean citation rate	MOCR [MOCR _X]	CPP [CPPe _X]
Journal-based relative citation rate	RCR [RCR _X]	CPP/JCSm
Subject-based relative citation rate	NMCR	CPP/FCSm
"Publication-strategy" indicator	NMCR/RCR	JCS/FCS

Allerdings machen diese Indikatorsysteme lediglich *eine* Dimension der evaluativen Bibliometrie aus. Es handelt sich hierbei meist um elementare statistische Funktionen wie z.B. Mittelwerte und relative Häufigkeiten. Statistische Tests zur Beurteilung der Signifikanz von Abweichungen werden leider relativ selten angewendet. Zwar sind die zugrundeliegenden Verteilungen diskret und meist zudem extrem schief, doch lässt sich kraft des zentralen Grenzwertsatzes die klassische mathematische Statistik auf die oben genannten Indikatoren anwenden (s. Schubert & Glänzel, 1983, Glänzel & Moed, 2002, van Raan, 2006a). Voraussetzung dafür ist, dass die zugrundeliegende Verteilung im Anziehungsbereich der Normalverteilung

ist, und dass die Publikationsmengen einen Umfang von etwa 30 Dokumenten oder mehr haben. Die Konsequenzen für die Anwendung auf verschiedene Aggregations-ebenen liegen damit auf der Hand. Näheres dazu wird im Abschnitt 2.2 besprochen.

Die obigen Erwägungen deuten auf einen echten Forschungshintergrund der Indikatorik hin. Deshalb an dieser Stelle einige kurze Bemerkungen zum Verhältnis von Forschung und Dienstleistung: Bibliometrische Forschung bedeutet keineswegs, Teile von Auftragsstudien auf Fachtagungen zu präsentieren oder in wissenschaftlichen Zeitschriften zu veröffentlichen, sondern bibliometrische Forschung heißt systematische „Grundlagenforschung“ zur Förderung des Fachgebiets und anwendungsbezogene Forschung zur methodischen Unterstützung der auszuführenden Evaluationen. Hierfür seien zwei Beispiele angeführt: die bibliometrisch-statistische Analyse der „späten Anerkennung“ (s. Glänzel et al., 2003, Glänzel & Garfield, 2004, van Raan, 2004) und die Untersuchung der Rolle der Autoren-Selbstzitierung in der dokumentierten wissenschaftlichen Kommunikation (z.B. Snyder & Bonzi, 1998, Aksnes, 2003, Glänzel et al. 2004, Glänzel & Thijs, 2004a,b, Thijs & Glänzel, 2006, Fowler & Aksnes, 2007). Allein die Zahl der methodischen Beiträge zum letztgenannten Thema zeigt die Notwendigkeit einer soliden Forschung, bevor Indikatoren standardmäßig in der Evaluation angewendet werden können.

Was muß ein Indikatorensystem nun leisten, um für die Anwendung in der Zukunft tauglich zu sein? Es muss zunächst den wichtigsten Kriterien der „Wissenschaftlichkeit“ genügen. Evaluative Bibliometrie sollte nicht allein auf Argumenten und tendenziös-willkürlicher Auswahl von unterstützendem Zahlenmaterial beruhen, sondern muss auf einem probaten System bibliometrischer Standardindikatoren fußen, das auch von anderen Anwendern als Referenz genutzt werden kann und Objektivität gewährleistet. Auch sollten Fragen der Reliabilität der Indikatoren und Validität der Methoden vollständig gelöst sein. Zudem müssen die Indikatoren durch mögliche Normierung sowohl Veränderungen in den verwendeten Datenbanken als auch „inflationäre Trends“ (Cronin, 2002, Persson et al., 2004) kompensieren können, um auch longitudinal-vergleichende Analysen zu ermöglichen. Die obigen Systeme (s. Tabelle 1) erfüllen prinzipiell diese Anforderungen, aber bedeuten wie erwähnt nur eine Dimension der evaluativen Bibliometrie. Andere wesentliche Aspekte sind deren Anwendbarkeit auf verschiedenen Aggregationsebenen und die methodische Bewältigung fachgebietsspezifischer Besonderheiten. Diese Fragen werden im folgenden kurz behandelt.

2.2 Die Rolle der Aggregationsniveaus

Neben der Verschiebung des Focus bibliometrischer Dienstleistungen von den Informationsdiensten hin zu evaluativen Anwendungen fand auch eine Verschiebung innerhalb der Forschungsevaluation statt. Stand in den 70iger und 80iger Jahren noch das Monitoring nationaler Forschungsleistung und deren internationaler

Vergleich im Vordergrund, so hat sich der Schwerpunkt auf tiefere Aggregationsebenen verlagert. Diese Verschiebung hat nicht nur einen inhaltlichen, sondern auch einen technischen Hintergrund. Die technischen Voraussetzungen für Verfügbarkeit, Speicherung und Aufbereitung von enormen Mengen bibliographischer Rohdaten haben sich in den 90iger Jahren sprunghaft verbessert und damit diesen Wechsel erst möglich gemacht.

Parallel zu dieser Aggregationsverschiebung kommt hinzu, dass an die Stelle der ursprünglich passiven Rolle der Bibliometrie, die sich lediglich in der Deskription von Zuständen und Trends ausdrückte, eine zunehmend aktivere Funktion getreten ist. Erklärtes Ziel war nun nicht nur die reine Beschreibung, sondern auch der Beitrag an Finanzierungs- und Förderungsmodellen. Diese neue Rolle und damit verbundene tatsächliche und potentielle Probleme wurden z.B. durch Butler (2004) für Australien und Debackere & Glänzel (2004) für Belgien beschrieben. Der Trend zur Disaggregation ist aber auf der Mesoebene noch nicht abgeschlossen; die Nachfrage nach Evaluationen von individuellen Wissenschaftlern auch mit bibliometrischen Mitteln ist in den letzten Jahren deutlich gestiegen. Bibliometriker hatten bereits in den 80iger Jahren klargestellt, dass Bibliometrie für evaluative Zwecke auf der Mikroebene aufgrund der kleinen zugrundeliegenden Publikationsmengen, der schlecht einzuschätzenden Fehler und dem Verlassen der individuellen Anonymität eigentlich nicht zu verantworten ist. Bibliometrie stößt auf der Ebene individueller Forscher sowohl bei den Wissenschaftlern als auch bei den Bibliometrikern auf Ablehnung. Damit sind die Grenzen klar gesteckt: die Forschungsgruppe sollte die kleinste evaluierbare Einheit bleiben.

Ein weiteres Ergebnis des Standard-Workshops 1995 war die Feststellung, dass eine jede Aggregationsebene ihre eigenen Standards und damit auch ihre eigene „Bibliometrie“ hat. Die unterschiedlichen Größenordnungen der „Stichprobenumfänge“ sind nur eine Ursache. Vor allem die Untersuchungen von Autor-Selbstzitierungen und wissenschaftlicher Mitautorschaft haben dies deutlich gemacht. Während Selbstzitierungen auf der Makroebene weniger problematisch sind, sollte auf der Mesoebene zumindest der Anteil der Selbstzitierungen als ergänzender Indikator berücksichtigt werden (Glänzel & This, 2004a,b, Thijs & Glänzel, 2006). Im Falle von Einzelautoren gibt es gleich mehrere Möglichkeiten, deren Selbstzitierungen zu berechnen (Glänzel et al., 2004).

Hinsichtlich der Mitautorschaft ist die Situation noch komplizierter. Auf der untersten Aggregationsebene ist die Koautorschaft als Maß der (intramuralen) wissenschaftlichen Zusammenarbeit fragwürdig (Laudel, 2002), auf der Mesoebene trotz geographischer Verzerrungen (Katz & Martin, 1997) akzeptabel und als Maß für internationale Kooperation bereits Standard. Aus den Erfahrungen resultiert also, daß für jede Aggregationsebene eine geeignete Zusammenstellung von Indikatoren gewählt werden sollte, die den Besonderheiten der Aufgabe gerecht wird, und dass die Aggregation bei der Interpretation der Ergebnisse berücksichtigt werden sollte.

2.3 Forschungsprofile und Strukturanalysen

Neben den Besonderheiten der Aggregationsebene spielen auch fachspezifische Eigenheiten eine bedeutende Rolle. Man vergleiche nur die Kommunikationsmuster in der Mathematik mit jenen in der Hochenergiephysik. Auf der einen Seite haben wir verhältnismäßig geringe Publikationsaktivität mit relativ geringen Zitationshäufigkeiten, niedrige Mitautorschaftstandards und einen hohen Anteil an Autor-Selbstzitierungen; auf der anderen Seite beobachten wir Hyperautorschaft, die die Frage nach individueller Mitautorschaft und Selbstzitierungen auf diesem Gebiet ad absurdum führt. Der Einfluss der unterschiedlichen Kommunikationspraktiken in verschiedenen Fachdisziplinen auf bibliometrische Indikatoren ist nicht zu überschätzen. Selbst die konsequente Anwendung von zeitschriften- und fachgebietsnormalisierten relativen Indikatoren (s. Tabelle 1) kann den Effekt nicht vollständig kompensieren. Als zweite Dimension der Indikatorik ergibt sich mithin die Notwendigkeit des Aufschlüsselns, des „Herunterbrechens“ in Fach- und Teilgebieten. Van Raan (1999) nennt dies „bibliometrische Spektroskopie“, Schubert und Braun (1986) entlehnten den Wirtschaftswissenschaften einen einfachen Balance-Indikator zur Bestimmung und Untersuchung von Fachgebietsprofilen.

In einem weiteren Schritt sollen Strukturanalysen auf der Grundlage von institutionellen Forschungsprofilen und institutioneller Profilklassifikation Methoden entwickeln helfen, um fachgebietsbedingte Verzerrungen zu vermeiden (Thijs & Glänzel, 2008). Das Ziel des Clusterings ist, Institute nach ihren Forschungsprofilen so zu gruppieren, dass Ähnliches mit Ähnlichem verglichen werden kann. Zwar bleiben große multidisziplinäre Einrichtungen immer noch ein Problem, aber die ansonsten oft benachteiligten spezialisierten Forschungseinrichtungen erhalten eine gerechtere Chance. Der dynamische Aspekt zeigt einen weiteren Vorteil dieser Methode: Ein möglicher national oder institutioneller Strukturwandel kann aufgrund von Cluster-Migration leicht entdeckt werden. So berichteten Leta et al. (2006) beispielsweise von einigen spektakulären Fällen von institutionellem Profilwandel und dessen Hintergründen in Brasilien.

2.4 Neueste Trends in der Indikatorenforschung

Die Entwicklung bibliometrischer Methoden und Evaluationsinstrumente ist damit jedoch nicht abgeschlossen; Pars pro toto sollen drei neue Entwicklungen die neuesten Trends in der Bibliometrie hier näher beschreiben. Die drei forschungsintensiven Themen machen erneut deutlich, daß Bibliometrie eine Disziplin zwischen Forschung und Dienstleistung ist und noch stets Aufgaben der Forschungsevaluation, Wissenschaftsinformation und Informetrie miteinander verknüpft.

Der *h-index* nach J.H. Hirsch (2005) wurde konzipiert, um die Forschungsleistung eines Wissenschaftlers mit Hilfe nur eines Indikators zu messen. Obwohl dieser Index aufgrund seines Konzepts eigentlich nicht in die Gruppe der zur Evaluation institutioneller oder nationaler Forschungsleistung geeigneten Indikatoren gehört, hat er seit seiner Veröffentlichung für Aufsehen unter den Bibliometrikern und

Fachwissenschaftlern gesorgt. Der wesentliche Grund dafür ist die leichte Berechenbarkeit auf der Basis individueller Zitationslisten. Allerdings gibt es keine wirkliche Möglichkeit zur Normalisierung und Standardisierung dieses Leistungsmaßes. Wie in der Einleitung und unter 2.1 festgestellt wurde, ist das jedoch ein wichtiges Kriterium für die Anwendung bibliometrischer Indikatoren. Gegenwärtig bemühen sich sowohl Mathematiker als auch Bibliometriker um das Verständnis seines mathematisch-statistischen Hintergrunds (Glänzel, 2006b, Egghe, 2006, Brurrel, 2006, Eghhe & Rousseau, 2006, Jin et al., 2007, Beirlant et al., 2007, Burrell, 2007) als auch um Interpretation und Anwendbarkeit des Indikators (Bornmann & Daniel, 2005, 2007, Glänzel, 2006c, van Raan, 2006b, Costas & Bordons, 2007). Diese Studie enthüllten zum Teil neue Eigenschaften dieses Indikators, die zu neuen Anwendungsbereichen und Erweiterungen führen. Von diesen sind vor allem die Anwendungsmöglichkeiten auf der Mesoebene interessant (Braun et al., 2005, Schubert & Glänzel, 2007, Glänzel, 2007c), da sie auch das Potential für einen Einsatz im internationalen Vergleich und bei der Evaluation wissenschaftlicher Einrichtungen haben. Die Erforschung der Einsatzmöglichkeiten des h-Index ist bei weitem noch nicht abgeschlossen. Deshalb sollte dieser Index vorerst nur unter Vorbehalt zur Anwendung kommen.

Ein unbestreitbarer Nachteil der weiter oben besprochenen Standardindikatoren ist deren begrenzte Fähigkeit, wichtige Eigenschaften bibliometrischer Verteilungen wiederzugeben. Eine Möglichkeit, diese Unzulänglichkeit auszugleichen, ist die Einteilung von Publikationsaktivitäts- und Zitierhäufigkeitsverteilungen in selbstregulierende Zonen. Die Zahl der Zonen kann frei gewählt werden; die Intervallgrenzen zwischen den Zonen und die Zahl der Elemente in den einzelnen Zonen dienen als Marker und Maßstäbe (*characteristic scores and scales*). Obwohl dieses Prinzip bereits van Glänzel & Schubert 1988 publiziert und bislang wegen des großen Rechenaufwands für die nötigen Iterationen selten angewendet wurde, erweist sich die Methode als äußerst stabil (Glänzel, 2007a). Langsam scheint auch sich das Prinzip durchzusetzen: ein ähnlicher Ansatz wurde unlängst von Zitt et al. (2007) vorgestellt.

Der dritte hier vorzustellende Trend bezieht sich auf die Fachgebietsbestimmung und -abgrenzung. Eigentlich steht diese Aufgabe vor jeglicher bibliometrischen Untersuchung, sie ist aber eines der meist kritisierten Teilprozesse der Bibliometrie. Neuere Ansätze versuchen Garfields Idee von der Ausweitung von Kernliteratur mit Hilfe von Zitationsverbindungen wieder aufzunehmen. Ausgehend von einer bereits validierten Ausgangsmenge kann durch vorgegebene Kriterien für Zitations-, Referenzverbindungen, Kozitationen oder bibliographische Verknüpfung mit Dokumenten der Ausgangsmenge, die Menge relevanter Publikationen schrittweise erweitert werden (Zitt & Bassecouard, 2006, Bassecouard et al. 2007, Glänzel et al., 2005). Diese Kriterien steuern gleichzeitig Abdeckung und Noise des Retrievals. Glänzel et al. (2006) nannten diese Methode der Fachgebietsbestimmung „bibliometriegestützte Datenbeschaffung“ (*Bibliometrics-Aided Data Retrieval*).

Neben der Anwendung in der Forschungsevaluation können die Ergebnisse auch traditionsgemäß für Wissenschaftsinformation und Informetrie genutzt werden.

3. Schlussfolgerungen

Bibliometriker haben in den 80iger und 90iger Jahren ein System von robusten Indikatoren entwickelt, die für den Einsatz auf der Makro- und Mesoebene, d.h. sowohl für Einsatz von Wissenschaftsindikatoren im internationalen Vergleich als auch in der Evaluierung wissenschaftlicher Einrichtungen geeignet sind. Selbst die Evaluation von Forschungsgruppen ist gegenwärtig bereits ein Standardprozedere. Eine Grenze sollte da gezogen werden, wo Zuverlässigkeit und Anonymität nicht mehr gewährleistet sind, nämlich bei der Beurteilung der Forschungsleistung von individuellen Wissenschaftlern.

Neue Forschungsansätze ebnen den Weg zur Überwindung bekannter Schwächen der Bibliometrie. Diese, die Einbeziehung strukturbasierter Methoden und die konsequente Verbindung von Forschung und Dienstleistung sorgen für die Zukunftstauglichkeit der Bibliometrie.

Literatur

AKSNES, D.W. (2003), A macro study of self-citation, *Scientometrics*, 56 (2), 235-246.

BASSECOULARD, E., LELU, A., ZITT, M. (2007), A modular sequence of retrieval procedures to delineate a scientific field: from vocabulary to citations and back. In: 11th International Conference on Scientometrics and informetrics (ISSI 2007). Madrid, Spain, 25-27 June 2007. 74-84.

BEIRLANT, J., GLÄNZEL, W., CARBONEZ A., LEEMANS, H. (2007), Scoring research output using statistical quantile plotting, *Journal of Informetrics*, 1 (3), 185-192

BORNMANN, L., DANIEL, H.D. (2005), Does the h-index for ranking of scientists really work? *Scientometrics*, 65 (3) : 391–392.

BORNMANN, L., DANIEL, H.D. (2007), Convergent validation of peer review decisions using the h index: Extent of and reasons for type I and type II errors, *Journal of Informetrics*, 1 (3), 204-213

BRAUN, T., GLÄNZEL, W., SCHUBERT, A. (1985), *Scientometric Indicators. A 32 country comparison of publication productivity and citation impact*. World Scientific, Singapore - Philadelphia (Hungarian Edition, Budapest, 1983)

BRAUN, T. GLÄNZEL, W. (1990), United Germany: The new scientific superpower? *Scientometrics*, 19, 513-521.

BRAUN, T., GLÄNZEL, W., SCHUBERT, A. (2005), A Hirsch-type index for journals. *The Scientist*, 19 (22) 8.

BRADFORD, S.C. (1934), Sources of information on specific subjects, *Engineering*, 137, 85-86.

BRODMAN, E. (1944), Choosing physiology journals, *Bulletin of the Medical Library Association*, 32, 479-83.

BURRELL, Q. L. (2006), Hirsch's h-index: a stochastic model. *Journal of Informetrics*, 1 (1) 16-25

BURRELL, Q. L. (2007), On the h-index, the size of the Hirsch core and Jin's A-index, *Journal of Informetrics*, 1 (2) 170-177.

BUTLER, L. (2004), What happens when funding is linked to publication counts? In: H.F.M. Moed, W. Glänzel, U. Schmoch (Eds), *Handbook of Quantitative science and Technology Research. The use of Publication and patent statistics in studies on S&T Systems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004, 389-405.

COSTAS, R., BORDONS, M. (2007), The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level, *Journal of Informetrics*, 1 (3), 193-203

CRONIN, B. (2001): Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? *JASIST*, 52, 558-569.

DEBACKERE, K., GLÄNZEL, W. (2004), Using a bibliometric approach to support research policy making: the case of the Flemish BOF-Key, *Scientometrics*, 59 (2), 253-276.

EGGHE, L., ROUSSEAU, R. (2006). An informetric model for the h-index. *Scientometrics*, 69 (1), 121-129.

EGGHE, L. (2006). Theory and practice of the g-index. *Scientometrics*, 69 (1), 131-152.

FOWLER, J.H., AKSNES, D.W. (2007), Does self-citation pay? *Scientometrics*, 72 (3), 427-437.

GARFIELD, E., SHER, I.H. (1963), New factors in the evaluation of scientific literature through citation indexing, *American Documentation*, 14, 199.

GARFIELD, E. (1981), Introducing the ISI Atlas of Science: Biochemistry and Molecular Biology, 1978/80, Current Contents, #42, October 19, 5-13.

GARFIELD, E. (1975), ISI's Atlas of Science may help students in choice of career in science, Current Contents, #29, July 21, 5-8.

GLÄNZEL, W., SCHUBERT, A. (1988), Characteristic Scores and Scales in assessing citation impact. J Inf Sci, 14, 123-127.

GLÄNZEL, W., KATZ, S., MOED, H.F. (1996), Proceedings of the Workshop on "Bibliometric Standards" Rosary College, River Forest, Illinois (USA), 11 June 1995, Scientometrics, 35 (2), 1996, 165-166.

GLÄNZEL, W. (1996), The need for standards in bibliometric research and technology, Scientometrics, 35 (2), 167_176.

GLÄNZEL, W., SCHLEMMER, B., THIJS, B. (2003), Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon, Scientometrics, 58 (3) 571-586.

GLÄNZEL, W., DEBACKERE, K. (2003), On the opportunities and limitations in using bibliometric indicators in a policy relevant context, In: R. Ball (Ed.), Bibliometric Analysis in Science and Research: Applications, Benefits and Limitations, Forschungszentrum Jülich, Germany, 225-236.

GLÄNZEL, W., GARFIELD, E. (2004), The myth of delayed recognition, The Scientist, 18 (11), 8-9

GLÄNZEL, W., MOED, H.F. (2002) Journal impact measures in bibliometric research, Scientometrics, 53 (2), 171-193.

GLÄNZEL, W., THIJS, B., SCHLEMMER, B. (2004), A bibliometric approach to the role of author self-citations in scientific communication, Scientometrics, 59 (1), 63-77.

GLÄNZEL, W., THIJS, B. (2004a), The influence of author self-citations on bibliometric macro indicators, Scientometrics, 59 (3), 281-310.

GLÄNZEL, W., THIJS, B. (2004b), Does co-authorship inflate the share of self-citations? Scientometrics, 61, (3), 395-404.

GLÄNZEL, W., VERBEEK, A., DU PLESSIS, M., VAN LOOY, B., MAGERMAN, T., THIJS, B., SCHLEMMER, B., DEBACKERE, K., VEUGELERS, R. (2005), Stem Cells

- Analysis of an emerging domain of scientific and technological endeavour, SOOS, KU Leuven.

Accessible via: http://www.steunpuntoos.be/rapportstamcellen_june2005.pdf

GLÄNZEL, W., DEBACKERE, K. (2005), Messen von wissenschaftlicher Kommunikation und Forschungsleistung. Möglichkeiten und Beschränkungen bibliometrischer Methoden. Mitteilungen der Österreichischen Bibliothekarinnen & Bibliothekare, 58 (2), 9-20.

GLÄNZEL, W., JANSSENS, F., SPEYBROECK, S., SCHUBERT, A., THIJS, B. (2006), Towards a Bibliometrics-Aided Data Retrieval for scientometric purposes. Poster presented at the "9th International Conference on Science & Technology Indicators", Leuven, Belgium, 7-9 September 2006.

GLÄNZEL, W. (2006a), The 'perspective shift' in bibliometrics and its consequences. Invited lecture delivered at the "I. International Conference on Multidisciplinary Informaion Sciences and Technologies" (InScit2006), Mérida, Spain, 25-28 October 2006.

GLÄNZEL, W. (2006b), On the h-index – A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, 67 (2) 315–321.

GLÄNZEL, W. (2006c), (On the opportunities and limitations of the h-index, in Chinese), *Science Focus*, 1 (1), 10-11.

GLÄNZEL, W. (2007a), Characteristic scores and scales. A bibliometric analysis of subject characteristics based on long-term citation observation, *Journal of Informetrics*, 1 (1), 92-102.

GLÄNZEL, W. (2007b), Some new applications of the h-index, *ISSI Newsletter*, 3 (2), 28-31.

HIRSCH, J. E. (2005), An index to quantify an individual's scientific research output, *Proceedings of the National Academy of Sciences of the United States of America*, 102 (46) 16569–16572.

GROSS, P.L.K., GROSS, E.M. (1927), College libraries and chemical education, *Science*, 66, 385 389.

JIN, B.H., LIANG, L.M., ROUSSEAU., R., EGGHE, L. (2007), The R- and AR-indices: Complementing the hindex. *Chinese Science Bulletin*, 52 (6) 855-863.

KATZ, J. S., MARTIN, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.

LAUDEL, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11, 3–15.

LETA, J., GLÄNZEL, W., THIJS, B. (2006), Science in Brazil. Part 2: Sectoral and institutional research profiles, *Scientometrics*, 67 (1), 87-105.

LOTKA, A.J., The frequency distribution of scientific productivity, *J. Washington Acad. Sci.*, 16, 1926, 317-323.

MERTON, R.K. (1973), The normative structure of science. In: R.K. Merton (Ed.), *The sociology of science: theoretical and empirical investigations*, The University of Chicago Press, Chicago.

MOED, H. F., DE BRUIN, R. E., VAN LEEUWEN, TH. N. (1995), New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications, *Scientometrics*, 33, 381-422.

NOYONS, E. (2001). Bibliometric mapping of science in a science policy context. *Scientometrics*, 50 (1), 83–98.

PERSSON, O., GLÄNZEL, W., DANELL R. (2004), Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies, *Scientometrics*, 60 (3), 421-432.

SCHUBERT, A., GLÄNZEL, W. (1983), Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics*, 5 (1), 59-74.

SCHUBERT, A., BRAUN, T. (1986), Relative indicators and relational charts for comparative-assessment of publication output and citation impact, *Scientometrics*, 9 (5-6), 281-291.

SCHUBERT, A., GLÄNZEL, W. (2007), A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics*, 1 (3), 185–192.

SMALL H. (1973), Co-citation in the scientific literature: a new measure of the relationship between two documents. *JASIS*, 24, 265-9.

SMALL H., GARFIELD E. (1985), The geography of science: Disciplinary and national mappings. *JASIS*, 11:147-159.

THIJS, B., GLÄNZEL, W. (2006), The influence of author self-citations on bibliometric meso-indicators. The case of European universities, *Scientometrics*, 66 (1), 71-80.

THIJS, B. GLÄNZEL, W. (2008), A structural analysis of publication profiles for the classification of European research institutes, *Scientometrics*, 74 (2), forthcoming.

VAN RAAN, A.F.J. (1999), Advanced bibliometric methods for the evaluation of universities, *Scientometrics*, 45 (3), 417-423.

VAN RAAN, A.F.J., VISSER, M.S., VAN LEEUWEN, TH.N., VAN WIJK, E. (2003), Bibliometric analysis of psychotherapy research: performance assessment and position in the journal landscape, *Psychotherapy Research*, 13 (4), 511–528.

VAN RAAN, A.F.J. (2004), Sleeping Beauties in science, *Scientometrics* 59 (3), 467-472.

VAN RAAN, A.F.J. (2006a), Statistical properties of bibliometric indicators: Research group indicator distributions and correlations, *JASIST*, 57 (3), 408-430.

VAN RAAN, A.F.J. (2006b), Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups, *Scientometrics*, 67 (3), 491-502.

PRICE, D. DE SOLLA, (1963), *Little Science, Big Science*, Columbia University Press, New York.

SEVINC, A. (2004), Manipulating impact factor: an unethical issue or an Editor's choice? *Swiss Medical Weekly*, 134, 410.

SNYDER, H., BONZI, S. (1998), Patterns of self-citation across disciplines. *Journal of Information Science*, 24, 431-435.

WEINBERG, A.M. (1961), Impact of Large-Scale Science on the United States, *Science*, 134, 161–165.

ZITT, M., BASSECOULARD, E. (2006), Delineating complex scientific fields by hybrid lexical-citation method: an application to nanoscience, *Information Processing & Management*, 42 (6), 1513-1531.

ZITT, M., BASSECOULARD, E., FILLIATREAU, G., RAMANANA-RAHARY S. (2007), Revisiting country and institution indicators from citation distributions: Profile performance measures. In: 11th International Conference on Scientometrics and Informetrics (ISSI 2007). Madrid, Spain, 25-27 June 2007. 797-802.

Merging information sources to obtain the impact factor of open access journals

Patrick Vanouplines, Ronald Beullens

Abstract

The Lund University Libraries do a great job in maintaining the Directory of Open Access Journals (DOAJ). Beside DOAJ, also JSTAGE and SciELO offer similar overviews. A combination of these three directories does not exist. Thomson Scientific has produced its Journal Citation Reports (JCR) for many years. The Journal Citation Reports give indications about the impact of individual journals. Scientists who want to get most impact from their research, and give most weight to their employers and institutions with their publications, try to select the most appropriate journal with the highest impact.

This paper reports about an ongoing study that combines information from the DOAJ, JSTAGE and SciELO with information from the Journal Citation Reports. The result is a list of open access journals with their impact factors. Such a list, if properly maintained, can serve as a guide to researchers who want their paper to be published in an open access journal with an impact that is as high as possible. Some interesting observations obtained during the study are discussed. The paper is related to the work by Marie E. McVeigh (Thomson Scientific, JCR & Bibliographic Policy Editorial Development Department).

Introduction

More than 230 organisations from all over the world have signed the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Scientists are encouraged, if not forced, to publish their research results in open access. An advantage of publishing in open access is better visibility and a higher rate of citation. But which journals are suitable for authors who wish to publish in open access with an impact that is as high as possible?

Congratulations ! You have just obtained a research grant. Problem: the contract stipulates that you have to publish your results in an open access journal. Solution: just find the right journal in one of the lists of open access journals that are maintained and published on the internet. There are three such lists that are wellknown: DOAJ, J-STAGE and SciELO. There is some overlap between these lists, but that does not pose a real problem.

Then the question arises: how do I find the most appropriate journal, with the highest impact factor? This assumes that there are open access journals that are considered in the Thomson Journal Citation Reports (JCR). Next question: where can I find this combined information? The answer to this last question is simple: nowhere! A few papers study exactly this problem: they contain lists of open access journals that appear in Thomson's Web of Science. But a simple table, with the impact factor (or

the immediacy index) next to the open access journal title is not readily available. This forms the subject of the investigation that we discuss here.

Data Collection

Besides the impact factors and the immediacy indexes obtained from the Thomson Journal Citation Reports, the data used in this study come from two main sources:

- directories of open access journals,
- the list of journals as used by McVeigh in a paper that investigated open access journals in the Thomson (ISI) citation databases (McVeigh, 2004).

In May and June 2006 we collected data about open access journals from three directories: the DOAJ, J-STAGE and SciELO. We compiled a single list that contains 2492 (de-duplicated) titles. More details are given in the following table (Tab. 1). Notice that by the end of July 2007, just over one year after our data collection, the DOAJ alone contained more than 2780 open access titles.

Table 1. Description of the data sources (directories).

<i>Directory</i>	<i>Data collection date</i>	<i>Number of open access journals</i>
DOAJ Directory of Open Access Journals http://www.doaj.org/	29 May 2006	2253
J-STAGE Japan Science and Technology Information Aggregator http://www.jstage.jst.go.jp/	1 June 2006	300
SciELO Scientific Electronic Library Online http://www.scielo.org/	1 June 2006	159
Number of unique (= de-duplicated) titles		2492

Lars Björnshauge (Director of Libraries at the Lund University, where the DOAJ originated and is being maintained) was so kind as to provide us with historical data on the number of titles in the DOAJ. Together with the data we found on the website of the DOAJ, we plot the evolution of the DOAJ, as shown in the following figure (Fig. 1).

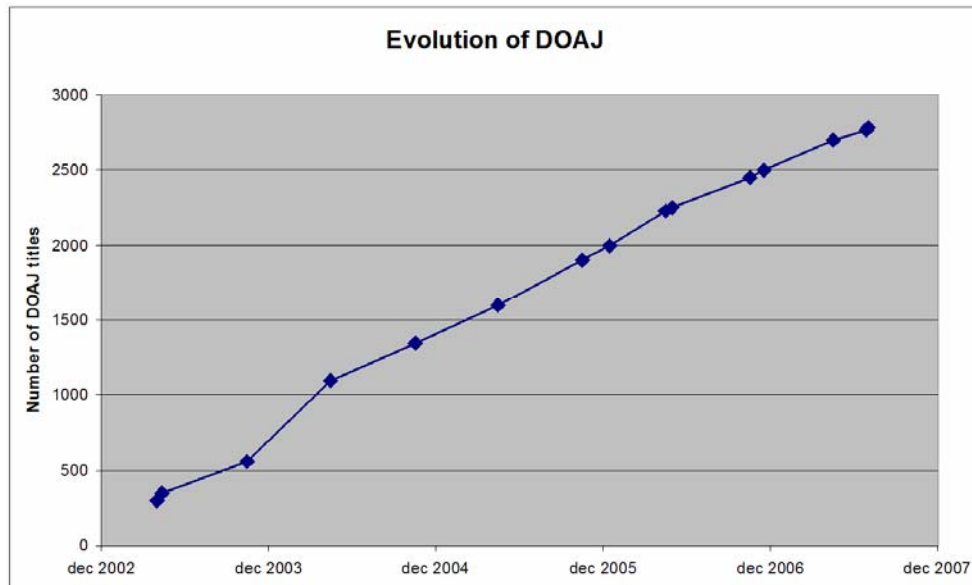


Figure 1. Evolution of the number of journal titles in the DOAJ.

In a personal communication, Lars Bjørnshauge pointed out that SciELO journals should be included in the DOAJ, provided that the journals comply with the DOAJ criteria. This was indeed experienced during the process of merging the three lists: in fact, no “new” journal title was added from SciELO.

Before we start processing the data we have to define the “open access journal” in the context of this paper. The DOAJ gives an answer to the frequently asked question “How do we define Open Access, Research Journal, quality control?” (DOAJ, 2007):

We define open access journals as journals that use a funding model that does not charge readers or their institutions for access. From the BOAI definition ... of “open access” we take the right of users to “read, download, copy, distribute, print, search, or link to the full texts of these articles” as mandatory for a journal to be included in the directory.

(BOAI = Budapest Open Access Initiative)

The selection criteria, based on DOAJ (2007), can be summarised as follows.

- Coverage
- All scientific and scholarly subjects are covered. Academic, government, commercial, non-profit private sources are all acceptable. The target group should be primarily researchers. A substantive part should consist of research papers. All languages are accepted.
- Access
- All content freely available, without delay (free user registration is acceptable).

- Quality
- The journal should exercise quality control on submitted papers through an editor, editorial board and/or a peer-review system.
- Journal registration
- The journal should have an ISSN.

Notice that the selection criterion “all content freely available” is very strict, and limits the number of journals to only the flavour of “*Unqualified open access*” in Willinsky’s nine flavours (Willinsky, 2003). Willinsky also describes the flavours:

- Dual mode (journals offered both in subscription-print and online in open access)
- Author fee (where authors pay to publish their paper)’
which are accepted as fully open access in the present paper.

In the present paper we do not accept as full open access the Willinsky flavours:

- Delayed OA (where articles become openly accessible some months after initial publication)
- Partial OA (with some articles in an issue are open access, others not).

The four other flavours (E-print archive, Per capita, Abstract and Co-op) are beyond the scope of the present study.

In fact, classification of open access could even be further specified and refined. Some open access journals offer their content in a form such that papers can be read online, and can be downloaded and saved, but cannot be printed. This form of open access is considered as acceptable in this study.

While performing the literature study for this paper we found the study by Marie McVeigh to be the only one relevant for the present investigation: McVeigh detects journals that are openly accessible and that are included in the Thomson Web of Science database. Indeed, we were also looking for “impact” indicators of journals that are open access. McVeigh studies open access journals regardless of subject area in her paper published in 2004 (but investigates differences regarding subject area), and provides a list of all selected journals in a seven-page appendix (McVeigh, 2004). This allowed us to use this material extensively in the present paper.

Data Processing

Before describing the processing of data we note that Thomson computes the JCR impact factor (I.F.) for year X as the number of cites in year X of articles published in years X-1 and X-2 divided by the number of articles published in years X-1 and X-2. The immediacy index (Im.In.) for year X is the number of cites in year X of articles published in year X divided by the number of articles published in year X. As an example we take the *Astronomical Journal* (Tab. 2).

Table 2. Example – basic data for the computation of impact factor and immediacy index for the Astronomical Journal.

Year X →	2006	2005	2004
Cites in 2006 to articles published in year X	743	2075	2643
Number of articles published in year X	480	449	523

For the journal data listed in the table (Tab. 2) we derive:

$$\text{I.F. (2006)} = (2075 + 2643) / (449 + 523) = 4.854;$$

$$\text{Im.In. (2006)} = 743 / 480 = 1.548.$$

This means that the 2006 edition of JCR (published in 2007) gives the impact factor (the number of articles in a journal and the number of citations to these articles) only if the journal was already considered in 2004 and 2005, but can already compute the immediacy index in the same year. The JCR data for year X normally become available in June of the year X+1.

Once the files with data on journals had been collected from DOAJ, SciELO and J-STAGE, these three files were merged (with de-duplication of the titles where necessary). Then journal title after journal title was meticulously manually checked in the JCR to see if data on the impact factor and the immediacy index exist for the considered journal. If the data about the impact factor and the immediacy index were found, they were copied into our file. This work was done twice: first for the Science Citation Index, and a second time for the Social Sciences Citation Index.

A comparison with the JCR Science Edition (2005) and the JCR Social Sciences Edition (2005) resulted in 273 titles present in the JCR Science Edition (2005) and 21 titles in the JCR Social Sciences Edition (2005). The combination of these two lists resulted in 291 unique titles.

The number of journals that appear on our list, or the McVeigh list (or on both lists) is 324. The number of unique journals on the McVeigh list is 237. That is one title less than the 238 listed in the Appendix A of McVeigh, due to one journal that is listed twice, but under another title description: “Nihon Reorogi Gakkaishi” and “Journal of the Society of Rheology Japan” are, respectively, the original Japanese title and the translated name. We did not retain 22 titles that appear in the McVeigh list, because they are not, or no longer, open access (according to our strict definition). There are 4 titles in the McVeigh list that did not appear in our list (see Tab. 3); these are journals that could (and in our opinion, should) be added to the DOAJ directory (note that 3 titles were already added to the DOAJ by mid-2007). We did not accept 7 of the McVeigh titles because they do not appear in the JCR (remember that McVeigh included titles in the Web of Science not necessarily in the JCR). Compared to the McVeigh list we added 88 titles. This is a very acceptable number: these are journals that are new open access journals, or journals that have moved to open access since the McVeigh study.

With the 291 unique titles that were obtained from the three open access journal directories, and the 4 titles (Tab. 3) that we added from the McVeigh list, we have now 295 titles in what we call the Final List, from which we can extract information.

Table 3. Journals that are open access following our definition, and that appear in the JCR, but that were not included in DOAJ in 2006.

Journal Title	ISSN	Included in DOAJ 2006	Included in DOAJ 2007
Annals of Saudi Medicine	0256-4947	No	No
Canadian Medical Association Journal	0820-3946	No	Yes
JASSS-The Journal of Artificial Societies and Social Simulation	1460-7425	No	Yes
Molecules, A Journal of Synthetic Organic Chemistry and Natural Product Chemistry	1420-3049	No	Yes

Information extraction

In the 295 journals in our Final List 32,884 articles were published in 2005, based on the JCR (version 2005). These articles received 486,989 citations. The impact factor ranges from zero (for some journals that were too recently added to the JCR to get an impact factor calculation) to almost 50 (49.794 for the journal “CA: A Cancer Journal for Clinicians”). The immediacy index ranges from zero to over 20 (21.300 again for the journal “CA: A Cancer Journal for Clinicians”); the journal with the second highest value has an immediacy factor of only 4.033.

The following table (Tab. 4) gives an idea of the number of journals in (rather) arbitrarily selected ranges for the immediacy index.

Table 4. Distribution of the number of open access journals in classes of impact factor and immediacy index.

Impact Factor	Number of journals	Immediacy Index	Number of journals
0.00 – 0.10	10	0.00 – 0.01	35
0.10 – 0.25	33	0.010 – 0.05	37
0.25 – 0.50	50	0.05 – 0.10	60
0.50 – 1.00	89	0.10 – 0.20	69
1.00 – 2.00	69	0.20 – 0.50	66
2.00 – 5.00	33	0.50 – 1.00	16
> 5.00	11	> 1.00	12

We grouped the 295 journals according to the same six broad subject areas that were used by McVeigh (2004): Chemistry; Physics, Engineering & Mathematics; Medicine; Life Sciences; Social Sciences; and Arts & Humanities. The following

figure gives the distribution over these six subject areas. The data for 2003 and 2004 were kindly provided by Marie McVeigh in a personal communication.

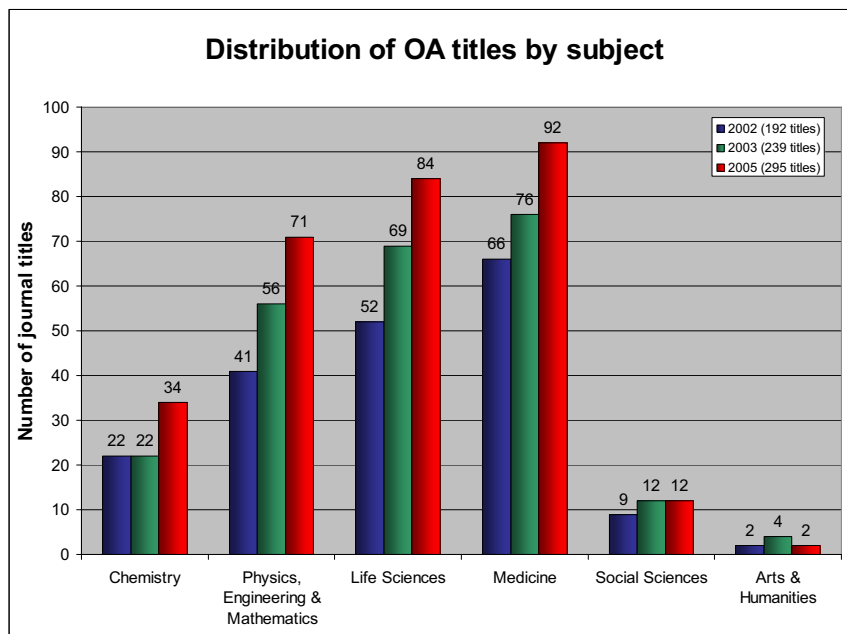


Figure 2. Evolution of the distribution of the selected 295 journal titles over subject areas.

A first observation regarding figure 2 is that the number of journal titles increases in subject areas for the science-technology-medicine (STM) group. Outside the STM group the number of titles is small and there is practically no increase, on the contrary. Regarding the subject group Arts & Humanities it should be noticed that no new titles appeared, and that the recent disappearance of two titles is due to our stricter definition of open access titles.

Besides these pure numbers it should be more important to have permanent access to the New List with 295 titles. Unfortunately, such a list, if not continuously revised, is already outdated the day after the collection of the basic data. But we had more in mind than just doing an academic exercise. The present study was rather done with a very practical target in mind. The aim was to demonstrate that:

- it is possible to collect basic data, and to construct a list with open access journal titles for which the impact factor is published,
- the result is of interest to researchers who are obliged to publish the results of their government-funded projects in an open access journal.

To make this idea operational we need to carry out the following work in the near future:

- from the academic point of view it is interesting to investigate the (new and disappearing) titles in the DOAJ and in the JCR;
- from the scholar and author's point of view a new service is needed.

The first item is discussed in the following while the entire next section is devoted to the second item.

When we embarked upon this study, more than a year ago, we were very happy that the JCR edition 2005 had just appeared: all recent data could be combined. During the course of the work, we found that all this is very time-consuming, and we saw ourselves spreading the work (in total probably no more than two months full-time) over a period of more than a year, due to circumstances. In the meantime, the JCR edition 2006 has appeared, and many journals have been added to the three directories. The DOAJ alone has grown from 2253 (end of May 2006) to over 2780 journal titles (end of July 2007).

What could have (and has) happened during these 14 months is:

- an increase of the number of titles in the open access directories;
- an increase of the number of journals considered in the JCR.

There was no time to check all previous 2253 journal titles in the new edition of JCR.

What we have done is:

1. check all the new DOAJ journal titles in the JCR (2006);
2. collect all the data about the (evolution of the) impact factor and the immediacy index of the new edition of the JCR.

Regarding point 1, our conclusion is that in the JCR 5 journal titles appear of the more than 500 additions to the DOAJ (all 5 of them in the Sciences Citation Index, and none in the Social Sciences Citation index).

Regarding point 2, we cannot present the results at the time of writing this paper, but we will discuss the evolution in the presentation of this work at the conference.

Future work

Only very recently, on 29 June 2007, the Lund University Libraries made available the Journal Info service, accessible at [<http://jinfo.lub.lu.se>]. This service intends to help the researcher in the selection of a journal for publication. Journal Info covers about 18,000 journal titles (by the end of July 2007). It is possible to search for an individual journal or to obtain information about a number of journals on a particular subject. It is clear that the authors of the present paper and the developers of the DOAJ are thinking along the same lines. The developers of the DOAJ are already producing results along these lines! We discuss this in the following.

The next figure (Fig. 3) gives a part (extract from the second page only) of the resulting output for a search for journal titles containing the word "information".

Merging information sources to obtain the impact factor of open access journals

Information Management Review ISSN: 87561557 Subject: General bibliographies --- Library and information science (General) --- Business and management --- Service management Publisher: Aspen Publishers
Information Manager: designed for organizational information needs ISSN: 01606123 Subject: Signals and systems --- General bibliographies --- Library and information science (General) Publisher: Information and Records Management
Information Polity: an international journal on the development, adoption, use and effects of information technology ISSN: 15701255 Subject: General bibliographies --- Library and information science (General) Publisher: IOS Press
Information Processing and Management Full info ISSN: 03064573 Subject: General bibliographies --- Library and information science (General) --- Business and management --- Planning and surveying Publisher: PERGAMON-ELSEVIER SCIENCE LTD
Information Processing Letters Full info ISSN: 00200190 Subject: Mathematics --- General bibliographies --- Library and information science (General) Publisher: ELSEVIER SCIENCE BV
Information Research: an international electronic journal ISSN: 13681613 Subject: General bibliographies --- Library and information science (General) Publisher: UNIV SHEFFIELD DEPT INFORMATION STUDIES
Information Retrieval Full info ISSN: 13864564 Subject: General bibliographies --- Library and information science (General) Publisher: KLUWER ACADEMIC PUBL
Information Sciences Full info ISSN: 00200255 Subject: General bibliographies --- Library and information science (General) Publisher: ELSEVIER SCIENCE INC
Information Sciences - Applications ISSN: 10690115 Subject: General bibliographies --- Library and information science (General) Publisher: Elsevier

Figure 3. Output screenshot (extract from the second page) given by the service Journal Info as a result of searching for journals with “information” in the title.

In the list of journals there is no indication that “Information Research: an international electronic journal” is indexed in the JCR. Clicking on the journal’s name leads to the following screenshot (Fig. 4):

The screenshot shows a web interface titled "Journal Info" with a blue header. Below the header, the journal's name "Information Research: an international electronic journal" is displayed. The page lists the following details: ISSN: 13681613, Publisher: UNIV SHEFFIELD DEPT INFORMATION STUDIES, Homepage: <http://informationr.net/ir/>, Subject: General bibliographies --- Library and information science (General), and First published year: 1976. Under the "Reader accessibility" section, it indicates "Open Access: Yes" with a green checkmark. The "Cost" section shows "Subscription price per article: FREE", "Subscription price per citation: FREE", "Profit-Status: Non-Profit", and "Publication Fees: No", all with green checkmarks. The "Quality" section shows "Databases indexing the journal: LISA --- Social Sciences Citations Index --- INSPEC" with a green checkmark, and "FRIDA score: Low" with a red X. At the bottom, there is a Creative Commons BY-NC-SA license logo and a "Close" button.

Journal Info

Information Research: an international electronic journal

ISSN: 13681613
Publisher: UNIV SHEFFIELD DEPT INFORMATION STUDIES
Homepage: <http://informationr.net/ir/>
Subject: General bibliographies --- Library and information science (General)
First published year: 1976

Reader accessibility
✓ Open Access: [Yes](#)

Cost
✓ Subscription price per article: FREE
✓ Subscription price per citation: FREE
✓ Profit-Status: [Non-Profit](#)
✓ Publication Fees: [No](#)

Quality
✓ Databases indexing the journal: LISA --- Social Sciences Citations Index --- INSPEC
✗ FRIDA score: [Low](#)

Provided by [Lund University Libraries](#), Head Office with support from the [National Library of Sweden](#).

[Close](#)

Figure 4. Particular information in the Lund database of Journal Info about the journal "Information Research: an international electronic journal".

We see (Fig. 4) that the journal is indexed in the Social Sciences Citation Index. Good, but not good enough for our purposes. At the time of writing there was no advanced search facility that allows the user to select only journals that are open access, nor to search for journals that are indexed in the Citation Indexes. Once this becomes possible in Journal Info, there is a practical answer to the main recommendation in this paper. Since all the necessary information seems to be present in the underlying database of Journal Info it is just a matter of time before a tool becomes available to the researcher who is "forced" to publish in an open access journal.

For some journal titles there is additional information, if available, under the hyperlinks "Journal eigenfactor", "Article Influence", and "FRIDA score". For the journal title "Information Sciences" (published by Elsevier Science Inc) we find:

Journal eigenfactor = 0.0057859
Article Influence = 0.31237
FRIDA score = low

The "journal eigenfactor" is an alternative to the impact factor; it is an estimate of the percentage of time that library users spend with that journal (see Eigenfactor, 2007). The "article influence" is the journal eigenfactor times 1000, divided by the number of articles in the journal over the considered period. The journal eigenfactor is relatively new, but might, to our opinion, become an indicator as important as the impact factor.

There is more good news regarding Journal Info. As shown (Fig. 4), this service also displays the fees for publishing in a given journal. There are no fees for publishing in, for example, "Information Research: an international electronic journal". Suppose that we want to publish the present paper in an open access journal. To find the list of potentially appropriate open access journals we need to be able to define an advanced search in Journal Info like:

Subject	= information science
Type of journal	= open access
Publication fees	= none
Sort on impact factor	= yes

Of course, with this search operation we expect the service Journal Info to come up with the journal "Information Research". It would maybe not show the journal "First Monday", because this journal:

- is not in the Journal Citation Reports,
- has (only) the subject descriptors "Computer Engineering; Signals and systems; Computer Science".

As pointed out by Willinsky (2003), there are flavours of open access. An indication of at least the most relevant flavours would be an interesting item to see in Journal Info. On the other hand, only very recently, Journal Info added information about alternative open access journals, as soon as the full information for a (non-open access) journal title is shown. This is almost an alternative solution to our recommendation.

The previous paragraphs describe a wish list, applied to the services of the Lund University Library services DOAJ and Journal Info. Such a description can be easily written for Thomson's JCR. In our opinion, it would be a big advantage to see indications of the open access flavours in the journals displayed by JCR. Searching directly in the JCR for open access journals in a given subject area would be equally desirable.

Recommendations

The main recommendation that we make is to provide a system that shows open access journals which are indexed in the Thomson Citation Indexes. This will, as explained above, help researchers who have to publish their research results in open access – as noted before, this is required by an increasing number of governmental research funding offices, especially in countries where the Berlin Declaration on Open Access has been signed. At this moment the researcher / author has no possibility of adequately identifying candidate journals. This could be a task for libraries in research institutes, but then huge amounts of work would be done again and again in a multitude of institutes. We see two players who can take up this task:

- the University Library of Lund, where the DOAJ and the service Journal Info are maintained,

- Thomson, which maintains the Journal Citation Reports.

A collaboration of these two players should be investigated. In the spirit of open access, the new service should, of course, be free of charge for the user.

A (possibly too) loose recommendation is that as much effort as possible should be made to include as many open access journals as possible that are of acceptable quality. A more realistic recommendation concerns the present number of open access journals in the subject areas of the Social Sciences, and, especially, Arts & Humanities: in two or three years there have been no new open access journals in these subject areas – this is more an alarm signal than a recommendation.

References

DOAJ (2007). FAQ. Online: [<http://www.doaj.org/doaj?func=loadTempl&templ=faq>]. Accessed: 27 July 2007.

Eigenfactor 2007. Online: [<http://www.eigenfactor.com/methods.htm>]. Accessed 6 August 2007.

McVeigh M (2004). Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns. A citation study from Thomson Scientific.

Online:

[<http://www.thomsonscientific.com/media/presentrep/essayspdf/openaccesscitations2.pdf>]. Accessed 30 July 2007.

Willinsky J (2003). The Nine Flavours of Open Access Scholarly Publishing. In: Journal of Postgraduate Medicine, 49, pp.. 263-267. Available online: [<http://www.jpgmonline.com/text.asp?2003/49/3/263/1146>]. Accessed 30 July 2007.

Bibliometrie als Teil eines Trenderkennungs-Systems in der Naturwissenschaft

Dirk Tunger

1. Einführung

"Der sprunghafte Anstieg des verfügbaren Wissens der Menschheit [...] sprengt jeden Maßstab. Die Zeitspanne zur Verdoppelung dieses Wissens schrumpfte von einstmals hundert auf nur noch sechs Jahre" (Warnecke, 1992, S. 92). Auf die Wissenschaft übertragen, formuliert Warnecke konkreter: "Verständlich wird diese explosionsartige Entwicklung, wenn man bedenkt, daß 90 Prozent aller jemals forschenden und lehrenden Wissenschaftler in der Gegenwart leben" (Warnecke, 1992, S. 92).

Die Anzahl wissenschaftlicher Publikationen in Zeitschriften, die Anzahl von Zitationen in diesen Zeitschriftenartikeln und die Anzahl angemeldeter Patente steigen kontinuierlich (Umstätter, 2004, S. 237). Das Angebot an verwertbaren Inhalten ist hoch, die Anwendung wird aber immer schwieriger: Das Problem ist die Erfassung aller Inhalte durch die potentiellen Nutzer.

Eine hierdurch bedingte mögliche Entwertung vorhandenen Wissens (vgl. hierzu Warnecke, 1992, S. 93), vor allem im technologischen Bereich, durch neues Wissen, thematisiert Vanini (Vanini, 1999, S. 16) und macht hierfür die Geschwindigkeit technologischer Entwicklung verantwortlich (Vanini, 1999, S. 35).

Die Vielfalt der Angebote ist zum einen so groß, dass die Übersicht sehr schnell verloren geht, und zum anderen ist der Zeitaufwand für die inhaltliche Erfassung aller Angebote zum großen Teil nicht zu bewältigen. Dies zeigt auch die hier dargestellte Entwicklung wissenschaftlicher Veröffentlichungen¹ in der Datenbank "Science Citation Index".

¹ Eine weitergehende Interpretation von Abbildung 1 findet sich in Abschnitt 3 dieser Veröffentlichung

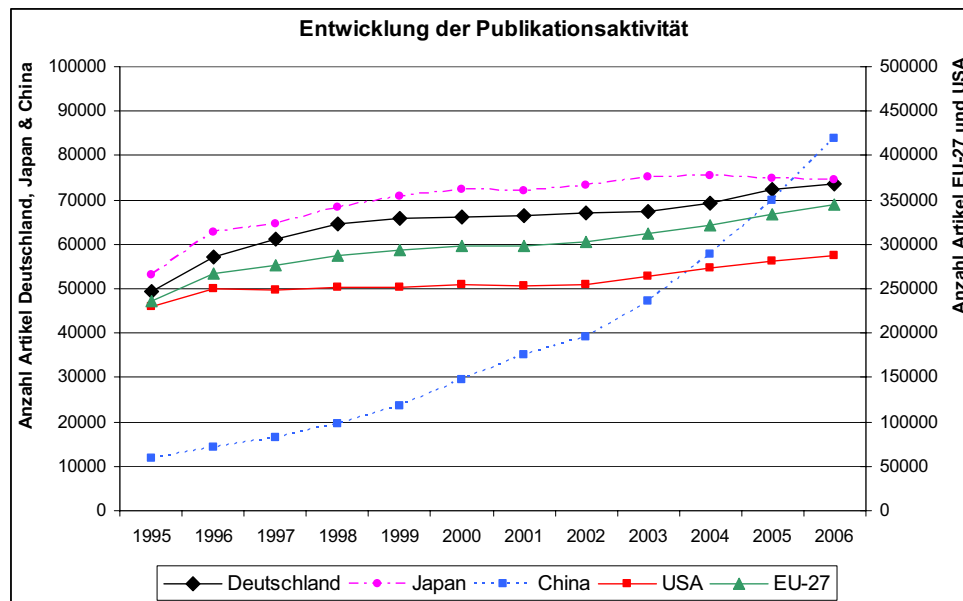


Abbildung 1: Anzahl wissenschaftlicher Veröffentlichungen (nur Research-Article) in der Datenbank "Science Citation Index" (SCI, SSCI, A&HCI)

Die Darstellung weist in eine ähnliche Richtung wie die Entwicklung der Anzahl wissenschaftlicher Zeitschriften in Abbildung 2 auf der Basis von de Solla Price (1963) und zeugt von der Zunahme wissenschaftlichen Outputs. Dies verdeutlicht die Notwendigkeit quantifizierbarer Aussagen zu Entwicklungen in Wissenschaft und Forschung (Schnurr, 2007) und hat den Einsatz von Bibliometrie² zu einem aktuellen Thema werden lassen, weniger in der von Nacke beschriebenen mathematischen Modellierung (Nacke, 1979, S. 46), sondern eher in einer von Ohly beschriebenen Form als technologisches Evaluationsinstrument (Ohly, 2004, S. 104).

Warum Bibliometrie heutzutage interessanter denn je ist, hat Derek de Solla Price bereits 1963 prognostiziert (De Solla Price, 1963): Er weist auf Wachstumsfunktionen in der Wissenschaft hin und verdeutlicht, wie stark die Anzahl wissenschaftlicher Zeitschriften zwischen 1655 und 1963 angestiegen ist und in welchem Maß sich die Entwicklung aus seiner Sicht fortsetzt (De Solla Price, 1963, S. 9).

² Der Begriff "Bibliometrie" wurde im Jahr 1969 durch A. Pritchard (Pritchard, 1969, S. 348f) verbreitet. Er leitet ihn vom vorher verwendeten Term "statistical bibliography" ab: "Therefore it is suggested that a better name for this subject [...] is bibliometrics, i.e. the application of mathematics and statistical methods to books and other media of communication" (Pritchard, 1969, S. 349). Eine frühere Nutzung des Begriffes Bibliometrie geht auf Paul Otlet im Jahr 1934 in dessen *Traité de Documentation* zurück (Umstätter, 2004, S. 237). Diodato verweist im "Dictionary of Bibliometrics" (Diodato, 1994) zur Definition des Begriffes "Bibliometrie" (Diodato, 1994, S. 13 – 15) auf zwei andere Quellen: "Bibliometrics is the application of various statistical analyses to study patterns of authorship, publication, and literature use" (Lancaster, 1977, S. 353). Oder: "Bibliometrics is the quantitative study of literatures as they are reflected in bibliographies" (White & McCain, 1989, S. 119).

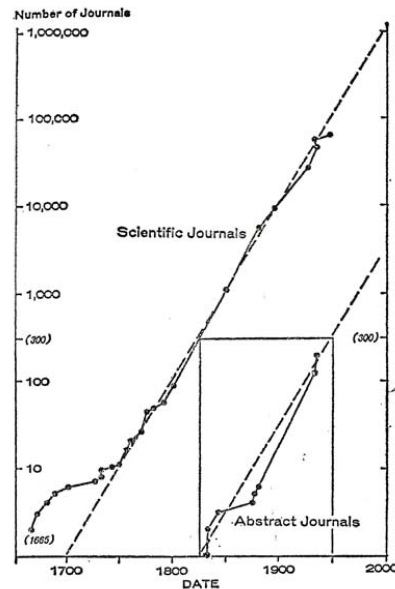


Abbildung 2: "Total number of Scientific Journals and Abstract Journals as a function of Date" (De Solla Price, 1963, S. 9)

Die Abbildung visualisiert die Anzahl der wissenschaftlichen Journals in Abhängigkeit der Zeit: "Numbers recorded here are for journals founded, rather than those surviving; for all periodicals containing any 'science' rather than for 'strictly scientific'" (De Solla Price, 1963, S. 9). Es geht de Solla Price in dieser Abbildung weniger um rein wissenschaftliche Zeitschriften und absolut zu betrachtende Zahlen wissenschaftlicher Zeitschriften als um die Visualisierung eines sich abzeichnenden und bis heute intakten Trends in der Wissenschaftskommunikation.

Mit dieser Darstellung hat er aus heutiger Sicht Recht behalten, das Wachstum der wissenschaftlichen Kommunikation hält auch weiterhin an, betrachtet man beispielsweise die Entwicklung von ausgewerteten Aufsätzen in der multidisziplinären Datenbank "Science Citation Index" (vgl. hierzu Abbildung 1).

"Die wissenschaftliche Produktion in Form von Publikationen wächst seit über dreihundert Jahren konstant mit ca. 3,5 % pro Jahr. Das entspricht einer Verdoppelungsrate von 20 Jahren" (Umstätter, 2004, S. 237).

Dies zeigt, dass Informationsspezialisten heute im Zentrum von gewaltigen digitalen Datenmengen stehen, die über Wissenschaft und ihren Output weltweit gesammelt werden. Sie sind als Informationsspezialisten grundsätzlich in der Lage, mit diesen Datenmengen umzugehen und aus ihnen fundierte, belastbare Informationen herauszukristallisieren (Ball, 2005; Kuhlen, Seeger & Strauch, 2004, S. 54ff).

Diese geschilderte Entwicklung wirft Fragen auf:

- Wie ist es möglich, den Überblick über wissenschaftliche Entwicklungen zu behalten, wenn der Output ständig ansteigt?

- Eignen sich Methoden der Bibliometrie für die Aufgabe, im Rahmen eines Trendbeobachtungssystems eingesetzt zu werden?

2. Einsatz von Bibliometrie / Ziel von Technologiebeobachtung

Mit einem zunehmenden Wachstum der Wissenschaft wird es für die Wissenschaftspolitik immer schwieriger, einen Überblick über aktuelle Entwicklungen zu behalten: Hierzu zählen u.a. die Entwicklung von Themen, die Reputation von wissenschaftlichen Einrichtungen und von Wissenschaftsregionen. Mit Hilfe von quantitativen Methoden der Bibliometrie wird der Versuch unternommen, einen objektivierten Einblick in wissenschaftliche Entwicklungen zu erhalten. Dieser Einblick soll fokussiert sein auf quantitative Parameter.

Ist ein Trend eine nur aus Zufällen bestehende Kette von Ereignissen, sind es Strategien oder lässt sich doch ein Muster erkennen (Rust, 2004, S. 8ff)? Trendforschung wurde von Igor Ansoff (Ansoff, 1976, S. 129ff) in der klassischen Wirtschaftswissenschaft eingeführt und im Englischen mit dem Begriff "Weak Signal Research" verknüpft. Ansoffs "Grundidee ist die Annahme, dass sich Diskontinuitäten im z.B. ökonomischen, technologischen, politischen oder sozialen Bereich noch vor ihrem tatsächlichen Eintreten bereits durch sogenannte 'weak signals' abzeichnen" (Loew, 2003, S. 29).

"Menschliches Handeln ist grundsätzlich zukunftsgerichtet und auf bestimmte Ziele orientiert. An die Stelle sicheren Wissens über die Zukunft treten Erwartungen der einzelnen Individuen. Diese beruhen auf Informationen prognostischer Art" (Rieser, 1980, S. 11).

Die Verringerung von Unsicherheit und die Erhöhung des Faktors Sicherheit, vor dem Hintergrund, eine "möglichst genaue Kenntnis über das Entstehen und die Wirkung zukünftiger Ereignisse zu erlangen", ist angestrebt. "Unabhängig von den Motiven liegt der wesentliche Nutzen dieses Anliegens in der Gewinnung von Zeit und Sicherheit. Zeit, um sich den eigenen Interessen gemäß optimal auf das Ereignis vorbereiten zu können, und Sicherheit, die Auswirkungen eigener Entscheidungen im Kontext des Ereignisses zu sehen" (Loew, 2003, S. 19).

Übertragen auf die naturwissenschaftliche Forschung heißt dies, Prospektionen für die Zukunft zu entwickeln. Zu diesen gehört es, u.a. Aussagen darüber zu treffen:

- Wie ein Thema sich unter zeitlichen Gesichtspunkten entwickelt
- Wie ein Thema sich unter regionalen / territorialen Gesichtspunkten entwickelt
- Welche Einrichtungen ein Thema führend bearbeiten / auf Veröffentlichungen zu einem Thema die größte Resonanz erhalten.

Um derartige Aussagen mit Hilfe bibliometrischer Methoden zu erzielen, sind geeignete Suchanfragen für die entsprechenden Datenbanken zu formulieren, die Daten zu prozessieren und auszuwerten. Hierbei handelt es sich um eine Umsetzung der groben Idee von Ansoff für die Wissenschaft. Sie stützt sich auf bereits vorhandene Erkenntnisse der Technikvorausschau zu quantitativen Analysen

(Steinmüller, 1997, S. 97-100; Stegelmann et al., 1988, S. 3f sowie Grupp & Schmoch, 1991, S. 1571-1615) und auf bereits vorhandene Erkenntnisse vorausgegangener Projektansätze, beispielsweise "Identifying Hot Spots" (Rothman, 1997, S. 15 - 17), "Use of Bibliometrics as a Technology watch technique" (Aguado-Monsonet, 1998) oder "Tracking and predicting growth areas" (Small, 2006, S. 595 – 610). Ebenfalls Anwendung findet die Beschreibung etablierter bibliometrischer Indikatoren (van Raan, 2004, S. 28 – 38).

Grupp & Schmoch nennen als "sinnvolle Suchkriterien nach Zukunftsgebieten" (Grupp & Schmoch, 1991, S. 1601) für quantitative Wissenschaftsanalysen:

- "Gebiete mit einem hohen Zuwachs in absoluter Menge
- Gebiete mit einem großen relativen Zuwachs
- Gebiete, die beide Kriterien gleichzeitig erfüllen"
- (Grupp & Schmoch, 1991, S. 1601)

3. Beispiele bibliometrischer Analysen vor dem Hintergrund von Trendbeobachtung

In diesem Abschnitt soll an drei Beispielen erläutert werden, in welcher Weise Bibliometrie als Trendbeobachtungstechnik zum Einsatz kommen kann. Die drei Beispiele stammen aus abgeschlossenen Projekten der Zentralbibliothek des Forschungszentrums Jülich³, die sich seit etwa vier Jahren intensiv mit diesem Thema befasst. Die Beispiele gliedern sich in zwei thematische Analysen (WissdeX Fullerene und materialwissenschaftliche Themen im SMART-Projekt) sowie eine formale (Publikationsoutput von Indien).

a) WissdeX Fullerene

Bei einem Thema können die zu beobachtenden Veränderungen in der Publikationsfrequenz, der Wahrnehmung des Themas, der bearbeitenden Länder oder Einrichtungen liegen. Auch denkbar sind inhaltliche Veränderungen, die u.a. durch eine Keywordanalyse ermittelt werden können. Der WissdeX Fullerene (Mittermaier et al., 2006, S.64f) gibt hierfür ein Beispiel: "Während zunächst die Grundlagenforschung der Fullerene im Vordergrund stand, rückte mit den Kohlenstoffröhrchen langsam eine Anwendung der Nanotechnologie und Nanoelektronik in greifbare Nähe" (Mittermaier et al., 2006, S.64).

Der Publikationsanteil zu den Grundlagen der Fullerene ist seit 1993 konstant geblieben, während die Zahl der Publikationen zu Kohlenstoffröhrchen seit diesem Zeitpunkt anstieg.

³ Siehe hierzu auch das entsprechende Informationsangebot im Internet unter: www.fz-juelich.de/zb/bibliometrie

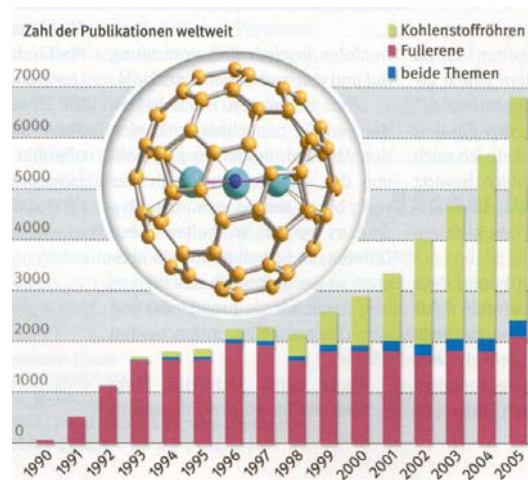


Abbildung 3: Entwicklung der Publikationsaktivität zu Fullerenen, unterschieden nach "Kohlenstoffröhren", Grundlagen der Fullerenen und einer Kombination aus beiden Themen (Mittermaier et al., 2006, S.64)

Der im Diagramm dargestellte Trend bezüglich der Grundlagen der Fullerenen wird weiter anhalten, solange es in diesem Bereich nicht zu fundamentalen neuen Entdeckungen kommen sollte. Gleiches gilt für den Trend der Kohlenstoffröhren: Auch dieser wird aus fachlicher Sicht noch anhalten, bis die wissenschaftliche Entwicklung ausgereizt ist.

b) Bibliometrie zur Unterstützung des SMART-Projektes

Es wird für die wissenschaftliche Planung immer wichtiger, implizites Wissen von Experten zu erhalten und mit objektivierte Ergebnissen anderer Erhebungstechniken zu vergleichen. Ein Beispiel hierfür ist das SMART-Projekt, das sich mit zukünftigen Entwicklungen in der Materialwissenschaft befasst (Schumacher et al., 2007, S. 75 – 77). Ganz bewusst wird ein Methodenmix eingesetzt, der wie folgt skizziert werden kann:

- Inhaltliche Auswertung von wissenschaftlicher Literatur (Review-Paper-Analysis)
- Darstellung von aktuellen wissenschaftlichen Forschungsschwerpunkten
- Ermittlung von Experten auf einem Gebiet
- Befragung der Experten mit Hilfe eines standardisierten Fragebogens
- Bibliometrische Analyse der Forschungsschwerpunkte
- Erstellen von Roadmaps für jeden einzelnen Forschungsschwerpunkt
- Organisation von Workshops
- Zusammenführung von subjektiven und objektiven Ergebnissen
- Publikation eines Gesamtergebnisses (Projektbericht)

Im SMART-Projekt, dessen primäres Ziel das Herausfiltern von Entwicklungspotential für die Zukunft in Teildisziplinen der Materialwissenschaft ist, werden die Methoden miteinander kombiniert, die am meisten zielführend und vielversprechend sind. Dabei ist bewusst gewollt, in immer neuen Evaluationsstufen die vorhergehenden Ergebnisse immer wieder zu überprüfen. Erst am Ende eines vielstufigen Prozesses steht ein Gesamtergebnis.

So lässt beispielsweise die Darstellung des Publikationsaufkommens zu 5 materialwissenschaftlichen Themen (Abbildung 4) Rückschlüsse auf die Veränderung der Bedeutung eines einzelnen Entwicklungsstranges in der Vergangenheit zu, woraus dann folgend Prospektionen für die Zukunft abgeleitet werden.

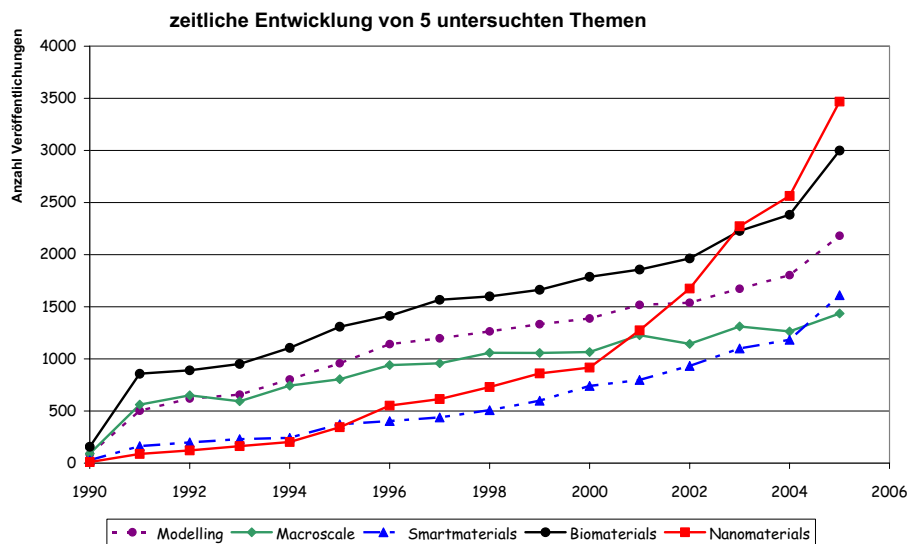


Abbildung 4: zeitliche Entwicklung von 5 untersuchten Themen des SMART-Projektes (Schumacher et al., 2007, S. 75 – 77)

c) Untersuchung des wissenschaftlichen Outputs eines Landes

Welche Aussagen können über den wissenschaftlichen Output eines Landes (in diesem Fall von Indien) getroffen werden? Dies ist die Leitfrage, die in der zu Grunde liegenden bibliometrischen Analyse (Mittermaier et al., 2007) im Mittelpunkt steht.

Kategorie	Indien: Anzahl Artikel 1996 - 2000	Indien: Anzahl Artikel 2001-2005	Veränderung
Biochemie	4030	5620	39%
Biologie	11106	14586	31%
Chemie	19397	27336	41%
Energie	1044	1090	4%
Geowissenschaften	5542	7495	35%
Informatik	1732	1967	14%
Ingenieurwissenschaften	11167	14040	26%
Landwirtschaft	7706	8403	9%
Materialwissenschaften	6229	9301	49%
Mathematik	2764	3146	14%
Medizin	16426	24287	48%
Physik	15293	19138	25%

Tabelle 1: Zuwachsraten in der wissenschaftlichen Publikationstätigkeit Indiens (Mittermaier et al., 2007, S. 10)

Es wird in Tabelle 1 beleuchtet, in welchem Maße sich die naturwissenschaftliche Publikationstätigkeit Indiens vom ersten Untersuchungszeitraum (1996-2000) zum zweiten Untersuchungszeitraum (2001-2005) verändert hat. Beide Zeiträume sind von identischer Länge. Um disziplinübergreifende Vergleiche zu vermeiden, wurde der naturwissenschaftliche Output in 11 Disziplinen eingeteilt, basierend auf den ISI-Subjectcategories. Neben dem Ergebnis in absoluten Zahlen ist eine prozentuale Veränderung ausgewiesen. "Alle Hauptdisziplinen konnten im Zeitraum 2001 – 2005 gegenüber den 5 Jahren zuvor (1996 – 2000) einen Zuwachs in der Anzahl der publizierten Artikel verbuchen. Die Materialwissenschaft und die Medizin sind in diesem Punkt federführend, aber auch die Biochemie hat aufgeholt. Fast schon als stagnierend kann die Entwicklung der Artikelproduktion für den Bereich Energie beschrieben werden. Im Vergleich zu Deutschland liegt Indien hier aber auch schon auf einem recht hohen Niveau, das bei 69 % der Artikelproduktion von Deutschland liegt. Ebenfalls einstellig ist der Zuwachs im Bereich der Landwirtschaft, hier gilt aber gleiches wie für die Energie-Forschung: hier hält Indien bereits ein recht hohes Niveau" (Mittermaier et al., 2007, S. 10).

Die Veränderung eines gemessenen Parameters, in diesem Fall die Artikelproduktion, lässt erkennen, in welche Richtung die Forschung sich bewegt. Im Originalreport folgt eine Unterteilung in die 170 Subjectkategorien des "Science Citation Index". Aus dieser Aufstellung ist auch der Rückgang einzelner Disziplinen erkennbar, z.B. der Tiermedizin oder der Zoologie.

Abbildung 1, die Darstellung der Publikationshäufigkeit von Research-Artikeln im "Science Citation Index" auf Ebene von Nationen, geht in eine ähnliche Richtung wie die Analyse über Indien. Es ist aus der Abbildung nicht nur ersichtlich, dass der

wissenschaftliche Publikationsoutput ständig ansteigt, sondern auch, wie einzelne Länder sich entwickeln. Es fällt auf, daß sich China über einen fast exponentiellen Anstieg seiner im Zitationsindex gelisteten Publikationen erfreuen kann, während der Anstieg bei den Ländern der EU-27 und den USA geringer ausfällt. Über die Wahrnehmung der chinesischen Artikel in Form der Zitation sagt dies aber nichts aus. Ebenso ist aus der Abbildung ersichtlich, dass die Publikationsaktivität der EU-27 derzeit höher ausfällt als bei den USA. Standen beide im Jahr 1995 bei einem ähnlichen Wert, so haben die EU-27-Staaten in den 17 Jahren des Beobachtungszeitraumes eine deutlich stärkere Steigerung erlebt. Diese Beobachtungen korrelieren mit Erhebungen des NSF in den USA (Mervis, 2007, S. 582) und von anderen Bibliometrie-Arbeitsgruppen (Glänzel, 2006, S. 59f).

Ausblick

An die Bibliometrie wird die Erwartung gerichtet, mit ihrer Hilfe einen Überblick über wissenschaftliche Tätigkeit zu erhalten. Aus diesem Grund wächst auch das Interesse an diesen Techniken (Hornbostel, 1997, S. 9) seitens des Wissenschaftscontrollings und Wissenschaftsmanagements. Diese Erwartung kann Bibliometrie erfüllen, wie anhand der ausgewählten Beispiele demonstriert wurde.

Bibliometrische Daten, die auf der statistischen Auswertung von wissenschaftlichen Veröffentlichungen beruhen, geben Hinweise für eine zukünftige Entwicklung, obwohl diese Daten die Gegenwart und die Vergangenheit repräsentieren. Es muss aber verdeutlicht werden, dass mit einem derartigen Trendanalysesystem keine Technik gemeint ist, die eine exakte Voraussage der Zukunft bezeichnet, sondern eine Technik, die eine grobe Vorstellung von Zukunft vermittelt. Die Beschäftigung mit der Zukunft bietet somit einen großen Vorteil: Man ist so auf eventuelle Entwicklungen vorbereitet, wenn man sich hiermit bereits im Vorfeld intensiver befasst hat.

Mit einer schnelleren Prozessierbarkeit wissenschaftlicher Daten wird auch der Umfang bibliometrischer Analysen zunehmen. Forschungslandkarten für größere Regionen werden erstellt. Wissenschaftliche Hotspots und Activity-areas werden länderübergreifend ermittelt und dargestellt.

Noch nicht möglich ist es, thematische wissenschaftliche Trends mit Hilfe einer alleinigen statistischen oder bibliometrischen Auswertung wissenschaftlicher Literatur ohne vorherigen qualitativen Input zu erkennen. Dies hat u.a. die Ursache, daß nur textbasierte Abfragen im "Science Citation Index" generiert werden können, die sich auf die vorgesehenen abfragbaren Felder (Titel, Abstract, Keywords und Adressangaben) und Kombinationen hieraus beziehen.

Trotz aller Objektivität wird aber auch in Zukunft gelten: "Was wir in der Zukunft wissen werden, können wir nicht wissen, denn sonst wüssten wir es ja" (Popper, 1974, S. XI).

Bibliometrie wird auch in Zukunft nur Anhaltspunkte einer Entwicklung liefern. Das gesamte System an Kommunikation ist so komplex und vielschichtig, dass wir es wahrscheinlich in Zukunft auch nicht in Echtzeit durchdringen und präzise Vorhersagen einer Entwicklung machen könnten.

Literatur

Aguado-Monsonet, M. A. (1998). Use of Bibliometrics as a Technology Watch technique. Application to the Analysis of the recent Developments of the Phtocatalysis. Seville, Paper prepared for the European Commission, EUR 18131 EN, Institute for Prospective Technological Studies.

Ansoff, I. H. (1976). "Managing Surprise and Discontinuity - strategic Response to Weak Signals (Die Bewältigung von Überraschungen - strategische Reaktionen auf schwache Signale)." Zeitschrift für betriebswirtschaftliche Forschung 28/1976: 129-152

Ball, R. and D. Tunger (2005). Bibliometrische Analysen - Daten, Fakten und Methoden. Grundwissen Bibliometrie für Wissenschaftler, Wissenschaftsmanager, Forschungseinrichtungen und Hochschulen. Jülich, Eigenverlag des Forschungszentrums Jülich.

De Solla Price, D. J. (1963). Little Science, Big Science. New York, Columbia University Press.

Diodato, V. (1994). Dictionary of bibliometrics. New York, Harworth Press.

Glänzel, W., K. Debackere and M. Meyer (2006). 'Triad' or 'Tetrad'? - On Global Changes in a dynamic world. 9th International Conference on Science & Technology Indicators, Leuven, 2006: 59-61

Grupp, H. and U. Schmoch (1991). Technologieindikatoren: Aussagekraft, Verwendungsmöglichkeiten, Erhebungsverfahren. Handbuch des Informationsmanagements im Unternehmen - Band II. H.-J. Bullinger. München, C.H. Beck'sche Verlagsbuchhandlung: 1571-1615.

Hornbostel, S. (1997). Wissenschaftsindikatoren – Bewertungen in der Wissenschaft. Opladen, Westdeutscher Verlag.

Kuhlen, R., T. Seeger and D. Strauch (2004). Grundlagen der praktischen Information und Dokumentation, Band 2: Glossar. München, Saur Verlag.

Lancaster, F. W. (1977). Measurement and evaluation of library services. Washington DC, Information Resources.

Loew, H.-C. (2003). Frühwarnung, Früherkennung, Frühaufklärung: Entwicklungsgeschichte und theoretische Grundlagen. Frühwarnsysteme. R. Schatz. Fribourg, InnovatioVerlag: 19-47.

Mervis, J. (2007). "U.S. Output Flattens, and NSF Wonders Why - Scientific Publishing." *Science* 317(5838): 582

Mittermaier, B., C. Plott, D. Tunger, U. Burkard and H. Lexis (2006). "WissdeX Fullerene." *Bild der Wissenschaft* 11/2006: 64-65

Mittermaier, B., D. Tunger, U. Burkard, S. Ramowsky and H. Lexis (2007). *Bibliometrische Analyse zum wissenschaftlichen Output von Indien*. Bonn, Internationales Büro des BMBF. Retrieved 15.5.2007, from http://www.internationale-kooperation.de/doc/Bibliometrischer_Report_Indien_dtsch_mit_Deckblatt_1971.pdf

Nacke, O. (1979). *Zitatenanalyse und verwandte Begriffe - Einführung in das Tagungsthema*. Deutscher Dokumentartag 1979. Deutsche Gesellschaft für Dokumentation (DGD). München, K.G. Saur Verlag: 45-50.

Ohly, H. P. (2004). *Bibliometrie in der Postmoderne. Wissensorganisation und Verantwortung: Gesellschaftliche, ökonomische und technische Aspekte; Proceedings der 9. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation*, Duisburg, 5. – 7. November 2004. H. P. Ohly, J. Sieglerschmidt and C. Swertz. Würzburg, Ergon Verlag: 103-114.

Popper, K. (1974). *Das Elend des Historizismus*. Tübingen, Verlag Mohr Siebeck.

Rieser, I. (1980). *Frühwarnsysteme für die Unternehmungspraxis*. München, V. Florentz Verlag.

Rothman, H. (1997). *Identifying 'Hot Spots' in Engineered Biocatalyst Research and Applications by Means of Bibliometrics*. Seville, Paper prepared for the European Commission, EUR 17342 EN, Institute for Prospective Technological Studies.

Rust, H. (2004). "Trends – Strategien oder Zufälle? - Badener Gespräche 2004." Retrieved 18.4.2007, from http://www.manstein.at/pictures/file_1084802655-bbe10db848b4897b172e63ffa6bfb70b.pdf.

Schnurr, E.-M. (2007). "Die Qual mit der Qualität." Retrieved 10.1.2007, from http://www.handelsblatt.com/news/Wissenschaft-Debatte/Geisteswissenschaften/_pv/_p/301117/_tft/_b/1202396/default.aspx/die-qual-mit-der-qualitaet.html.

Schumacher, G., D. Tunger, A. Smith, S. Preston and B. Knott (2007). "Materials Research in Europe: Mapping Excellence and Looking Ahead." *JOM: the journal of the Minerals, Metals & Materials Society* 59(2): 75-77

Small, H. (2006). "Tracking and predicting growth areas in science." *Scientometrics* 68(3): 595-610

Stegelmann, H. U., H. P. Peters, G. Stein and E. Münch (1988). Die Früherkennung technologischer Chancen - Realisierung und Perspektiven; Programmgruppe Technik und Gesellschaft, Reihe JÜL-Spez, Band 433. Jülich, Eigenverlag der Kernforschungsanlage Jülich GmbH.

Steinmüller, K. (1997). Grundlagen und Methoden der Zukunftsforschung; Szenarien, Delphi, Technikvorausschau; Werkstattbericht 21. Gelsenkirchen, Sekretariat für Zukunftsforschung.

Tunger, D. (2007). Bibliometrische Verfahren und Methoden als Beitrag zu Trendbeobachtung und -erkennung in den Naturwissenschaften (bisher unveröffentlicht). Regensburg / Jülich, Universität Regensburg / Forschungszentrum Jülich. Dissertationsschrift zur Erlangung des Doktorgrades der Philosophischen Fakultät.

Umstätter, W. (2004). Szientometrische Verfahren. Grundlagen der praktischen Information und Dokumentation, Band 1: Handbuch zur Einführung in die Informationswissenschaft und -praxis. R. Kuhlen, T. Seeger and D. Strauch. München, Saur Verlag: 237-243.

Van Raan, A. (2004). Measuring Science. Handbook of Quantitative Science and Technology Research. H. F. Moed, W. Glänzel and U. Schmoch. Dordrecht, Kluwer Academic Publishers: 19-50.

Vanini, S. (1999). Halbwertszeit von technologischem Wissen: Maßkonzepte und Implikation für die Technologieplanung. Hamburg, Dr. Kovac.

Warnecke, H.-J. (1992). Die Fraktale Fabrik - Revolution der Unternehmenskultur. Heidelberg, Springer Verlag Heidelberg.

White, H. D. and K. W. McCain (1989). "Bibliometrics." *Annual Review of Information Science and Technology* 24: 119-186

Teile dieses Aufsatzes sind entnommen aus: Tunger, 2007

Indicators for Footprints through science – Designing the Journal Application Level (JAL)

Miloš Jovanović

Abstract:

The question of how fundamental research and applicable technologies are connected has been discussed for a long time. One possible answer might be the "Footprint Analysis", through which the knowledge transfer between articles via citations is monitored. The "Footprint Analysis" will use a combination of bibliometric indicators, one of which will be the "Journal Application Level". This indicator is composed of journal data to allow a characterization of journals by telling whether they publish articles which are more application or science-oriented.

Die Frage wie Grundlagen- und angewandte Forschung miteinander verbunden sind wird schon seit langem diskutiert. Eine mögliche Antwort könnte die „Footprint Analyse“ liefern, welche den Wissenstransfer zwischen Artikeln über Zitationen beobachtet. Die „Footprint Analyse“ wird eine Kombination bibliometrischer Indikatoren nutzen, einer soll der „Journal Application Level“ sein. Dieser Indikator besteht aus Journaldaten, die eine Charakterisierung von Journalen erlauben sollen hinsichtlich der Frage, ob eine Zeitschrift eher Artikel aus der Grundlagen- oder angewandten Forschung veröffentlicht.

Einleitung:

Schon seit längerem gibt es Untersuchungen zu der Frage, wie angewandte- und Grundlagenforschung, oder auch Wissenschaft und Technologie, zusammenhängen und was für Wechselwirkungen zwischen den beiden bestehen (z. B. Bhattacharya S., 2003; Meyer M., 2000; Schmoch U., 1997). Indikatoren aus diesen Untersuchungen könnten genutzt werden, um eine Aussage zu treffen darüber wie sehr eine technische Disziplin in der Grundlagenforschung verhaftet ist, oder wann Technologien zur Anwendung gelangen. Hierbei werden nicht nur wissenschaftliche Artikel berücksichtigt, sondern auch Patente (Schmoch, U., 2006). Ergebnisse aus diesen Studien treffen auf Interesse bei Entscheidungsträgern, ähnlich wie bei anderen bibliometrischen Indikatoren. Der Vorteil liegt dabei auf der Hand: Entscheidungsträger, die Fördermittel zu vergeben haben, könnten mit Hilfe verlässlicher bibliometrischer Verfahren eine Entscheidung darüber fällen, ob sie mit ihren Mitteln Grundlagen- oder angewandte Forschung fördern möchten. Ein solches Verfahren wird derzeit am Fraunhofer INT im Rahmen einer Dissertation entwickelt, die sogenannte „Footprint Analyse“ (Jovanovic, M., 2007). Einen Teil dieses Verfahrens macht die Erhebung und Bestimmung des „Journal Application Level“ (JAL) aus. Die Arbeit am JAL beruht dabei auf einer Studie, die im letzten Jahr in Kooperation der Universitäten Düsseldorf und Graz entstanden ist (Schlögl, 2006).

Methodik

In der Studie von Schlögl (et al.) wurden Journals der Library and Information Science (LIS) daraufhin untersucht, ob man sie auf einer Skala in „Practitioner“- und „Academic“-Journals einteilen könnte. „Academic“ bedeutete hier, dass die Autoren und Leser an Universitäten forschten, „Practitioner“ forschten an nicht-staatlichen Institutionen. Die Studie fand mit Hilfe statistischer Methoden einige Indikatoren, aus denen sich folgende Hypothesen ableiten ließen:

- Das Vorhandensein von Werbung spricht dafür, dass es sich um ein Practitioner-Journal handelt
- Je größer das Editorial/Advisory Board ist, desto wissenschaftlicher ist die Zeitschrift
- Je länger die durchschnittliche Bibliographie der Autoren, desto wissenschaftlicher die Zeitschrift
- Je kürzer das Citing Half Life des Journals, desto praktischer

Der JAL soll nun aus den oben genannten Indikatoren zu einem einzigen Indikator zusammengefasst werden. Hierzu wurden die oben genannten Kennwerte für 60 Journale der Biologie, Chemie, Physik und Ingenieurwissenschaften erhoben, um in einem ersten Schritt festzustellen, ob die Hypothesen, die für die LIS-Journals gemacht wurden, auch auf andere Disziplinen übertragbar sind. Zusätzlich wurden die Journale noch nach eigenen Aussagen über die Leserschaft und die Zielgruppe untersucht. Mit statistischen Methoden sollen diese fünf Kennwerte nun gewichtet werden und in einen einzigen Indikator einfließen, den JAL. Die Auswahl der statistischen Methoden wird derzeit mit Mitarbeitern der Universität Wuppertal vorgenommen. Am Ende soll der JAL, so die Daten dafür sprechen, ein Indikator dafür sein, ob ein Journal eher ein „Academic“ oder ein „Practitioner“-Journal ist.

Diskussion

Generell besteht die Frage, ob die Hypothesen der Untersuchung mit den LIS-Journals auch auf andere Disziplinen übertragbar sind, da beispielsweise der Indikator „Werbung“ in der Biologie nicht nutzbar ist, um zwischen „Practitioner“- oder „Academic“-Journals zu unterscheiden. Dort gibt es viele eindeutig Grundlagenforschung veröffentlichende Zeitschriften, die jedoch auch Werbung im Blatt haben. Die statistische Auswertung wird zeigen, ob sich die fünf Kennwerte überhaupt nutzen lassen für einen einzelnen Indikator, oder ob nicht andere dafür erhoben werden sollten.

Literatur:

Bhattacharya, Sujit (et al.): „Characterizing intellectual spaces between science and technology“ in: „Scientometrics“; Vol. 58; Nr. 2; 2003; S. 369-390

Jovanovic, Milos: „Footprints through science – Using Citations to Assess the Path towards Applicability“ in: „ISSI Newsletter“; Vol. 3; Nr. 2; 2007; S. 16

Meyer, Martin: „Does science push technology? Patents citing scientific literature“ in: „Research Policy“; Nr. 29; 2000; S. 409-434

Schlögl, Christian; Stock, Wolfgang G.: „Practitioners and academics as authors and readers: The case of LIS Journals“; unveröffentlichtes Manuskript

Schmoch, Ulrich: „Indicators and the relations between science and technology“ in: „Scientometrics“; Vol. 38; Nr. 1; 1997; S. 103-116

Schmoch, Ulrich (et al.): „Typical development paths of knowledge-based technologies“ in: „Book of Abstracts; 9th International Conference on Science & Technology Indicators“; Leuven; 2006; S. 129-130

Wissenschaftsindikatoren als Management Tool

The ISI Web of Knowledge as a Management Tool

James Pringle

A Half-Century of Research Management

It is sometimes hard to believe that the idea of using citations to understand scholarship is now more than fifty years old. It was in fact in 1955 when Dr. Eugene Garfield first described the concept of a citation index for the sciences, and in 1964 that the Science Citation Index was launched. (Garfield 1955) Since that time, the ISI Citation Databases have grown to encompass nearly 40 million records and 700 million cited references, spanning more than a century of scholarly communication from 1900 to 2007.

The possibility of using the insights derived from citations, and especially the number of citations received by scholarly works, for evaluative purposes was soon recognized. The ability to count citations offered a way to avoid complete dependence on personal opinion, and promised to add a new layer of understanding to a simple count of the number of works that researchers could publish, or what has sometimes been termed the “least publishable units” of scholarship.

Early users began to recognize something important about citations, an attribute that has continued to fuel their power as evaluative indicators. Citations, to use terms from the language of today’s “Web 2.0” world, are the most fundamental units of “user-generated content” in the world of formal scholarship. In today’s world of web links, social tagging, and community-based ranking, they remain the best indicators that a scholarly article is playing a role in the continually growing network of scholarship. There is much debate over what exactly they mean—and an immense literature on the sociology of citation behavior. But, compared with newer alternatives like social tagging or article downloads, their characteristics are well understood.

To be well understood does not mean that they are not misused. There is a large literature dedicated to studying the misuse of citation data. Thomson Scientific, the company which now has custodianship of the ISI Citation Databases, advises care in using citation data to evaluate, and never to use it as a substitute for listening to views of experts knowledgeable in the field - peer reviewers.

Following Best Practices

A few years ago, an ISI analyst, David Pendlebury, listed ten important rules for citation-based evaluation, which I have summarized and updated below to help inform our understanding. (Pendlebury 2002) These rules are really just common sense, and represent an appeal that Thomson Scientific constantly makes to evaluators to approach their data with due consideration in order to gain value from its use and not be misled. These guidelines can be very helpful when planning or interpreting citation-based analysis.

Consider whether available data can address the question

There is a natural tendency to use the ISI Citation Databases to answer all questions about research impact for all levels and types of scholarship. Those who use our databases must appreciate the characteristics of the data and its best applications. For example, questions involving large-scale international comparisons in the life sciences, where journal literature is the primary means of scholarly communications, can be more effectively addressed than questions about the value the work of a researcher in German literature, where citations to articles may represent only a small fraction of the total impact of the work.

Choose publication types, field definitions, and years of data

Citation patterns vary according to the type of publication (review article, research article, letter to the editor, etc.) and by area of scholarly research. They accrue at different rates according to scholarly field. Those using citation data as management tool must recognize these distinctions.

Decide on whole or fractional counting

Increasingly, we work in an interconnected world of scholarly collaboration. Today's research articles may credit the work of numerous scholars from many locations around the world. The ISI citation databases include all authors and addresses for each of the documents it indexes, and most of Thomson Scientific's own analyses provide full credit to each author and institution for the resulting article, those who ask precise questions may find that a different way of allocating or weighting contributions may be appropriate.

Judge whether data require editing to remove "artifacts"

A variety of considerations must be taken into account when preparing citation data for analysis. For example, name homonyms must be disambiguated (Smith J and Suzuki T may actually represent two or two hundred people in the chosen dataset). Likewise, the various sub-units of an institutional name may be represented in a variety of ways, and the names of institutions may have changed over time and may not yet be adequately reflected in the data. While Thomson Scientific and others are developing ever-better algorithmic methods of unifying and normalizing this data, they will often require hand editing or checking by knowledgeable individuals to obtain best results.

Compare like with like

This rule is one of the most important, but often the most difficult to apply. Everyone is immediately impressed by a large number of citations, but such numbers often prove meaningless to the questions we are trying to answer. Comparing the total number of citations of a senior researcher who has produced many papers with that of a young and brilliant newcomer whose citations are yet to accrue is deceptive.

Likewise, comparing immunologists with economists is deceptive, because of different citation practices in these two fields.

Use relative measures, not just absolute counts

While total counts have their place, generally averages are better, and averages that enable comparisons to “baseline” indicators that take into account the citation patterns in fields, by year of publication, by journal, article type, and other relevant parameters are best. A minor but common error, for example, is to rely on the “impact factor” alone when judging a journal’s impact, rather than comparing the journal’s impact factor to others in the same field.

Obtain multiple measures

In general, more rather than fewer measures enables the evaluator to examine the question from multiple perspectives.

Recognize the skewed nature of citation data

This is one of the most fundamental and common errors in the use of citation measures for evaluation. A large proportion of the articles in any field, category, or journal are never cited, and a few receive a large number of citations. The distribution of citations is far from resembling a statistical “bell curve”. Because this distribution is almost universally present but often forgotten, it is often best to use carefully chosen simple measures appropriate to a specific question, rather than esoteric or complex ones, unless careful study calls for them. For example, taking the “impact factor”, which is a method of averaging a skewed distribution to compare journal impact, and creating additional concepts like average impact factor or total impact factor, and applying these thoughtlessly to evaluation of individuals, is a perilous, though frequent, activity of evaluators. I would not say that such measures should not be used, but only that careful thought should be given to whether they actually help answer the question that is being posed by the evaluator.

Confirm data collected are relevant to question

When looking at results, ask again whether the data have been appropriately chosen and treated, or whether artifacts or false assumptions have crept in.

Ask whether the results are reasonable

Use the quantitative tool of citation analysis to assist and inform qualitative discussion, not to replace it. If the numbers seem to be at odds with expert opinion, investigate these anomalies and ask further questions, do not simply use them as a substitute for informed opinion or peer review.

The New Interest in Evaluation

If these “rules” seem somewhat basic, I have included them because today’s academic climate of “hyper-evaluation” makes it ever easier to forget them. The last

decade has seen an increasing, and unprecedented, level of interest in the use of citation data for evaluation.

This interest is driven partly by the greater availability of this data through Internet-based tools such as the Web of Science™ and new citation resources such as the Astrophysics Database, CiteBase, and Google Scholar. More importantly, it is driven by the dynamics of the university community itself. Funding pressures become more intense each year. To take the example of NIH funding in the United States, only 19% of grant proposals are projected to receive funding in 2007, compared with 31% in 1998.(Zerhouni 2007)

At the same time, the global nature of today's scholarship means that new centers of excellence are constantly emerging in new places, older centers that have not successfully kept pace may lose out, and there are more opportunities than ever for the researchers and research teams with the greatest reputations for success. One commentator has pointed out that top researchers today are becoming celebrities to be sought after on a global scale by institutions seeking to strengthen their place in the global research hierarchy and garner more funding in a regionally and globally competitive arena.(Brody 2007)

Administrators seeking to understand these trends, to manage and plan their institutional strategies in the face of them, and to evaluate their current faculty, departments, and investments, naturally seek hard, outcomes-based evidence to help them. Indicators of numbers of publications and of the citations to those publications have a proven "track record" in providing this assistance, and so are turned to with ever greater frequency.

Though the hunger for outcomes-based metrics appears to be universal, the way in which citation metrics are implemented tends to reflect the structure of research administration by country and region, and the administrators' experience with such metrics. In the UK, national debate surrounds the implementation of a national program of metrics-based evaluation, the Research Assessment Exercise, in an environment where national funding of universities is the key determinant of their growth. In the United States, with its diverse mix of privately endowed and state-funded institutions, there is a more ad hoc approach and less interest in citation metrics for standardized evaluative activities. Those that exist tend to be trend analyses on a national scale, such as those conducted by the National Science Foundation (<http://www.nsf.gov/statistics/>). In some Asian countries, interest in metrics has often focused on straightforward reward systems for those researchers who publish in "high-impact" journals.

Multiplication of Metrics

Greater availability of data and the intense interest in outcomes at all levels have made today's world an exciting laboratory of new questions, approaches, and methodologies. Administrators in some countries are seeking to allocate resources at complex levels of a university funding hierarchy that may involve national government agencies providing funding to research centers that form part of individual or multiple

universities. These administrators pose new questions that stretch the limits of indexing systems. To take one example, the standardized address information captured in the ISI citation databases includes both main organization and sub-organization. The question of which entity is the “sub-org” and which is the “main org” is not always a simple one of asking the right administrator - it may reflect a complex situation in which several universities or agencies seek to claim credit for the outcomes of the same research!

New metrics are emerging almost daily. Although I have not made a statistically valid study, I would guess that a new alternative to the “journal impact factor” appears at least every month in the literature. These new metrics fall into several categories:

Correctives and additions to existing measures:

For example, using one set of measures - cited half-life - to correct or refine another - the journal impact factor - to create a new metric, the “cited half-life impact factor”.(Sombatsompop, Premkamolnetr et al. 2004) Such metrics refine our understanding of the thing we are currently measuring, without breaking new ground.

New types of citation-based metrics calculated in new ways:

For example, the “h-index”, which corrects the total citation count for a set of papers by looking for the point where the number of citations to a paper crosses the number of papers published. (Hirsch 2005) Another example is the Eigenfactor, which calculates a journal ranking based on the entire matrix of journal - journal citations provided by the Journal Citation Reports™, in an approach similar to that used by Google to construct the PageRank™ algorithm.(Bergstrom 2007) These metrics may enable us to measure new things, or refine our understanding of the things we are already measuring. The “h-index” has been proposed as a measure of individual research output, and the Eigenfactor as new approach to measuring journal impact.

New types of metrics based on data other than citations:

Examples of these metrics include a “usage factor” and other ways of counting downloads to derive a “reader-based” measure of journal and article impact.(Bollen and Van de Sompel 2006; Shepherd 2007) More broadly, institutional evaluation may bring together disparate additional data, including total value of grants received, number of degrees awarded, and patents received.

Web of Science Adapts to New Management Needs

In this exciting period of change in practices and interest in evaluation, it should not be surprising that the ISI citation databases are evolving, and that the focus of this evolution is the ISI Web of KnowledgeSM, which is Thomson Scientific’s premier Web platform. As this platform has evolved, TS has increasingly found the concept of “workflow” helpful in understanding how citation data are used for evaluation and how their use can be improved. We use this term in a very general way to refer to sets of tasks generally performed by specific types of analysts to achieve specific goals

using information. In this general way, a student just beginning her studies at a university is learning a workflow process as she is being trained in a specific set of steps to research and write a paper, and a scientist follows a number of workflow process not only in experimentation, but also in preparing a scientific paper for submission, or writing a grant proposal.

Seen in this light, there are many workflow processes that are of particular interest in the use of ISI citation databases as a management tool. Three such processes illustrate how we are trying to make sure that our citation data fits into them.

- A library management team is evaluating a journal collection
- A university department is evaluating a colleague for promotion.
- A university administrator wants to know how well a research team is performing in an international context.

As we have developed solutions for each of these different “workflows”, we have tried to integrate tools into the Web of Knowledge where appropriate, create separate products or modules where the task is more complex, use Web services or other techniques to help integrate citation data into external systems, or recognize that the analysis requires place large collections of citation data in other environments using a custom data feed.

Evaluating a journal collection:

Thomson Scientific has for many years produced the Journal Citation Reports, and this tool continues to evolve in line with the changing needs of collection managers. The much-discussed “journal impact factor” remains a key statistic, complemented by the immediacy index, cited and citing journal half-life, total citations. Collection managers and publishers often ask for specific additional data to help them with specific tasks, and we have responded by adding new metrics to this resource. Examples include charts showing the citation pattern over time along with percent self-citations and cited half-life; a five-year impact factor trend graph, and a journal relatedness factor that helps managers see other journals that are closely related by citation to the journal they are analyzing.

As we delved more deeply into the workflows of collection managers, we recognized that they need to analyze more indicators of a journal’s importance to their institution. They seek measures that reflect the actual publishing activity. These administrators want to know where their own researchers are publishing, what journals are citing their work, and what they are actually downloading, and to have this data presented in management reports that enable them to make decisions easily.

For this reason, we created a new module in the Web of Knowledge, which we call the Journal Use Reports®. This module integrates COUNTER-compliant usage data, citation data, and publication data about the research activities in an institution to show how journals are actually used within an institution. Its focus is on the way decisions are actually made by the library administration and the tools and data that support that decision.

Promotion and Tenure

When a university is evaluating individuals for promotion or other rewards, citation data often plays a part. This is one of the most delicate tasks that academic administrators can undertake, and an area where Thomson Scientific constantly recommends that citation data be used only to inform the discussion, not as a definitive tool.

In this area, it has become apparent to us that virtually every academic institution around the world has become engaged in a time-consuming effort to collect publication records that include citation data and to integrate these into evaluation programs. The ISI citation databases, which offer consistently indexed citation content that are comparative on a global basis, are quite rightly essential tools in this process. To help with these various tasks, we have added several functions to the Web of Science, including:

- Author Finder: To help locate a researcher of interest.
- The Distinct Author Identification System (DAIS): Which algorithmically disambiguates author names and associates publications with the appropriate research. This is a finding aid rather than a definitive list, since author participation is ultimately needed to determine if the associated works are truly authored by the same author.
- Citation Report: Which provides a concise report on a set of papers, including total citations, average cites per paper, and such article-level statistics as an “h-index”.
- ResearchID.com: Thomson Scientific is working with a number of institutions to beta test and prepare the general release of a new environment in which researchers can maintain their own publication lists, drawing on the bibliographic and citation data in the Web of Science.

Evaluating Research Teams

Given the intensity of the global challenges facing research administrators and the availability of consistently indexed citation data, it is no wonder that they turn to the ISI citation databases for such tasks as analyzing the contributions of a research team. To assist them, Thomson Scientific has extended the tools it offers within the Web of knowledge, but also recognized the need for careful custom research.

Initial Views and Exploration

- Analyze Tool: The Web of Science provides a flexible analysis tool with bar charts and simple statistics for on-the-fly comparison of outputs in any subject area defined by keyword searching. It provides a straightforward method to identify other centers of excellence related to the one under study.
- Citation Report: Useful for analyzing both individuals and teams, this report provides a visual look at the total publications and citation impact of the team.

Profiling Within Institutions

- Journal Use Reports: One of the features of the JUR is its ability to support administrator-constructed profiles of an institution or a research team within the institution. Doing so enables an administrator to create precise comparisons with similar institutions according to publications, citations, journals the team is citing, and journals they are using.

A Global View

- Essential Science Indicators: Has for many years been a high-level guide to global research excellence, enabling administrators to compare institutional research across scientific fields. This type of comparative analysis requires the use of baseline data, such as expected citation rates, which are

Detailed and Precise Analysis

- Custom Analysis: Because the kinds of questions administrators need to know can be so varied, it is likely that custom data sets will be needed. Some institutions have the ability to make good use of limited data sets from the Web of Science for their internal analyses. More commonly, administrators will want Thomson Scientific or a third party with access to the ISI citation databases to provide custom citation data processed and organized to their needs. This approach helps with resolving some of the difficulties in working with such data - such as the need to unify institutional addresses. Institutional Indicators provide baseline indicators, and Institutional Citation Reports provide metrics based on custom-selected data processed to exact specifications.

References

Bergstrom, C. (2007). from <http://www.eigenfactor.org/methods.htm>.

Bollen, J. and H. Van de Sompel (2006). Usage Impact Factor: the effects of sample characteristics on the usage-based impact metrics.

Brody, W. R. (2007). "College Goes Global." *Foreign Affairs* 86(2): 122-133.

Garfield, E. (1955). "Citation Indexes for Science - New Dimension in Documentation through Association of Ideas." *Science* 122(3159): 108-111.

Hirsch, J. E. (2005). "An index to quantify an individual's scientific research output." *PNAS* 102(46): 16569-16572.

Pendlebury, D. (2002). *The ISI Database and Bibliometrics: Uses & Abuses in Evaluating Research*. Thomson ISI Symposium. Tokyo.

Shepherd, P. T. (2007, September 10, 2007). "Final Report on the Investigation into the Feasibility of Developing and Implementing Journal Usage Factors." from <http://www.uksg.org/usagefactors/final>.

Sombatsompop, N., N. Premkamolnetr, et al. (2004). "A modified method for calculating the impact factors of journals in ISI Journal Citation Reports: Polymer Science Category in 1997-2001." *Scientometrics* 60: 217-235.

Zerhouni, E. (2007). "NIH at the Crossroads." from <http://grants1.nih.gov/grants/award/award.htm>.

Einsatz bibliometrischer Analysen im EU-Projekt zur Technologiefrüherkennung SMART

Show-Ling Lee-Müller, Gerd Schumacher

Zusammenfassung

SMART (Foresight Action for Knowledge-Based Multifunctional Materials Technology) ist eine Specific Support Action im 6. EU-Forschungsrahmenprogramm zur Früherkennung von wissensbasierten multifunktionalen Werkstofftechnologien.

Ziel des SMART-Projektes ist es, einen europäischen Kompetenzatlas in Materialwissenschaften zu erstellen und Werkstoffforschungsthemen von hoher Relevanz für die nächsten zwei Jahrzehnte zu identifizieren. Neben der Durchsicht von Übersichtsartikeln, Expertenbefragungen sowie Expertenrunden kam dabei auch die bibliometrische Analyse zum Einsatz.

Mit Hilfe bibliometrischer Methoden wurde auf dem Gebiet der Materialwissenschaften die Stellung Europas im internationalen Vergleich ermittelt („benchmarking“). Dabei wurde die Publikationsaktivität und der Wirkungsgrad der Veröffentlichungen betrachtet. Die Analysen erfolgten für fünf verschiedene thematische Bereiche: Nanomaterialien, intelligente Werkstoffe, Biomaterialien, maßgeschneiderte Werkstoffe für Hochleistungsanwendungen sowie Modellierung. Diese Untersuchungen wurden auch auf Städteebene durchgeführt und resultierten in entsprechenden Rankinglisten.

Mittels der Bibliometrie konnte eine sehr gute Übersicht der Materialforschungsaktivitäten gewonnen werden. Da bibliometrische Analysen sich mehr auf die Statistik als auf den wirklichen Inhalt der Veröffentlichungen konzentrieren, trugen die parallel durchgeführten Befragungen von über 200 Experten zur Abrundung des Gesamtbildes wesentlich bei.

Abstract

SMART (Foresight Action for Knowledge-Based Multifunctional Materials Technology) is a Specific Support Action financed within the Sixth Framework Programme which is aimed at the early identification of knowledge-based multifunctional materials technology.

The aim of the SMART project is to create a European map of excellence in materials science and to identify most relevant materials research topics which will be of importance in the next two decades. In addition to the data screening of review articles, expert interviewing and expert discussions, bibliometric analyses were also employed in SMART.

With the aid of these bibliometric methods, the position of Europe in materials science was benchmarked on a global scale, whereby both publication activities and the impact of the research paper were used as indicators. The analysis was carried out for the five materials research categories: Nanomaterials, Smart Materials, Bio-

conceptional materials, Tailored Macroscale Materials for high-performance applications and Modelling. These investigations were also carried out on a city level and resulted in corresponding ranking lists.

By using bibliometrics, a very good overview of materials research activities was obtained. Since bibliometric analysis concentrate more on the statistics rather than the actual content of the publications, the parallel interviewing of more than 200 experts was essential to complete the whole picture.

Einführung

SMART ist eine Specific Support Action im 6. EU-Forschungsrahmenprogramm zur Früherkennung von wissensbasierten multifunktionalen Werkstofftechnologien. Ziel des SMART Projektes ist, der wissenschaftlichen Fachwelt und der Europäischen Kommission Auskunft über spezifische Stärken und Schwächen in der Europäischen Materialtechnologie zu geben sowie ein Bild der zukünftigen Materialwissenschaften zu zeichnen. Dazu wurde einerseits ein europäischer Kompetenzatlas in den Materialwissenschaften erstellt sowie andererseits Werkstoffforschungsthemen von hoher Relevanz für die nächsten zwei Jahrzehnte identifiziert. Die SMART-Strategie besteht aus einer zweifachen Vorgehensweise:

- traditionelle Forschungsprognose (forecast), welche die Entwicklung aktueller Forschungsergebnisse extrapoliert
- innovative Technologiefrüherkennung (foresight), welche die Megatrends berücksichtigt und verschiedene mögliche Szenarien durchspielt

Insgesamt kann der SMART-Arbeitsvorgang wie folgt beschrieben werden (Abb.1).

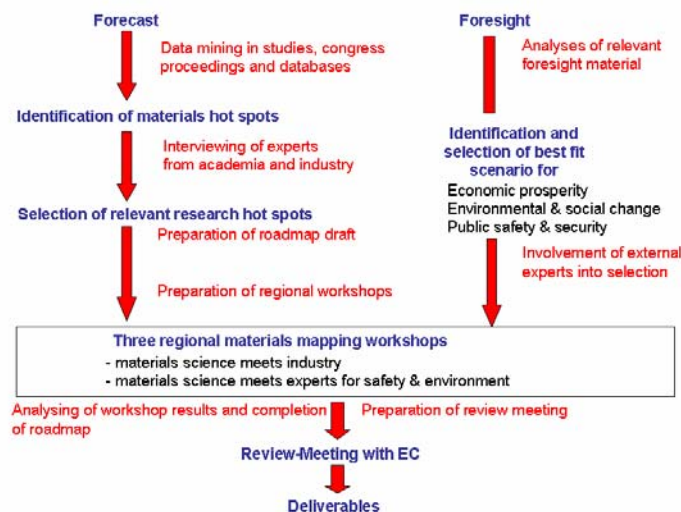


Abbildung 1: Der SMART-Arbeitsvorgang

Das Projekt begann im April 2005 mit einem Gesamtbudget von ca. 480.000 Euro für die Dauer von zwei Jahren. Das Konsortium bestand aus fünf Partnern aus vier verschiedenen Ländern (Deutschland, Frankreich, Großbritannien und Slowakei). Die bibliometrischen Analysen wurden von der Zentralbibliothek des Forschungszentrums Jülich (Tunger, Plott und Mittermaier) durchgeführt. Es gab für diese Arbeiten kein separates Arbeitspaket, sondern sie waren Bestandteile von mehreren Arbeitspaketen:

- Rastersuche nach Materialforschungszentren durch die Analyse von Datenbanken, Studien und Konferenzbänden (WP 1)
- Überprüfung des Kompetenzatlas der Materialforschung (map of excellence) und Ausarbeitung von Empfehlungen für zukünftige Forschungsaktivitäten (WP 9)
- Analyse von Angaben über Engpässe und der Leistungsvergleich der europäischen Aktivitäten (WP 10)

Das erste Arbeitspaket lieferte die Grundlagen für die bibliometrischen Analysen. In den anderen beiden Arbeitspaketen wurde mittels Bibliometrie eine europäische Karte von Spitzenforschungszentren der Materialwissenschaften erstellt und die Stellung Europas im internationalen Vergleich ermittelt („international benchmarking“). Insgesamt kann man den Anteil der bibliometrischen Studien im SMART-Projekt auf ca. 20% beziffern.

Der Fokus des SMART-Projekts lag in der Diskussion der Ergebnisse beider Arbeitsvorgänge „foresight“ und „forecast“ auf drei regionalen Workshops, mit dem Ziel, thematische Roadmaps für die Materialforschung zu erstellen. Zu diesen Workshops, welche 2006 in London, Lissabon und München stattfanden, kamen jeweils ca. 40 Experten zusammen.

Methode

Bibliometrie beschreibt die Anwendung von statistischen Methoden zur Untersuchung von wissenschaftlichen Publikationen. Mit Hilfe bibliometrischer Methoden können Trends in forschungsrelevanten wissenschaftlichen Themenbereichen aufgespürt werden, z.B. welche Werkstoffe in der Zukunft an Bedeutung gewinnen werden. Bibliometrie liefert Zahlenmaterial, (z.B. die arbeitsgruppen-, instituts- oder städtebezogene Veröffentlichungsaktivität oder Zitationsrate) mit dem aktuellen Fragestellungen (z.B. wo im internationalen Vergleich hervorragende Forschungsleistungen erbracht werden) beantwortet werden können. Um das Risiko von Fehlinterpretationen durch einzelne Ausreißer (sog. Untergrundrauschen) sowie einseitige Interpretation zu reduzieren, sollten bibliometrischen Analysen immer in Kombination mit Fachexperten Anwendung finden.

Für die bibliometrischen Analysen wurde ausschließlich der Science Citation Index von Thomson Scientific verwendet. Er eignet sich sehr für die Naturwissenschaften, deckt einen großen Bereich ab und liefert die abgefragten Daten präzise und schnell.

Die im SMART-Projekt verwendeten bibliometrischen Parameter sind wie folgt:

- Anzahl der Veröffentlichungen – Number of Publications [P]
- Dieser Parameter gibt die absolute Anzahl von Veröffentlichungen einer definierten Gruppe (z.B. Autor oder Forschungsinstitut) im beobachteten Zeitraum wieder und somit die Publikationsaktivität dieser Gruppe. „P wird ebenso als ein Größenindikator einer Forschungseinheit betrachtet“ (Van Raan, A., 2004).
- Anzahl der Zitate – Number of Citations [C]
- Dieser Parameter gibt die absolute Anzahl der Zitate wieder, die von der Größe der Forschungseinheit und/oder vom beruflichen Hintergrund des Wissenschaftlers abhängt sowie von der Anzahl der Veröffentlichungen.
- Zitationsrate – Cites per paper [CPP]
- Dieser Parameter misst den Wirkungsgrad der Veröffentlichung und somit die Wahrnehmung der Publikation von anderen Wissenschaftlern. Vorteil von CPP ist die Unabhängigkeit von der Größe der betrachteten Forschungseinheit und gibt „die durchschnittliche Anzahl von Zitate pro Veröffentlichung“ wieder (Van Raan, A., 2004).

Als Vorarbeit zu den bibliometrischen Analysen wurden im Arbeitspaket 1 mehr als 40.000 Datensätze mit Hilfe der Literaturdatenbank erhoben sowie über 300 Übersichtsartikel analysiert. Die dabei identifizierten 260 Materialforschungsthemen wurden in 36 Materialforschungsgebiete zusammengefasst. Die Literaturanalyse wies für die Werkstoffklassen

- Nanomaterialien (Nanomaterials)
- Intelligente Werkstoffe (Smart Materials)
- Biomaterialien (Bio-conceptional materials)
- Maßgeschneiderte Werkstoffe (Tailored Macroscale)
- Modellierung (Modelling)

eine hohe Relevanz für die Materialforschung aus. Diese fünf Werkstoffklassen wurden als Suchkategorien (wiederum unterteilt in insgesamt 23 Unterklassen) für die bibliometrische Analyse herangezogen (Abb.2).

Knowledge based multifunctional materials	Nano-materials	nanoparticles & -crystals
		nanocomposites
		nanofibres & -rods
		nanotubes/ fullerenes
		thin films & spintronic materials
	Smart materials	shape memory materials
		functional fluids & gels
		piezo-, ferro-, pyroelectric materials
		magneto- & electrostrictive materials
		electroactive polymers
		electro-, photo- & thermochromic materials
		tunable dielectrics
	Bio-materials	bioinspired materials
		biohybrids
		biodegradable materials
		bioinert materials
		soft matter
		bioactive materials
	Tailored Macroscale Materials for high performance applications	structural materials for extreme environments
		functional materials for extreme environments
		energy efficient materials
		electromagnetic materials
	Predictive Modelling of Materials	embedded coupled multiscale approach

Abbildung. 2: Einteilung der Materialforschung in 5 Suchkategorien mit insgesamt 23 Unterklassen

Ergebnisse

Die Anzahl der wissenschaftlichen und technischen Veröffentlichungen ist in den vergangenen Jahren stetig gewachsen. Dies gilt ebenso für die Materialwissenschaften (Abb.3), wobei vor allem auf den Gebieten der Nanomaterialien, Biomaterialien und intelligenten Werkstoffe ein exponentieller Anstieg zu verzeichnen ist.

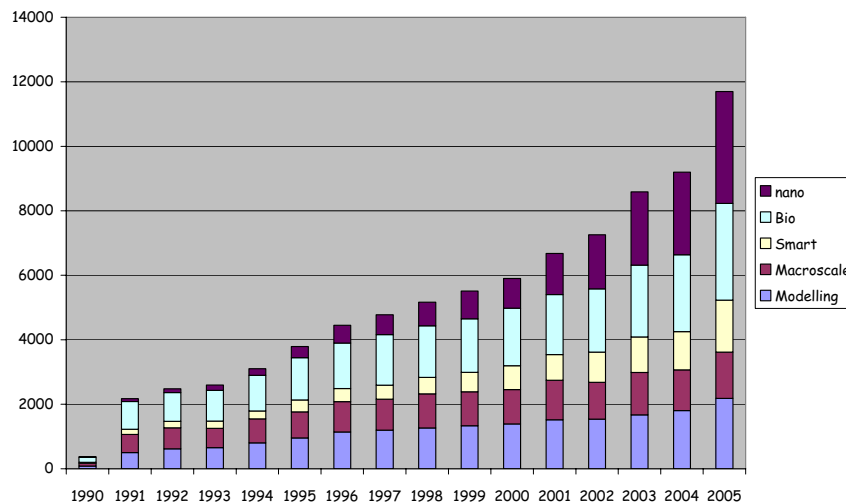


Abbildung 3: Gesamtzahl der Publikationen in den für SMART relevanten Themen

Die bibliometrischen Analysen wurden für die internationalen Leistungsvergleiche in der Materialforschung eingesetzt. Dabei wurde die Veröffentlichungsrate unter Berücksichtigung der Einwohnerzahl sowie der ökonomischen Stärke der jeweiligen Länder betrachtet. Europa befindet sich u.a. im Wettbewerb mit den USA, Japan, China sowie Südkorea und liegt auf einem der vorderen vier Plätze in vier der fünf SMART-Themen (Tabelle 1). Bei den intelligenten Werkstoffen befindet sich Europa auf dem fünften Platz.

Die Analyse wurde auch auf Städteebene durchgeführt, um die Zentren mit hoher Publikationsaktivität zu identifizieren. Das könnte ein möglicher Hinweis auf ein Spitzenforschungszentrum der Materialwissenschaften sein. Beispielsweise führen Beijing, Shanghai und Singapur die Ranking-Liste für Nanomaterialien und intelligente Werkstoffe an und lassen US-amerikanische Städte dabei zurück. Bei dieser Art von Mapping werden regionale Zentren sichtbar, deren Länder ihre Forschungsaktivitäten auf wenige Standorte konzentrieren.

Eine etwas veränderte Rankingliste ergibt sich, wenn man neben der Publikationsaktivität auch die Beachtung dieser Publikationen durch die Fachwelt, indiziert durch die Zitationsrate, betrachtet. So werden die Publikationen der Zentren Berkeley (USA), Ufa (Russland) und Karlsruhe (BRD) im Bereich der Nanomaterialien am häufigsten zitiert, obwohl sie vergleichsweise wenige Publikationen aufweisen. Ein Blick in die Inhalte der Publikationen zeigt, dass an diesen drei Standorten bekannte und exzellente Forscher auf dem Gebiet der Nanomaterialien arbeiten.

Nanomaterials				
Land	Number of articles	Inhabitants	GDP [Million \$]	Number of articles per 10 Million inhabitants
South Korea	333	48.289.100	448.979	69
USA	1796	290.342.600	10.450.000	62
Japan	627	127.214.500	4.597.140	49
EU-27	2302	460.000.000	11.506.369	47
Russia	273	144.526.300	312.496	19
China	1157	1.286.975.500	1.278.900	9

Tabelle 1: Globaler Vergleich der Publikationsaktivität auf dem Gebiet der Nanomaterialien

Betrachtet man nur Europa für sich, so wurde für Deutschland, Italien, Frankreich und Großbritannien eine hohe Publikationsaktivität in allen fünf Materialforschungsgebieten festgestellt. Polen und Spanien haben eine signifikante Publikationsanzahl auf dem Gebiet der Maßgeschneiderten Werkstoffe. Weiterhin wird in Spanien auf den Gebieten von Nano- und Biomaterialien sowie intelligenten Werkstoffe viel publiziert. Schweden weist eine deutliche Aktivität im Bereich der Modellierung auf. Diese Befunde sind vergleichbar mit den Ergebnissen der Untersuchungen zum globalen Mapping auf dem Gebiet der Biomaterialien, welche von der Universität Leiden und vom Fraunhoferinstitut 2001 durchgeführt wurden (Van Raan, A., Van Leeuwen, T., 2001).

Die Mapping-Ergebnisse anhand der Publikationsaktivität enthalten keinerlei Information über die Qualität der Forschung an diesen Standorten. Eine qualitative Aussage erfolgte vielmehr durch die Experteninterviews, die auch Fragen zur Reputation der Forschungsstandorte beinhalteten.

Diskussion / Interpretation

Anhand der oben beschriebenen Auszüge der Ergebnisse zeigt sich, dass die mittels Bibliometrie ermittelte Aktivität und Wahrnehmung von Forschung immer zu interpretieren ist. Dennoch liefern die bibliometrischen Analysen die statistische Grundlage für die Erstellung eines europäischen Kompetenzatlas der Materialforschung. Eine Veröffentlichung darüber wurde bereits publiziert (Schumacher, Tunger, Smith, Preston, Knott, 2007).

Die Ergebnisse aus den bibliometrischen Analysen, den Experteninterviews und der Durchsicht der Früherkennungsstudien wurden herangezogen, um die Themen für die Roadmapping-Workshops zu formulieren. Die Durchführung solcher Workshops war das Hauptziel des SMART-Projektes. Unter Berücksichtigung der für die Bibliometrie als Suchkategorien verwendeten fünf weitreichenden Werkstoffklassen und der in den Früherkennungsstudien ermittelten Herausforderungen und Megatrends wurden für die drei Workshops folgende Themen ausgewählt:

- Materials powering Europe
- Materials for a Better Life
- Materials for a Safe Europe

Der Prozess zur Erstellung einer thematischen Roadmap besteht aus vier Schritten in denen (1) die derzeitige Situation (2) die zukünftige Zielrichtung (3) die Hindernisse in der Weiterentwicklung und (4) die Handlungsempfehlungen zur Überwindung dieser Hindernisse identifiziert und diskutiert werden. In wechselnden Kleingruppen wurden die Inhalte des vierstufigen Roadmapping-Prozesses von den Wissenschaftlern erarbeitet. In die Beschreibung des Status quo sowie der zukünftigen Richtung flossen dabei auch die Erkenntnisse aus den Review-Artikeln und Foresightstudien sowie die Ergebnisse der bibliometrischen Analysen ein. Leider lässt sich der Umfang schwer quantifizieren. Qualitativ trug die Bibliometrie zur Identifizierung der innerhalb der Workshops diskutierten Themen bei und konnte über die Publikationsaktivität auch deren Aktualität belegen. Beispielsweise wurden auf dem Workshop „Materials for a Safe Europe“ vor allem Themen wie Sensoren, Werkstofftechnologie, Verpackung, Identifikationsüberprüfung, Sicherheitsvorkehrung sowie gesellschaftliche Belange adressiert. In Abb. 5 ist die Roadmap „Werkstoffe für die Sicherheit“ dargestellt. Hier sind die Forschungsthemen in drei verschiedenen Bereichen (1) themenübergreifende Belange (2) persönlich Sicherheitsvorkehrungen und (3) Identifikationsüberprüfung auf einer Zeitskala gegliedert.

Die Arbeiten auf der Seite der Forschungsprognose im SMART-Projekt wären ohne die bibliometrischen Studien nicht denkbar. So konnten die wichtigen Materialforschungsthemen sowie die Experten zu diesen Forschungsgebieten identifiziert werden. Die identifizierten relevanten Publikationen schufen die Grundlage, um im Endbericht des SMART-Projektes (Schumacher, G., S. Preston, A. Smith, P. Sajgalik, 2007) die Roadmapping-Kapitel inhaltlich zu untermauern.

ROADMAP IN MATERIALS FOR SECURITY

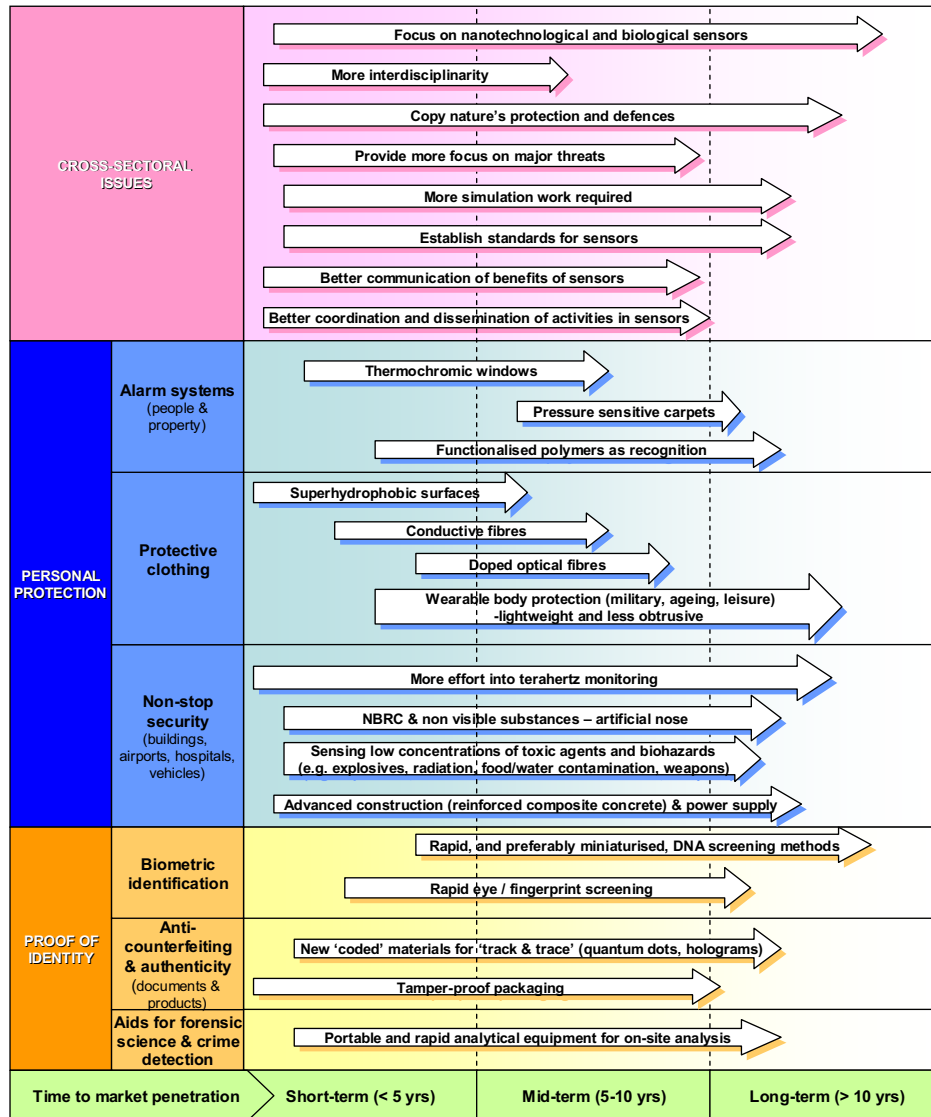


Abbildung 5: Die auf dem Workshop „Materials for a Safe Europe“ erstellte Roadmap

Literatur

Van Raan, A. (2004). Measuring Science. Handbook of Quantitative Science and Technology Research. H. F. Moed, W. Glänzel and U. Schmoch. Dordrecht, Kluwer Academic Publishers: 19-50.

Van Raan, A., T. Van Leeuwen, 2001, CWTS (www.cwts.leidenuniv.nl)

Schumacher, G., D. Tunger, A. Smith, S. Preston and B. Knott (2007). "Materials Research in Europe: Mapping Excellence and Looking Ahead." JOM: the journal of the Minerals, Metals & Materials Society 59(2): 75-77

Schumacher, G., S. Preston, A. Smith, P. Sajgalik (2007). „Future Perspectives of European Materials Research“, Verlag des Forschungszentrums Jülich.

Messen, Steuern, Regeln – zum Controlling der Helmholtz-Forschung

Henning Möller

Zusammenfassung

Mit Gründung des Helmholtz-Vereins im Jahr 2001 fand ein Paradigmenwechsel in der Steuerung der 15 nationalen Großforschungseinrichtungen, heute Helmholtz-Zentren, statt: Programmorientierte Förderung (PoF) und wissenschaftsadäquates Controlling sind an die Stelle von Globalsteuerung und Erfolgskontrolle getreten. Heute entscheiden forschungspolitische Vorgaben und internationale Evaluationen der insgesamt 29 zentrenübergreifenden Helmholtz-Programme über Forschungsbudgets für jeweils 5-jährige Planungsperioden. Wissenschaftliche Exzellenz ist hierbei die Herausforderung, Wettbewerb und Kooperation zwischen den Helmholtz-Zentren sind strategische Erfolgsfaktoren. Der hohe Anteil staatlicher Finanzierung wirft die Frage nach Nutzen und gesellschaftlicher Rechtfertigung auf. Ein zertifiziertes wissenschaftsadäquates Controlling mit integriertem Kennzahlensystem soll der Qualitätssicherung und Steuerung der Helmholtz-Forschung dienen, aber auch den gesellschaftlichen Nutzen und damit die Reputation der Helmholtz-Zentren belegen. Ausgehend von der Entwicklungsgeschichte von Planung und Erfolgskontrolle in den Großforschungseinrichtungen werden das neue Steuerungsinstrument der Programmorientierten Förderung und das wissenschaftsadäquate Controlling vorgestellt und die Problematik quantitativer Messgrößen in der Wissenschaft diskutiert. Nach 5-jähriger praktischer Erfahrung können das wissenschaftsadäquate Controlling in der Helmholtz-Gemeinschaft und seine Eignung als Steuerungs- und Regelungsinstrument noch keinesfalls abschließend beurteilt werden.

Abstract

With the foundation of the Helmholtz Association in 2001, paradigms changed in the controlling of the 15 National Research Centers, now Helmholtz Centers: Program-oriented funding and science-adequate controlling have replaced global guidance and success monitoring. Today, research policy targets and international evaluations of the 29 center-overlapping Helmholtz programs decide on research budgets for planning periods of 5 years. Scientific excellence is the challenge, competition and cooperation of the Helmholtz Centers are strategic factors of success. The high proportion of state funding raises the question of benefit and societal justification. A certified science-adequate controlling with an integrated system of benchmarks shall ensure quality assurance and controlling of Helmholtz research, but also confirm the benefit for the society and, hence, the reputation of the Helmholtz Centers. Based on the history of planning and controlling of the former National Research Centers the new controlling instruments of program-oriented funding and science-adequate

controlling are introduced and the problem of quantitative measurement in science is discussed. After five years of practical experience, however, science-adequate controlling in the Helmholtz Association and its suitability as a control instrument cannot yet be assessed conclusively.

Einführung

Ein kürzlich in der Frankfurter Allgemeinen Zeitung erschienener Artikel mit der Überschrift „Die Krankheit der Wissenschaft“ beginnt mit dem Satz: „Eine Krankheit hat die Wissenschaft befallen: Die Evaluitis.“ (Osterloh, Frey 2007). Weiter geht es: „Spitzenforschung braucht Zeit, um zu bewertbaren Ergebnissen zu führen. Dauerkontrollen begünstigen Mittelmaß und verführen zu Manipulationen. Wonach sollen sich die Geldgeber dann richten?“ Ganz ohne Evaluationen gehe es aber nicht, räumen die Verfasser ein. „Ziel müsse aber in erster Linie eine Bewertung der Qualität des Forschungsprozesses sein, nicht der Ergebnisse.“ Damit ist schon fast alles zum Controlling in der Wissenschaft gesagt.

Die quantitative Erfassung und Beurteilung von Forschungstätigkeit ist ein Bedürfnis, das mit zunehmender Distanz der Geldgeber zum Förderobjekt wächst. Verspricht man sich doch von Messen, Steuern, Regeln, kurz dem Controlling in der Forschung, Effizienz, Zielorientierung und nützliche Ergebnisse. Die letzten Prüfinstanzen sind Parlament und Öffentlichkeit, die mit Recht nach dem Nutzen der eingesetzten Haushalts- und Steuergelder für die Forschung fragen. Mit inhaltlichen Kleinigkeiten tut man sich auf dieser Ebene schwer und die Maßzahl oder der moderne Benchmark wird zum einzigen, scheinbar objektiven Erfolgskriterium. Damit befinden wir uns mitten im Dilemma der Beurteilung des Nutzens oder, abstrakter ausgedrückt, der gesellschaftlichen Relevanz von Wissenschaft und Forschung: Was gestern nützlich war, kann aber heute schon des Teufels sein. Die früheren „Kern“forschungszentren Jülich und Karlsruhe kennen das aus langjähriger Leidensgeschichte.

Mit Gründung des Helmholtz-Vereins im Jahr 2001 fand ein Paradigmenwechsel in der Steuerung der 15 deutschen Großforschungseinrichtungen, heute Helmholtz-Zentren, statt (Wissenschaftsrat 2001). Waren bis dato die so genannte Globalsteuerung der Zuwendungsgeber und wissenschaftliche Erfolgskontrolle die Steuerungsinstrumente, so weist schon die Begrifflichkeit in der neuen Helmholtz-Welt den Paradigmenwechsel aus: Programmorientierte Förderung und wissenschaftsadäquates Controlling sind heute die Leitelemente für die Steuerung der Helmholtz-Zentren. Die Beschreibung dieses neuen Systems der Helmholtz-Gemeinschaft in allen seinen Verästelungen und bürokratischen Auswüchsen würde den Rahmen dieses Vortrages sprengen. Im Folgenden soll daher der Fokus auf das neu eingeführte wissenschaftsadäquate Controlling der Helmholtz-Forschung gerichtet sein. Dabei sollen auch einige grundsätzliche Überlegungen zum Controlling in der Wissenschaft einfließen. Hilfreich ist hierbei zunächst ein kurzer Rückblick auf die Entwicklung von Planung und Erfolgskontrolle in den

Großforschungseinrichtungen, die jetzt schon vielfach auf ein fünfzigjähriges Bestehen zurückblicken.

Zur Entwicklung von Planung und Erfolgskontrolle in den Großforschungseinrichtungen

Anfang der 1960er Jahre, in der Gründungsphase der Großforschungseinrichtungen, formulierte der damalige Staatssekretär im Forschungsministerium in der lesenswerten Schrift „Die Großforschung und der Staat“ alle Schlagworte, die uns auch heute in der renovierten Helmholtz-Gemeinschaft bewegen (Cartillierie 1963). Damals fielen solche Stichworte wie „mangelnde Flexibilität“, seinerzeit auf die Bundesforschungsanstalten gemünzt, „Anziehungskraft für Spitzenkräfte“, „Personalaustausch“ und natürlich auch „Kontrolle über die Verwendung öffentlicher Gelder“. In dieser Frühzeit stand die Frage der Bewertung der Forschungsleistung und des wissenschaftlichen Erfolges noch nicht im Vordergrund der Diskussion. Offenbar existierte seinerzeit noch ein gemeinsames Grundverständnis zur programmatischen Zielsetzung und forschungspolitischen Absicht zwischen den wichtigen Partnern Staat, Wirtschaft und Wissenschaft. Angesichts der damaligen forschungspolitischen Akteure wurde wissenschaftliche Exzellenz stillschweigend vorausgesetzt. Eine Überprüfung und laufende Kontrolle ihrer wissenschaftlichen Arbeit außerhalb der Wissenschaftsgemeinschaft dürften die Gründungsväter der Großforschung daher entrüstet von sich gewiesen haben.

Was ist aber über die Jahre aus diesem Vertrauensverhältnis zwischen den wichtigen Akteuren der staatlichen Großforschung, Zuwendungsgebern und Wissenschaftlern geworden? Die Soziologen Rosewitz und Schimank formulierten dies Ende der 1980er Jahre wie folgt (Mayntz et al. 1988): „Erst interventionsresistente Umwelt-inadäquanz konstituiert Verselbständigung. Bevor sie attestiert wird, müssen also ernsthafte Versuche insbesondere politischer Akteure, die wahrgenommene Umwelt-inadäquanz des betreffenden Teilsystems durch entsprechende Steuerungsmaßnahmen zu beheben, fehlgeschlagen sein“. In schlichter Übersetzung dürfte das heißen, der Staat zahlt und die Großforschung hat es sich in der Nische gemütlich gemacht. Dieses Verdikt aus dem Max-Planck-Institut für Gesellschaftsforschung legte ein über die Jahre gewachsenes Problem der Divergenz von Globalsteuerung, Freiheit der Forschung, Forschungsfinanzierung und Nützlichkeit frei (Hohn, Schimank 1990). Der Großforschungsinsider wird der holzschnittartigen soziologischen These nur schwer folgen können, dennoch muss konstatiert werden, dass im Laufe der 1980er Jahre der Grundkonsens zur Aufgabenstellung mancher Großforschungszentren zwischen den wichtigen Akteuren wohl auseinandergedriftet war (Weule et al. 1994). Waren in der Gründungsphase Staat und Großforschung noch mit gemeinsamen Zielen gestartet, so hatte sich jetzt ein diffuses Mit- und Gegeneinander entwickelt. Folgerichtig rückten Steuerung und Erfolgskontrolle dann im Laufe der 1990er Jahre als Thema zunehmend in den Vordergrund.

Trotz des anfänglichen Vertrauensvorschlusses hatten sich über die Jahre ausdifferenzierte Verfahren der Planung und Erfolgskontrolle mit regelmäßigem Berichtswesen, vielfältigen Evaluationsverfahren, Strukturkommissionen und wissenschaftlichen Beiräten in den einzelnen Großforschungseinrichtungen etabliert (HGF 1997). Entsprechend der unterschiedlichen Aufgabenstellung und Struktur der rechtlich unabhängigen Großforschungseinrichtungen variierten folgerichtig die individuellen Planungs- und Steuerungsinstrumente. Mit Gründung des Helmholtz e.V. im Jahr 2001 ging eine Vereinheitlichung von Planung und Kontrolle einher und seither sind die programmorientierte Förderung und das wissenschaftsadäquate Controlling beherrschende Elemente für die Steuerung der heutigen, nach wie vor rechtlich unabhängigen Helmholtz-Zentren (Möller 2003, Popp 2003).

Steuerung und Regelung der Helmholtz- Forschung – Die programmorientierte Förderung und das wissenschaftsadäquate Controlling

Die Konzeption eines zentrenübergreifenden Planungs- und Steuerungsinstrumentes, das, wie bei der programmorientierten Förderung, vor allem auch Budgetkonsequenzen haben sollte, war nicht ganz neu. Ähnliche Verfahren wurden bereits in den 1990er Jahren für die Universitäten in Großbritannien, den Niederlanden und den ETH-Bereich in der Schweiz eingeführt (HEFCE 1996, VSNU 1994, Lange 1997). Die Diskussion in der Helmholtz-Gemeinschaft zur künftigen Zentrensteuerung startete jedoch unter einer mehr industrieorientierten Überschrift, nämlich dem Konzept der Balanced Score Card. Dahinter steckt die Idee, dass aus hochaggregierten Unternehmenszielen operative Ziele auf Arbeitsebene abgeleitet werden, die dann mit entsprechenden Kennwerten gemessen werden können. Also aus der vom Vertreter verkauften Stückzahl von Konservendosen wird schließlich Shareholder Value. Beim Versuch, ein Balanced-Score-Card-System in der Helmholtz-Forschung zu etablieren, hat sich bald herausgestellt, dass in der vielgestaltigen Forschungswelt der Helmholtz-Zentren mit Bau und Betrieb von Großgeräten, Vorsorgeforschung, grundlagen-, anwendungsorientierter und industrienaher Forschung einheitliche unternehmerische Ziele nur schwer definierbar sind. Die Helmholtz-Forschung ist komplex, inhomogen und oft sehr langfristig angelegt und unterscheidet sich dadurch von der Forschung in der Max-Planck-Gesellschaft oder Fraunhofer-Gesellschaft, die sich entweder am Kriterium der wissenschaftlichen Exzellenz (MPG) oder der Industrierelevanz (FhG) einheitlicher bewerten lässt.

Ergebnis eines längeren Diskussionsprozesses in der Helmholtz-Gemeinschaft war schließlich die Einführung eines eher klassischen und wissenschaftsüblichen Peer-Review-Verfahrens, das sich auf einen akzeptierten Kanon von Kennwerten stützt. Diese Kennwerte finden sich durchgehend auch in jährlichen Programm- und Fortschrittsberichten wieder.

Das Instrument der programmorientierten Förderung lässt sich in seinen Grundzügen wie folgt skizzieren. Auf der Grundlage von forschungspolitischen Vorgaben der

Zuwendungsgeber begutachten externe, international besetzte Gutachterkommissionen die zentrenübergreifenden Helmholtz-Programme und geben Finanzierungsempfehlungen für 5-Jahresperioden. Nach einem mehrstufigen Entscheidungsprozess leiten sich hieraus schließlich die Budgetanteile der Zentren aus dem verfügbaren Budget für die einzelnen Programme ab. Die zentralen Entscheidungskriterien für die Gutachter sind in der Reihenfolge der Bedeutung 1. Wissenschaftliche Exzellenz, 2. strategische Relevanz und 3. Angemessenheit des Mitteleinsatzes. Sowohl für die internationalen Gutachter als auch für die Begutachteten war neu an dem Verfahren, dass am Ende des aufwändigen Evaluierungsprozesses eine bedeutende Budgetentscheidung in einer Größenordnung von einigen 100 Mio € pro Jahr und Programm über die Planperiode von 5 Jahren steht.

Die in den Helmholtz-Zentren wohlbekannten Peer-Review-Verfahren nach althergebrachtem Muster hatten sich bisher weitgehend auf die qualitative Bewertung erzielter Forschungsergebnisse (ex post) und der daraus abgeleiteten Handlungsoption für die Zukunft (ex ante) gestützt. Böse Worte fielen in der Regel dabei nicht und entsprechende Gutachten haben das Beste gelobt und dem Mittelmäßigen Besserung empfohlen. Kritik war nur für den Insider zwischen den Zeilen zu finden. Das neue Begutachtungsverfahren der Helmholtz-Gemeinschaft vergibt dagegen differenzierte Noten auf der Skala 1 (aussondern) bis 7 (internationale Spitze). Mit dem Prinzip „das Beste muss das Gute verdrängen“ soll Mittelmaß nun keine Überlebenschance mehr haben. Nach den Erfahrungen der ersten Evaluationsrunde wurde der Evaluationsprozess mit Ernst und Fairness aller Beteiligten durchgeführt. Die Gutachter schienen sich auch ihrer verantwortungsvollen Aufgabe, programmatische und finanzielle Entscheidungen von großer Tragweite zu treffen, sehr bewusst zu sein.

Während die programmorientierte Förderung und der Begutachtungsprozess bewusst auf eine zukunftsgerichtete strategische Zielsetzung (ex ante) ausgerichtet sind, versucht das wissenschaftsadäquate Controlling die getroffenen Versprechungen zu verifizieren. Hierzu dienen ein jährlicher zentrenübergreifender Programmbericht an den Helmholtz- Senat sowie ein zentrenindividueller Fortschrittsbericht an Zuwendungsgeber und Aufsichtsorgane. Diese Berichte werden mit Kennzahlen unterlegt, die im folgenden etwas eingehender diskutiert werden sollen.

Messung der Helmholtz- Forschung – Die Kennzahlen

Die Messung von Forschung, d.h. die quantitative Bewertung von Forschungsprozessen und -ergebnissen, ist eine eigene Wissenschaft, allerdings eine empirische (Weingart et al. 1991, Rip 1997, Erche 2002). Wenn klare Ziele definiert sind, lässt sich der Erfolg in der Regel auch einfach messen: Beim Fernsehen ist jedem die Quote bekannt, die zudem noch zeitnah, am Tag nach der Sendung vorliegt. Das Analogon der Quote legt zugleich das Dilemma der Messung von Wissenschaft und Forschung offen: Der Erfolg oder Nutzen von

wissenschaftlichen Ergebnissen stellt sich oft erst mit erheblicher Zeitverzögerung ein und ist dann vielleicht auch gar nicht mehr eindeutig auf einen Urheber, das kann auch eine Forschungsgruppe oder -institution sein, zurückzuführen. Der Idealfall des Nobelpreisträgers, der manchmal erst nach Jahrzehnten für eine Bahn brechende Arbeit ausgezeichnet wird, dürfte eher die Ausnahme sein. In der Regel vollzieht sich Forschung als ein mehr oder weniger zielgerichteter Prozess mit wichtigen und unwichtigen Zwischenergebnissen im Umfeld der jeweiligen wissenschaftlichen Gemeinschaft. Die Beurteilung von wichtig und unwichtig kann oft auch nicht zeitnah getroffen werden. Unter dem Gesichtspunkt der gesellschaftlichen Relevanz kann sich wichtig und unwichtig auch einer objektiven Beurteilung entziehen, wie es das Beispiel der Kerntechnik gezeigt hat. Entsprechend schwierig ist die Messung und Beurteilung von wissenschaftlicher Arbeit und ebenso variantenreich sind quantitative Kennzahlensysteme in der Forschung. Dies kann auch als ein Symptom der Hilflosigkeit gewertet werden. Dennoch lassen sich bei nüchterner Überlegung einige forschungstypische Messgrößen ableiten.

Relativ einfach können die Inputs in den Forschungsprozess gemessen werden: In der Regel Geld als Maß für die wissenschaftliche Ausstattung und das wissenschaftlich aktive Personal als Maß für das intellektuelle Kapital. Als Output manifestieren sich die Ergebnisse wissenschaftlicher Arbeit primär in Wort und Schrift, also Publikationen, Vorträge, Patente und nicht zu vergessen in der Lehre. Der inzwischen etablierte Forschungszweig der Bibliometrie mit szientometrischen Methoden belegt, dass der Publikation als Ergebnis wissenschaftlicher Arbeit eine dominante Rolle zukommt (van Raan 2003). Typisch für den Ingenieurbereich ist die Entwicklung von Prototypen und Verfahren einschließlich IT-Entwicklungen. Schließlich sind Konzeption, Bau und Betrieb von Großexperimenten sowie die Bereitstellung von Orientierungswissen und Handlungsoptionen in der Vorsorgeforschung charakteristisch für die Helmholtz-Forschung.

Die schlichte Quantifizierung dieser Output-Größen besagt aber nichts über deren Qualität. Eine Publikation kann ein Übersichtsartikel im Wissenschaftsteil einer Tageszeitung oder eine nobelpreiswürdige Entdeckung in einer international renommierten Fachzeitschrift sein. Bei Patenten ist es ganz ähnlich: Liegt eine Hightechlösung für ein nicht vorhandenes Problem vor oder legt ein Patent die Grundlage für eine Innovation, die eine breite wirtschaftliche Verwertung verspricht? Die jährliche Messung der Forschungsergebnisse der Helmholtz-Programme mit insgesamt 20 quantitativen Erfolgsindikatoren orientiert sich an den genannten forschungstypischen direkten Outputgrößen – Publikationen und Technologietransfer – sowie an einigen indirekten Qualitätskennwerten wie DfG-Förderung, Kooperationen und Gastwissenschaftler. Der Kennwert „Drittmittelerträge“ fällt aus der wissenschaftstypischen Systematik heraus und bedarf einer besonderen Betrachtung.

Die Schwierigkeiten bei der Messung des wissenschaftlichen Outputs beginnt bei der qualitativen Bewertung der nackten Zählgröße. Beim Kriterium wissenschaftliche

Produktivität und Exzellenz wird der referierten Publikation ein hoher Rang eingeräumt. Der Wert von Patenten lässt sich an der Zahl von Lizenzen oder Lizenzeinnahmen messen. Keine der Bewertungsgrößen ist jedoch unproblematisch. Zitierkartelle bei Publikationen oder eine naturwissenschaftslastige Auswahl von gelisteten referierten Publikationsorganen beim Information Sciences Institute (ISI) können etwa im Ingenieursbereich schon zu Fehlurteilen führen. Selbst bei der Bewertung von Patenten mit der eher harten Währung von Lizenzen und Lizenzeinnahmen kann es zu Fehlbewertungen kommen, wenn eine Lizenz zum Beispiel nicht verwertet wird, sondern aus wettbewerblichen Gründen im Panzerschrank eines Unternehmens versauert. Jeder einzelne Messwert mit seiner Ausdifferenzierung in Bewertungsgrößen ist bei genauerer Untersuchung mit Definitions- und Beurteilungsproblemen behaftet.

Die Erfahrung zeigt, dass die Aussagekraft von Kennwerten zunimmt, wenn einerseits Kennwertprofile und andererseits Zeitreihen über die Jahre betrachtet werden. Diese beiden zusätzlichen Auswerteparameter lassen bei aller Unschärfe des singulären Kennwertes Stärken- und Schwächenprofile von Forschergruppen oder Forschungsprogrammen erkennen. Welches Symptom allerdings hinter einem qualitativen und zeitlichen Kennwertmuster steckt, kann nur die fachliche Beurteilung im Peer-Review-Verfahren erbringen (Wissenschaftsrat 2002, 2005, Jansen 2007). Es ist ähnlich wie in der medizinischen Diagnose: Messwerte können nicht das ganzheitliche Fachurteil des Arztes ersetzen sondern nur Indizien liefern.

Bewertung der Helmholtz-Forschung – Reputation und Relevanz

Hinsichtlich Aufgabenstellung und Auftrag der Helmholtz-Zentren verdienen zwei Beurteilungsgrößen eine besondere Betrachtung: Dies sind die wissenschaftliche Reputation und die gesellschaftliche Relevanz. Diese beiden Merkmale sollen hier bewusst diffus verstanden werden. Im Detail werden sich wissenschaftliche Exzellenz und gesellschaftliche Relevanz immer auf die eine oder andere Weise belegen lassen. Hier geht es aber um mehr, nämlich den Ruf der Helmholtz-Forschung an sich.

Auslöser für die Gründung des Helmholtz e.V. und die Einführung der programmorientierten Forschung mit wissenschaftsadäquatem Controlling waren u.a. diese beiden forschungspolitischen Merkmale, an denen sich die Helmholtz-Zentren ebenso wie auch andere Forschungsinstitutionen in der Öffentlichkeit messen lassen müssen. Mit Recht wird man die Frage stellen, ob sich wissenschaftliche Reputation und gesellschaftliche Relevanz überhaupt mit Indikatoren messen lassen. Nun ist es aber leider so, dass in der heutigen Welt der „Fakten, Fakten, Fakten“ quantitative Benchmarks besser kommunizierbar sind als komplexe Forschungsergebnisse. Zugespißt und auf den Punkt gebracht wird hier gerne auf zwei beliebte Kenngrößen zurückgegriffen, nämlich die Zahl der Nobelpreise als Nachweis für die wissenschaftliche Reputation und die Höhe der Drittmitteleinnahmen als Nachweis für den gesellschaftlichen Nutzen (Relevanz). Wissenschaftliche Exzellenz wird mit Preisen honoriert, Innovationen gehen mit Ertrag in die wirtschaftliche Verwertung.

Beide Kennwerte sind leicht vermittelbar und öffentlichkeitswirksam. Bei beiden Kennwerten hat die Helmholtz-Gemeinschaft allerdings starke Konkurrenz, auf der einen Seite bei der Max-Planck-Gesellschaft und auf der anderen Seite bei der Fraunhofer Gesellschaft. In dem breit angelegten Forschungsfeld dazwischen bewegen sich die Helmholtz-Zentren.

Zusammenfassung und Ausblick

Was haben nun nach fast fünfjähriger Übung die programmorientierte Förderung und das wissenschaftsadäquate Controlling im Helmholtz e.V. gebracht? Für die Helmholtz-Forschung ist charakteristisch, dass sie vielgestaltig ist und sich auf mittel- und längerfristiger Zeitskala abspielt. Ob sich die programmorientierte Förderung und das wissenschaftsadäquate Controlling als Steuer- und Regelinstrument bewährt haben, kann heute noch nicht abschließend beurteilt werden. Die jetzt zu Ende gehende erste 5-jährige Förderperiode hat jedenfalls gezeigt, dass sich die Helmholtz-Forschung beim Hauptkriterium „wissenschaftliche Exzellenz“ im internationalen Vergleich nicht verstecken muss. Bleibt die Frage der strategischen Ausrichtung. Inwieweit sich das wissenschaftsadäquate Controlling hier gegenüber dem Althergebrachten zu einem überlegenen Steuerungsinstrument entwickelt hat, muss sich noch erweisen. Im komplexen Verhaue von zuweilen konkurrierender, aber unerlässlichen Zentrensteuerung und der übergeordneten Helmholtz-Programmsteuerung und trotz der grassierenden Evaluitis hat sich die Helmholtz-Forschung einstweilen als robust und überlebensfähig erwiesen.

Wie auch internationale Erfahrungen bestätigen, bedürfen Controlling-Instrumente der kontinuierlichen Überprüfung und Weiterentwicklung hinsichtlich Aussagekraft und Validität. Das in der Wissenschaft geübte und übliche Peer-Review-Verfahren ist unersetzbar und kann durch quantitative Indikatoren nur gestützt werden. Die vielgestaltige Helmholtz-Forschung führt in das Dilemma, dass der Nachweis von wissenschaftlicher Exzellenz und gesellschaftlicher Relevanz mit einem griffigen und öffentlich vermittelbaren Messwert nur schwer darstellbar ist.

Neuere Ansätze des Benchmarking orientieren sich nicht am Forschungsooutput sondern am Forschungsimpact, also dem unmittelbaren Nutzen von Forschungsergebnissen bei den verschiedenen Adressaten. In der Bibliografie ist diese Sichtweise bereits durch die Messung von Zitaten verankert. In vielen anderen Bereichen ist dies erheblich schwieriger. Wie lässt sich zum Beispiel messen, ob ein Forschungsergebnis in die industrielle Produktion eingeflossen ist? Sind durch Forschung Arbeitsplätze geschaffen worden oder hat Vorsorgeforschung staatliches Handeln beeinflusst? Hier tut sich ein neues Feld für die Definition von Indikatoren auf.

Bei allem Messen, Steuern, Regeln tut der zahlenverliebte Controller gut daran, gelegentlich in sich zu gehen und über die Maxime von Hermann von Helmholtz zu sinnieren: „Wer bei der Verfolgung von Wissenschaften nach unmittelbarem Nutzen fragt, kann ziemlich sicher sein, dass er vergebens jagen wird.“

Literatur

Margit Osterloh, Bruno S. Frey, Die Krankheit der Wissenschaft, Frankfurter Allgemeine Zeitung vom 21.07.2007

Wissenschaftsrat, Systemevaluation der HGF – Stellungnahme des Wissenschaftsrates zur Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren, Drs. 4755/01, Berlin, 19.01.2001.

Wolfgang Cartellieri, Die Großforschung und der Staat, Sonderdruck aus Heft 4 der Schriftenreihe „Forschung und Bildung“ des Bundesministers für wissenschaftliche Forschung, Heidelberg, 1963.

Renate Mayntz et al., Differenzierung und Verselbständigung – Zur Entwicklung gesellschaftlicher Teilsysteme, Campus Verlag, Frankfurt/New York, 1988, S. 297.

Hans-Willy Hohn, Uwe Schimank, Konflikte und Gleichgewichte im Forschungssystem-Akteurkonstellationen und Entwicklungspfade in der staatlich finanzierten außeruniversitären Forschung, 1990 Campus Verlag, Frankfurt/Main.

Hartmut Weule et al., Gutachten zur Zusammenarbeit GFE/Industrie-Identifikation und Bewertung der thematischen Ansatzpunkte für eine engere Kooperation der Industrie mit der KFA Jülich und dem KfK Karlsruhe, Mai 1994.

HGF-Dokumentation 15, Begutachtungen in der Helmholtz-Gemeinschaft – Grundsätze und Verfahren, Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren, Bonn 1997.

H. Möller, Kontrolle ist gut, Vertrauen ist besser – Zur Entwicklungsgeschichte von Steuerung und Erfolgskontrolle in der Helmholtz-Gemeinschaft, in: Technikfolgenabschätzung – Theorie und Praxis, Nr. 1, 12. Jahrgang, März 2003, Institut für Technikfolgenabschätzung und Systemanalyse, Forschungszentrum Karlsruhe.

M. Popp, Erste Schritte in die Programmorientierte Förderung – ein Abenteuerbericht, in: Technikfolgenabschätzung – Theorie und Praxis, Nr. 1, 12. Jahrgang, März 2003, Institut für Technikfolgenabschätzung und Systemanalyse, Forschungszentrum Karlsruhe.

1996 Research Assessment Exercise- Criteria for Assessment, November 1995, hefce- Higher Education Funding Council for England.
Quality Assessment of Research, Protocol 1994, VSNU – Vereniging van Samenwerkende Nederlandse Universiteiten, Utrecht/ The Netherlands.

Hans Rüdiger Lange, Die Praxis leistungsorientierter Forschungsfinanzierung in Großbritannien Eine Fallstudie am Fachbereich Physik der Universität St. Andrews, Wissenschaftsmanagement 3, Mai/Juni 1997.

Weingart, Peter, Wissenschaftsindikatoren als soziale Konstruktion und ihre Realität, in: Weingart et al. (Hg.), Indikatoren der Wissenschaft und Technik – Theorie, Methoden, Anwendungen, Campus Verlag, Frankfurt/New York, 1991.

Arie Rip, Higher Forms of Nonsense in The Future of the Peer Review System, September 1997, Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Den Haag/ Nederland.

Bettina Erche, Evaluation der Evaluation und so weiter - Universitätssysteme im Stress, Neue Zürcher Zeitung, Nr. 24, 30.01.2002

A.F.J. van Raan, The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments, in: Technikfolgenabschätzung – Theorie und Praxis, Nr. 1, 12. Jahrgang, März 2003, Institut für Technikfolgenabschätzung und Systemanalyse, Forschungszentrum Karlsruhe.

Wissenschaftsrat, Empfehlungen zu Querschnittsbegutachtungen in der Forschung, Drs. 5360/02, Berlin, 12. Juli 2002.

Wissenschaftsrat, Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates, Drs. 6966-05, Bremen, 11. November 2005.

Dorothea Jansen (Hrsg.), Forschungspolitische Thesen der Forschergruppe Governance der Forschung: „Rahmenbedingungen für eine leistungsfähige öffentlich finanzierte Forschung“ anlässlich der Tagung „Neue Governance für die Forschung“, Berlin-Brandenburgische Akademie der Wissenschaften, 14. und 15. März 2007.

Liste der Autoren

Dr. Rafael Ball

Forschungszentrum Jülich GmbH
Zentralbibliothek
52425 Jülich
E-Mail: r.ball@fz-juelich.de

Lydia Bauer

Hochschule für Technik und Wirtschaft (HTW) Chur
Swiss Institute for Information Research
Ringstraße / Pulvermühlestraße 57
7004 Chur (Schweiz)
E-Mail: lydia.bauer@fh-htwchur.ch

Ronald Beullens

Vrije Universiteit Brussel
University Library
Pleinlaan 2
1050 Brussel (Belgien)

Prof. Dr. Christoph Bläsi

Universität Erlangen-Nürnberg
Institut für Buchwissenschaft
Harfenstraße 16
91054 Erlangen
E-Mail: christoph.blaesi@buchwiss.uni-erlangen.de

Nadja Böller

Hochschule für Technik und Wirtschaft (HTW) Chur
Swiss Institute for Information Research
Ringstraße / Pulvermühlestraße 57
7004 Chur (Schweiz)
E-Mail: nadja.boeller@fh-htwchur.ch

Dr. Jan Brase

Technische Informationsbibliothek Hannover
DOI Registrierung
Welfengarten 1b
30167 Hannover
E-Mail: jan.brase@tib.uni-hannover.de

Prof. Dr. Ir. Koenraad Debackere

Katholieke Universiteit Leuven
Diensten Algemeen Beheer
Krakenstraat 3
3500 Leuven (Belgien)
E-Mail: koenraad.debackere@abh.kuleuven.be

Dr. Michael Diepenbroek

Universität Bremen
MARUM – Zentrum für Marine Umweltwissenschaften
Postfach 330 440
28334 Bremen
E-Mail: mdiepenbroek@pangaea.de

Prof. Dr. Dr. Wolfgang Glänzel

Katholieke Universiteit Leuven
Faculteit ETEW
Dekenstraat 2
3000 Leuven (Belgien)
E-Mail: wolfgang.glaenzel@econ.kuleuven.de

Dr. Hannes Grobe

Stiftung Alfred-Wegener-Institut für Polar- und Meeres-
forschung in der Helmholtz-Gemeinschaft
Am Handelshafen 12
27570 Bremerhaven
E-Mail: hgrobe@awi-bremerhaven.de

Dr. Peter Haber

Universität Basel
Historisches Seminar
Hirschgässlein 21
4051 Basel (Schweiz)
E-Mail: peter.haber@digitalpast.net

Dr. Christian Hänger

Universitätsbibliothek Mannheim
Im Schloss
68131 Mannheim
E-Mail: christian.haenger@bib.uni-mannheim.de

Dr. Leni Helmes

Fachinformationszentrum Karlsruhe
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Leni.Helmes@fiz-karlsruhe.de

Prof. Dr. Josef Herget

Hochschule für Technik und Wirtschaft (HTW) Chur
Swiss Institute for Information Research
Ringstraße / Pulvermühlestraße 57
7004 Chur (Schweiz)
E-Mail: josef.herget@fh-htwchur.ch

Sonja Hierl

Hochschule für Technik und Wirtschaft (HTW) Chur
Swiss Institute for Information Research
Ringstraße / Pulvermühlestraße 57
7004 Chur (Schweiz)
E-Mail: sonja.hierl@fh-htwchur.ch

Jan Hodel

Pädagogische Hochschule der Nordwestschweiz
Institut für Forschung und Entwicklung
Kasernenstraße 20
5000 Aarau (Schweiz)
E-Mail: jan.hodel@fhnw.ch

Dr. Annette Holtkamp

DESY
Notkestraße 85
22607 Hamburg
E-Mail: Annette.Holtkamp@desy.de

Susanne von Itter

European Association of Development Research
and Training Institutes (EADI)
Kaiser-Friedrich-Straße 11
53113 Bonn
E-Mail: itter@eadi.org

Miloš Jovanović

Fraunhofer INT
Am Appelsgarten 2
53864 Euskirchen
E-Mail: Milos.Jovanovic@int.fraunhofer.de

Dr. Jens Klump

Geoforschungszentrum Potsdam
Telegrafenberg A17
14473 Potsdam
E-Mail: jklump@gfz-potsdam.de

Christine Krätzsch

Universitätsbibliothek Mannheim
Im Schloss
68131 Mannheim
E-Mail: christine.kraetzsche@bib.uni-mannheim.de

Dr. Harald Krottmaier

Technische Universität Graz
Institut für Computergraphik und Wissensvisualisierung (CGV)
Inffeldgasse 16c
8010 Graz (Österreich)
E-Mail: harald.krottmaier@tugraz.at

Show-Ling Lee-Müller

Forschungszentrum Jülich GmbH
Projektträger Jülich
52425 Jülich
E-Mail: ptj-smart-ssa@fz-juelich.de

Dr. Steffen Leich-Nienhaus

DaimlerChrysler AG
CFM – Corporate Information and Research Management
70546 Stuttgart
E-Mail: steffen.leich@daimlerchrysler.com

Dominic Mainz

Heinrich-Heine-Universität Düsseldorf
Institut für Informatik, Lehrstuhl für Bioinformatik
Universitätsstraße 1
40225 Düsseldorf
E-Mail: mainz@cs.uni-duesseldorf.de

Indra Mainz

Heinrich-Heine-Universität Düsseldorf
Institut für Physikalische Biologie, AG Biocomputing
Universitätsstraße 1
40225 Düsseldorf
E-Mail: mainzi@biophys.uni-duesseldorf.de

Dr.-Ing. Henning Möller

Forschungszentrum Karlsruhe
Leiter der Stabsabteilung Planung, Außenbeziehungen
und Erfolgskontrolle
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Henning.Moeller@pae.fzk.de

Dr. Wiebke Oeltjen

Universität Hamburg
Regionales Rechenzentrum RRZ,
MyCoRe Geschäftsstelle
Schlüterstraße 70
20146 Hamburg
E-Mail: wiebke.oeltjen@uni-hamburg.de

Ingo Paulsen

Heinrich-Heine-Universität Düsseldorf
Institut für Informatik, Lehrstuhl für Bioinformatik
Universitätsstraße 1
40225 Düsseldorf
E-Mail: paulsen@cs.uni-duesseldorf.de

Prof. Dr. phil. Dr. med. habil. Ernst Pöppel

Ludwig-Maximilians-Universität München
Institut für Medizinische Psychologie
Goethestraße 31
80336 München
E-Mail: ernst.poeppel@med.uni-muenchen.de

Jim Pringle

Thomson Scientific
3501 Market Street
Philadelphia, PA 19014-3302 (USA)
E-Mail: james.pringle@thomson.com

Matthias Razum

Fachinformationszentrum Karlsruhe
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Matthias.Razum@fiz-karlsruhe.de

Ute Ruznak

Fachinformationszentrum Karlsruhe
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Ute.Ruznak@fiz-karlsruhe.de

Dr. Gerd Schumacher

Forschungszentrum Jülich GmbH
Projektträger Jülich
52425 Jülich
E-Mail: g.schumacher@fz-juelich.de

Elena Semenova

Humboldt-Universität zu Berlin
Hermann-von-Helmholtz-Zentrum für Kulturtechnik
Unter den Linden 6
10099 Berlin
E-Mail: elena.semenova@culture.hu-berlin.de

Martin Stricker

Humboldt-Universität zu Berlin
Hermann-von-Helmholtz-Zentrum für Kulturtechnik
Unter den Linden 6
10099 Berlin
E-Mail: martin.stricker@rz.hu-berlin.de

Dirk Tunger

Forschungszentrum Jülich GmbH
Zentralbibliothek
52425 Jülich
E-Mail: d.tunger@fz-juelich.de

Patrick Vanouplines

Vrije Universiteit Brussel
University Library
Information and Documentation
Pleinlaan 2
1050 Brussel (Belgien)
E-Mail: Patrick.Vanouplines@vub.ac.be

Anne-Katharina Weilenmann

Schweizerische Nationalbibliothek
Hallwylstraße 15
3003 Bern (Schweiz)
E-Mail: anne-katharina.weilenmann@nb.admin.ch

Karin Weishaupt

Institut Arbeit und Technik
Forschungsschwerpunkt Wissen und Kompetenz
Munscheidstraße 14
45886 Gelsenkirchen
E-Mail: weishaupt@iat.eu

Katrin Weller

Heinrich-Heine-Universität Düsseldorf
Institut für Sprache und Information,
Abteilung für Informationswissenschaft
Universitätsstraße 1
40225 Düsseldorf
E-Mail: weller@uni-duesseldorf.de

Register

A

Ähnlichkeitsfunktion 129
 Akzeptanz 128, 193, 195, 197, 200
 Anschauungswissen *Siehe* Wissen:
 bildliches
 Ansoff 238
 Anwendungswissen 135, 142
 Archivierung 148, 150, 151, 156, 164
 arXiv.org 177

B

Bayerischen Staatsbibliothek München
 176
 Bedürfnisse 202
 begriffliches Wissen *Siehe* Wissen
 explizites
 Bibliometrie 209, 211, 212, 213, 214,
 215, 216, 217, 235, 236, 238, 239,
 243, 263, 265, 266, 269
 Bibliothek 87, 90, 91, 126
 Bibliothekssystem 147, 148
 BibSonomy 124, 128, 131
 Biologie 147, 148
 Blended Learning 81, 83
 Blog 97
 Blogsoftware 100
 Buchhandlung 141

C

CAH *Siehe* Computer Aided
 Historiography
 CERN 182
 COCOON 164
 Collexis 127, 128, 130
 Community 35
 Community of Practice 98
 Community-Bildung 117
 Computer Aided Historiography 75, 78
 Controlling
 wissenschaftsadäquates 273, 274,
 276, 277, 279, 280
 Copy-Paste Wissenschaft 111
 Cyberscience 25, 27
 Cyberwissenschaft *Siehe*
 Cyberscience

D

DaimlerChrysler AG 107
 Datenmanagement 149
 Datenzentren 147, 148
 Datenzentrum 163
 DDC *Siehe* Dewey Decimal
 Classification
 del.icio.us 100, 102, 126
 Department for International
 Development UK 98
 Deutsche Forschungsgemeinschaft
 61, 62, 123, 149, 151, 161, 163,
 169, 176, 193, 198
 DevExpertise/DevResearch 100
 Dewey Decimal Classification 64, 126
 DFG *Siehe* Deutsche
 Forschungsgemeinschaft
 DFID *Siehe* Department for
 International Development UK
 DIAMOND 83, 86
 Didactical Approach for Media
 Competence Development 83,
 Siehe DIAMOND
 Digital Object Identifier *Siehe* DOI
 DILIGENT 28
 Directory of Open Access Journals
 Siehe DOAJ
 DOAJ 223, 225, 227, 229, 230, 233
 DOI 152, 156, 159, 161, 162, 164
 DOIDB 164, 165
 Dokumentenserver *Siehe* Repository
 Dr. Städtler 136

E

EADI *Siehe* Europäischer
 Dachverband der
 Entwicklungsforschungsinstitutionen
 EADI Portal 95, 102
 ECDPM *Siehe* European Centre for
 Development Policy Management
 Eigenfactor 257
 EL@ND 95, 98, 100
 eLearning 81, 83
 Entwicklungsforschung 95, 97, 99,
 101
 Entwicklungszusammenarbeit 95, 96,
 99, 103
 Erd- und Umweltwissenschaft 164

Erfolgskontrolle *Siehe* Controlling
Erlanger Buchwissenschaft 135, 137, 140
Erstanwender 136
eSciDoc 48, 49, 50, 52, 53, 54, 55, 56
eScience 25, 26, 28, 33, 47
Euforic 97, 100
Euforic.org 95
Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen 193
Europäischer Dachverband der Entwicklungsforschungsinstitutionen 95, 97, 98, 100, 101, 103
European Centre for Development Policy Management 99, 100
European Consensus on Development 96
European Libraries Network for Development *Siehe* EL@ND
Europe's Forum on International Cooperation *Siehe* Euforic
Evaluation 209, 212, 213, 214, 215, 217, 253, 255, 274, 277, 280
Experteninterview 135, 140

F

Fachinformation 109, 110, 111, 120
Fachkommunikation
 historische 71, 73, 74
 wissenschaftliche 78, 79
Fachportal 118
Feldversuch 135, 138, 140, 142
FIZ Karlsruhe 48, 56
Flickr 126
Focuss.eu 95, 99
Folksonomy 38
Footprint Analyse 247
Forschungsplan 101
Forschungspraxis 72
Forschungsprognose 264

G

Geisteswissenschaft 71, 193, 201, 202
Geisteswissenschaften 77
Geodateninfrastruktur 154
GeoForschungsZentrum Potsdam 164

GEOSS 154
Geowissenschaft 147, 148, 159
Geschichtsschreibtisch
 virtueller 75
Geschichtswissenschaft 71, 72, 74
Gesellschaftswissenschaft 201
Global Earth Observing System of Systems *Siehe* GEOSS
Global Spatial Data Infrastructures *Siehe* GSDI
Großforschungseinrichtung 273, 274, 275
Großunternehmen 136
GSDI 154

H

Hardware-Transferprozess 139
Helmholtz-Gemeinschaft 273, 274, 276, 279, 280
h-index 215, 216, 257
Historiker 72, 73, 75
HistoryGuide 73
Hochenergiephysik 177, 178, 179, 215
Hochschulbibliothek 123, 127, 132
Hochschule für Technik und Wirtschaft Chur 86
H-Soz-u-Kult 73
Humboldt Universität 193

I

ICSU *Siehe* International Council for Scientific Unions
Impact Factor 210, 223, 226, 228, 232, 255
Indexierung
 automatische 123
Indexsystem 130
Indikator 209, 211, 215, 216, 217, 247
Indikatorensystem 212, 213
Information Retrieval 76, 112
Informationsangebot 113
 personalisiertes 115
Informationsautonomie 111
Informationsbeschaffung 114
Informationsflut 25, 123
Informationsmanagement 110
Informationsversorgung 107, 109, 110, 121

Innovation 135, 136
 Innovationsprozess 139
 Insecurityforum.org 98
 INSPIRE 156
 Institut Arbeit und Technik
 Gelsenkirchen 202
 Institute of Development Studies 97
 Institute of Social Studies 97
 International Council for Scientific
 Unions 148, 156
 International DOI Foundation 162
 Interoperabilität 188, 189
 ISI Citation Database 253, 254
 ISI Web of Knowledge 257, 258
 ISI Web of Science 259

J

JAL *Siehe* Journal Application Level
 Japan Science and Technology
 Information Aggregator *Siehe*
 JStage
 JCR *Siehe* Journal Citation Reports
 Journal Application Level 247
 Journal Citation Reports 223, 226,
 227, 228, 229, 231, 233
 Journal Eigenfactor 232
 Journal Impact Factor 257, 258
 Journal Use Reports 258
 J-STAGE 223, 227

K

Kennzahl 276, 277, 279, *Siehe auch*
 Indikator
 Kennzahlen 273
 Klassifikation 126, 129, 189
 KMU 135, 136, 137, 139
 Knowledge & Data Engineering Group
 der Universität Kassel 124
 Knowledge Intermediaries 102
 Knowledge-enhancing Helix 81, 83,
 84, 89, 91
 Knowledge-Enhancing Helix 85, 87
 Kollaboration 82
 Kommunikation *Siehe auch*
 Fachkommunikation
 organisationelle 115, 120
 Kommunikationsebenen 115
 Kompetenz
 kollaborative 81

Kompetenzatlas 263, 264, 269
 Konferenz-Weblog 98
 Kulturwissenschaft 71

L

Literaturwissenschaft 194

M

Management Tool 258
 Materialwissenschaft 240, 241, 263,
 264, 265, 267, 269
 Mathematik 215
 Max Planck Digital Library (MPDL) 56
 Max-Planck-Gesellschaft (MPG) 48
 Medienbranche 135
 Medienkompetenz 81, 83
 Medienunternehmen 135
 Medizin 194, 195
 Meeresforschung 150
 Metadaten 150, 151, 153, 155, 156,
 163, 185
 Metadaten-Management 185, 187,
 188, 190
 Metamedium 114
 Mobile IT 135, 137, 139
 MyCoRe 185, 186, 187, 188, 190

N

Nationallizenz 123, 127
 Naturwissenschaft 159, 160, 200, 201,
 235
 Newsfeed 100, *Siehe auch* RSS
 Niedersächsische Staats- und
 Universitätsbibliothek Göttingen 73

O

OECD Development Centre 99
 Ontologie 33, 39, 40, 41, 63
 der Wissenschaftsdisziplinen 61,
 62, 67, 68
 Ontoverse 41
 OntoWiki 41
 Open Access 179, 184, 194, 197, 198
 Open Access Publishing 177
 Open Access-Geschäftsmodell 180,
 195

Open Access-Journals *Siehe* Open Access-Zeitschrift
Open Access-Publikation 180, 182, 195
Open Access-Studie 198
Open Access-Zeitschrift 193, 195, 199, 202, 223, 225, 228, 233
Overseas Development Institute 97, 99

P

PANGAEA® 147, 149, 151, 153, 154, 155
Parallelmedium 114, 117, 119
Parzival-Projekt 28
Peer Review 178, 277
Praxisbeispiel 95
Preprints 177, 178
Primärdaten 200
 wissenschaftliche 147, 148, 151, 154, 159, 160, 161, 166
Pritchard 236
PROBADO 169, 170, 172, 173, 174, 175
Produkt-Lebenszyklus 136
Prozessverbesserung 135, 137, 142
Publikation 148, 151, 161
Publikationsagent 164
Publikationsaktivität 236, 240, 268, 269
Publikationslandschaft 178
Publikationssystem 153
Publishing Network for Geoscientific and Environmental Data *Siehe* PANGAEA®

Q

Qualitätskontrolle 129

R

Recherchesystem 127, 128
Rechtfertigung
 gesellschaftliche 273
Regensburger Verbundklassifikation 64, 130, 131, 132
Relevanz 263, 264, 266, 274, 278, 279

Repository 171, 175, 177, 193, *Siehe auch* Datenzentrum
Reputation 279
Research in Motion 136
Royal Tropical Institute 99
RSS 97, 99
RSWK 131
RVK *Siehe* Regensburger Verbundklassifikation

S

Sacherschließung 123, 124, 126, 132
Schlagwortnormdatei 129, 131
Schlüsselkompetenz 82, 83, 84, 86
Schlüsselqualifikation *Siehe* Schlüsselkompetenz
Scholarly Semantic Web 30
SciELO 223, 225, 227
Science Citation Index 210, 235, 236, 253, 265
Scientific Electronic Library Online *Siehe* SciELO
SCOAP 177, 178, 179, 180, 182, 183, 184
Second-Product-Lifecycle 135, 136, 138
Semantic Web 39, 77
Serviceverbesserung 135, 137, 142
SMART-Projekt 240, 241, 263, 264, 265, 268, 269
Social bookmarking 100
Social Navigation 38
Social Network 35, 37
Social Sciences Citation Index 232
Social Software 35, 82, 126
Social Tagging 38, *Siehe auch* Tagging
Social Web 78
Software-Applikation 139
Soziales Netz *Siehe* Social Network
Sozialwissenschaft 127, 194, 202
Sponsoring Consortium for Open Access Publishing in Particle Physics *Siehe* SCOAP
Stiftung Alfred-Wegener-Institut für Polar- und Meeresforschung
 Bremerhaven 149
Suchetechnik 112
Such-Infrastruktur 116
Suchmaschine 77, 170

fachbezogene 99
 Suchtechnik 112
 Suchtechnologien 29
 SWD *Siehe* Schlagwortnormdatei
 Systemarchitektur 169, 172, 173, 174
 Systematik 65, 66
 SyynX 127

T

Tagging 123, 128, *Siehe auch* Social
 Tagging
 collaboratives 123, 124, 125, 126
 Taggingssystem 125
 TECH-CYCLE 135
 Technische Informationsbibliothek
 Hannover 148, 159, 162, 166, 176
 Technische Universität Graz 176
 Technische Universitätsbibliothek
 Hannover 153
 Technologiebeobachtung 238
 Technologiefrüherkennung 263, 264
 Teilchenphysik 178
 Thomson Scientific 253, 257, 258
 TIB *Siehe* Technische
 Informationsbibliothek Hannover
 T-Mobile 136
 Topic Map 81, 83
 Transfermodalität 136
 Trend 237, 238
 Trenderkennungssystem 235, 239,
 243

U

UDK *Siehe* Universale
 Dezimalklassifikation
 Umweltforschung 150
 Universale Dezimalklassifikation 64
 Universität Bonn 175
 Universität Erlangen-Nürnberg 135,
 136
 Universität Lund 194, 223
 Universität München 194
 Universität Oldenburg 176
 Universitätsbibliothek Lund 230, 233
 Universitätsbibliothek Mannheim 123,
 127, 128, 132
 University of Arkansas at Little Rock
 194
 Unternehmen

kleine und mittlere *Siehe* KMU

V

vascoda 65
 Verlag 141
 Verlagssystem 147, 148
 virtual H-Desk 79
 Volltextserver 123, 127
 Vorgehensmodell 138, 140

W

Wandel
 digitale 71
 WDC *Siehe* World Data Center
 WDC-MARE *Siehe* World Data Center
 for Marine Environmental Sciences
 Web2.0 33, 35, 38, 78, 82, 95, 97,
 100, 101, 103, 107
 Wiki 81, 83, 86
 WIKINGER 41
 Wikipedia 41, 126
 Wirtschaftsinformatik 194
 Wirtschaftswissenschaft 127
 WissdeX Fullerene 239
 Wissen
 bildliches 11, 15, 16, 18
 explizites 11, 12, 16, 17, 21
 implizites 11, 13, 16, 17, 20
 Wissenschaftsdisziplin 66
 Wissenschaftskommunikation 7, 81,
 82, 86, 91, 95, 98, 109, *Siehe auch*
 Fachkommunikation
 Wissensmanagement 98, 123
 Wissens-Transferprozess 139
 Wissensvision 26
 World Data Center 148, 151, 152, 156
 World Data Center for Marine
 Environmental Sciences 147, 148,
 150, 154

X

XING 37

Z

Zeitschrift 179, 180
elektronische 123
Zeitschriftenverlag 178

Zentralbibliothek des
Forschungszentrums Jülich 265
Zentrum für Marine
Umweltwissenschaften der
Universität Bremen 149
Zitation 214, 216, 253, 254, 256

Sponsoren

Wir danken unseren Sponsoren!

Ex Libris

Vom **Bibliothekssystem**
zum **Informationsservice**

Ihre Vorteile:

- ✓ Schnelles und komfortables Finden
- ✓ Best in Class eResourcemanagement
- ✓ Library 2.0 mit Primo
- ✓ Umfangreiches Lösungsportfolio im Markt
- ✓ Optimale Integration

Ex Libris (Deutschland) GmbH

Gasstraße 18, Haus 2 · D-22761 Hamburg

Tel. +49 (0)40 / 89 809-0 · Fax +49 (0)40 / 89 809-250

E-Mail: info@exlibrisgroup.com

www.exlibrisgroup.com





MASSMANN
INTERNATIONALE
BUCHHANDLUNG GMBH

Massmann Internationale Buchhandlung
Luruper Chaussee 125
D-22761 Hamburg
Telefon 040/7670 04 0
Telefax 040/7670 04 10
E-Mail info@massmann.de
Internet www.massmann.de

Stéphanie Rübner
»I hold this letter in my hand«
Scherenschnitt, 31x32 cm, 2005

1. **Naturwissenschaft und Technik – nur für Männer?
Frauen mischen mit!**
Auswahl-Bibliographie Wissenschaftlerinnen (1999), 28 Seiten
ISBN: 978-3-89336-246-2
4. **Schweißen & Schneiden**
Wissenschaftliche Veröffentlichungen des Forschungszentrums Jülich
(1997), 16 Seiten
ISBN: 978-3-89336-208-0
5. **Verzeichnis der wissenschaftlich-technischen Publikationen**
des Forschungszentrums Jülich
Januar 1993 - Juli 1997 (1997), ca. 100 Seiten
ISBN: 978-3-89336-209-7
6. **Biotechnologie**
Wissenschaftliche Veröffentlichungen der Institute für Biotechnologie
des Forschungszentrums Jülich
Januar 1992 - Juni 1997 (1997), 48 Seiten
ISBN: 978-3-89336-210-3
7. **Verzeichnis der wissenschaftlich-technischen Publikationen**
des Forschungszentrums Jülich
1997 bis 1999 (2000), 52 Seiten
ISBN: 978-3-89336-260-8
8. **Kompodium Information**
Teil I: Archive, Bibliotheken, Informations- und Dokumentationseinrichtungen
Teil II: Ausbildungsstätten, Fort- und Weiterbildungsaktivitäten, Informations-
dienste, Presse- und Nachrichtenagenturen, Verlagswesen und Buchhandel,
Einrichtungen des Patent- und Normungswesen, Publikationen
von G. Steuer (2001), 1130 Seiten
ISBN: 978-3-89336-286-8
9. **Die Zukunft des wissenschaftlichen Publizierens**
Der Wissenschaftler im Dialog mit Verlag und Bibliothek
Jülich, 28. bis 30. November 2001. 40 Jahre Zentralbibliothek. Konferenz und
Firmenausstellung
Tagungsprogramm und Kurzfassungen (2001), 50 Seiten
ISBN: 978-3-89336-292-9
10. **Die Zukunft des wissenschaftlichen Publizierens**
Der Wissenschaftler im Dialog mit Verlag und Bibliothek
Jülich, 28. - 30.11.2001. Tagungsprogramm und Vorträge (2002), 184 Seiten
ISBN: 978-3-89336-294-3 (broschiert)
ISBN: 978-3-89336-295-0 (CD)

11. **Bibliometric Analysis in Science and Research**
Applications, Benefits and Limitations
2nd Conference of the Central Library, 5 – 7 November 2003, Jülich, Germany
Conference Proceedings (2003), 242 pages
ISBN: 978-3-89336-334-6
12. **Bibliometrische Analysen – Daten, Fakten und Methoden**
Grundwissen Bibliometrie für Wissenschaftler, Wissenschaftsmanager,
Forschungseinrichtungen und Hochschulen
von R. Ball, D. Tunger (2005), 81 Seiten
ISBN: 978-3-89336-383-4
13. **VIRUS – Sicher im Netz?**
2. Internationale Konferenz zur Virtuellen Bibliothek des Goethe-Instituts
Brüssel
herausgegeben von R. Ball, C. Röpke, W. Vanderpijpen (2005), 137 Seiten mit
beiliegender CD-ROM
ISBN: 978-3-89336-377-3
14. **Knowledge eXtended**
Die Kooperation von Wissenschaftlern, Bibliothekaren und IT-Spezialisten
3. Konferenz der Zentralbibliothek, 2. – 4. November 2005 Jülich
Vorträge und Poster (2005), 392 Seiten
ISBN: 978-3-89336-409-1
15. **Qualität und Quantität wissenschaftlicher Veröffentlichungen**
Bibliometrische Aspekte der Wissenschaftskommunikation
von M. Jokić, R. Ball (2006), 186 Seiten
ISBN: 978-3-89336-431-2
16. **Bibliotheken in Singapur**
von B. Mittermaier (2006), 114 Seiten
ISBN: 978-3-89336-449-7
17. **Libraries in Singapore**
by B. Mittermaier (2007), 112 pages
ISBN: 978-3-89336-450-3
18. **WissKom 2007: Wissenschaftskommunikation der Zukunft**
4. Konferenz der Zentralbibliothek, Forschungszentrum Jülich,
6. - 8. November 2007, Beiträge und Poster
herausgegeben von R. Ball (2007), 300 Seiten
ISBN: 978-3-89336-459-6

Sponsoren der WissKom2007:



Dietmar Dreier Wissenschaftliche Versandbuchhandlung GmbH

EBSCO Information Services GmbH

Elsevier Science GmbH

ExLibris (Deutschland) GmbH

FIZ Karlsruhe

Harrassowitz Booksellers & Subscription Agents

MASSMANN – Internationale Buchhandlung GmbH

Mayersche Buchhandlung GmbH & Co. KG

MIRA Consulting GmbH

Wolters Kluwer Medical Research

Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

