

Institutionelle und fachliche Repositorien

Steigern Repositorien die Nutzung und den Impact von Publikationen?

Ulrich Herb

Saarländische Universitäts- und Landesbibliothek (SULB)



Institutionelle und fachliche Repositorien



Themen

- OA-Aktivitäten der SULB als Betreiberin eines institutionellen und eines disziplinären Repositories
- Repository-Services
- Zusammenspiel institutioneller und disziplinärer Repositories

Die SULB - Aufgabenspektrum

- Universitätsbibliothek der Universität des Saarlandes (UdS)
- Sondersammelgebietsbibliothek für die Psychologie
- Landesbibliothek des Saarlandes

Die Repositories der SULB bis 2003

- **SOVA** – Saarbrücker Online Volltext Archiv
Beginn der Entwicklung: 1999
Ein Server für alle Aufgaben?
- DPI – Digitale Psychologie Information
Gemeinschaftsprojekt mit dem Zentrum für Psychologische Information und Dokumentation (ZPID), gefördert von DFG (2002 bis 2005)
u.a. Einrichtung des Psychologie-Volltextservers
PsyDok als separates Angebot für die psychologische Community

2003

Neustrukturierung der Server-Infrastruktur: Ein eigenes Repository pro Funktion

SciDok – Service für die Universität
Institutional Repository der Universität des Saarlandes UdS
<http://scidok.sulb.uni-saarland.de>

PsyDok – Service des SSG
Disciplinary Repository für die Psychologie
<http://psydok.sulb.uni-saarland.de>

SaarDok – elektronische Pflichtexemplare
Archivserver für landesspezifische Schriften
<http://saardok.bsz-bw.de/>



Die OA-Repositories

SciDok und PsyDok stellen eine technisch zeitgemäß ausgestattete Infrastruktur zum Self-Archiving dar (OPUS-basiert, mit DINI-Zertifikat).

Diese Angebote sollen weiterentwickelt und ergänzt werden, aber:

Wichtiger als technische Voraussetzungen ...

... ist die Schaffung eines OA-freundlichen Milieus durch

1. Politische Maßnahmen
-> UdS und SSG
2. Öffentlichkeitsarbeit
3. Services der SULB
4. Ausbau bestehender Angebote

IR: Politische Maßnahmen

- Schaffung institutioneller Rahmenbedingungen zur Verankerung von Open Access auf dem Campus

Initiativen der SULB an der UoS:

Eingaben beim/ Diskussionen mit dem Präsidium über ...

- Verabschieden einer Open Access Policy
- Unterzeichnung der Berlin Declaration
- Installation eines OA-Lenkungsgremiums

IR: Content-Akquise

1) Fortlaufend: Durchführen von Informationsveranstaltungen für Fachrichtungen bzw. Lehrstühle/ Tingen

- Aufhänger: Elektronisches Publizieren von Dissertationen, Elektronischer Semesterapparat, neugestaltete Jahresbibliographie, Informationen über neue Datenbanken durch zuständige Fachreferentin/ zuständigen Fachreferenten

2) Verzahnung von SciDok und Jahresbibliographie

<http://jahrbib.sulb.uni-saarland.de>

- Beide Systeme laufen unter OPUS
- Die Meldung in der Jahresbibliographie und die Veröffentlichung in SciDok sind in einem Arbeitsschritt möglich
- SULB-Mitarbeiter können Dokumente zu den Metadaten in der Jahresbibliographie nachträglich in SciDok einspielen

IR: Zwischenbilanz

- Server ist am Campus als OA-Server bekannt
- Verzahnung mit Jahresbibliographie wirkt sich positiv aus
- Noch kein starkes Statement der Universitätsleitung zu Open Access (OA-Policy, OA-Lenkungsgremium)
- Absprache über Einstellen von Preprints und Postprints mit einzelnen Lehrstühlen gelungen
- Attraktivität des Servers korreliert mit Einbindung in (fachspezifische) Datenbanken & Suchmaschinen: Web Citation Index (Thomson Scientific), IO-Port, CiteSeer, Google Scholar, Scirus...

DR: Politische Maßnahmen

Rahmenbedingungen sind wesentlich von der Deutschen Gesellschaft für Psychologie **DGPs** bestimmt.

- Kooperation mit dem Zentrum für Psychologische Information und Dokumentation (ZPID)
- 28.09.2004: Mitgliederversammlung der DGPs stimmt für die Unterzeichnung der Berliner Erklärung
- wünschenswert: Abstimmung zwischen SULB und DGPs über Open Access
- wünschenswert: Empfehlung von PsyDok als Repository für die Psychologie
- Erwähnung von PsyDok in den „[Informationen und Empfehlungen beim aktuellen UrhG](#)“

DR: Content-Akquise

- Wiederholtes Anschreiben aller Institute, Fachbereiche, Lehrstühle
- Kontaktaufnahme zur Deutschen Gesellschaft für Psychologie (DGPs)
- Anschreiben von Prüfungsämtern
- Anschreiben von Fachschaften
- Anschreiben von Produzenten grauer Literatur
- Anschreiben DFG-geförderter Projekte
- Teilnahme an Fachtagungen

DR: Spezielle Hindernisse

Das Problem:

Publisher:	Hogrefe & Huber
Pre-print:	✗ author cannot archive pre-print (ie pre-refereeing)
Post-print:	✓ author can archive post-print (ie final draft post-refereeing)
Conditions:	<ul style="list-style-type: none">• On author's or employers web site only after publication• Publisher version cannot be used• Publisher copyright must be acknowledged with set phrase• Must link to publisher's journal page• Set phrase must appear "This article does not exactly replicate the final version published in the journal ?[Add title of Journal]?. It is not a copy of the original published article and is not suitable for citation."• On personal Webpage or employers Website but not "in any repositories other than the publisher's own"
Mandated OA:	✗ Wellcome Trust (Non-compliant); ✗ ESRC (Non-compliant); ✗ HBM (Non-compliant); ✗ NIH (Non-compliant)
Copyright:	Not available online
RoMEO:	This is a RoMEO blue publisher

- Defensive Open-Access-Policies
- Teils nur Self-Archiving auf Institutional Repositories zulässig

DR: Zwischenbilanz

- Verlage sperren disciplinary repositories meist über SHERPA aus
- Server ist in der Community etabliert
- Nach Erwähnung von PsyDok in den „Informationen und Empfehlungen der IuK-Kommission“ der DGPs verstärkt Postprint-Veröffentlichungen
- Wird z.T. als Plattform für Primärveröffentlichungen genutzt
- Autoren wünschen Aufwertung von PsyDok mit Verlagsfunktionen
- Attraktivität des Servers durch Einbindung in **fachspezifische** und andere anerkannte Datenbanken: Web Citation Index (Thomson Scientific), Current Web Contents (Thomson Scientific), Scirus/Scopus, Psyndex, PsychLinker, PsychSpider, Google Scholar ...

IR & DR: Services

- Autorenunterstützung per Mail und Telefon
- Einspielen von Dokumenten durch SULB, Prüfung der Metadaten
- Erstellen von Portalen
- Aktive Recherche nach Dokumenten und Dokumentsammlungen/ Reihen
- Dokumente auf Seiten der Wissenschaftler mit SHERPA-Liste abgleichen und einspielen durch SULB (Zustimmung der Wissenschaftler vorausgesetzt), UdS: Einträge der Jahresbibliographie mit SHERPA abgleichen
- Steigern der Sichtbarkeit durch fortlaufende Integration in Suchmaschinen, Datenbanken, Nachweissysteme
- Möglichkeit der Druckausgabe im Print on Demand

IR+DR: ein Zusammenspiel

Lessons learned

- Nutzer beider Repositories/ Autoren wollen in Ihrem Fachkontext sichtbar sein
- DR genießt höheren Stellenwert bei Wissenschaftlern: Fachausstatter vs. Gemischtwarenhandel
- DRs können prinzipiell auch Bibliographiefunktionen übernehmen
- PsyDok: Metadaten-Harvesting psychologischer Dokumente verteilter IRs und Nachweis der Daten in PsyDok, Pilotpartner: KOPS der UB Konstanz

IR+DR: ein Zusammenspiel

Perspektive des IR:

- Absprachen mit disziplinären Datenbank Anbietern und Suchmaschinen entfallen, Dokumente aus IRs gelangen über PsyDok-Schnittstelle in Psyndex, Psychspider, ggf. PsychLinker, Psychcrawler
- Werbeargument bei Contentakquise: Dokumente der IRs werden im Fachkontext sichtbar und rezipierbar

Perspektive des DR:

- Umständliche Contentakquise via Post und Mail entfällt
- Dokumente, deren Volltext nach SHERPA-Vorgaben nicht auf einem DR vorgehalten werden dürfen, können dennoch nachgewiesen werden.

IR+DR: ein Zusammenspiel

Technik:

- Harvesting basiert auf OAI-PMH
- Dublettenkontrolle erfolgt über Identifier (URN, ggf. DOI)
- Geänderte Metadaten im IR werden auch im DR aktualisiert
- Bedingung: IR muss Zitierfähigkeit der Dokument garantieren, DINI-Zertifizierung

Perspektiven

- PsyDok

Verstärktes Metadaten-Harvesting

Feintuning der Schnittstellen

Metadatenaustausch auch mit anderen DRs?

- SciDok, PsyDok: Wissenschaftler wünschen sich Aufwertung der OA-Angebote mit Verlagsmerkmalen

-> UdS wird OA-Verlag für UdS und für das Fach Psychologie gründen

Fragen?

Anregungen?



Kommentare?



Steigern Repositorien die Nutzung und den Impact von Publikationen?

Themen

- Stellenwert der Impact Maße
- Impact Maße: Die Referenzen
- Traditionelle Impact Maße und Self-Archiving
- Sind Self-Archiving-spezifische Impact Maße modellierbar?
- Mögliche Standards

Impact und Karrieren

Karrierechancen der Wissenschaftler hängen von ihrem Publikationsverhalten ab.

Reputation der Wissenschaftler wird meist gleichgesetzt mit der Qualität der Journals, in denen sie publizieren.

Qualität der Journals wird gleichgesetzt mit deren Journal Impact Factor (JIF).

Je höher der JIF der Journals ist, in denen publiziert wird, desto größer die Karrierechancen.

Impact Maße: Der Platzhirsch, Journal Impact Factor JIF

Zahl der Zitate im laufenden Jahr auf Artikel der vergangenen zwei Jahre

Zahl der Artikel der vergangenen zwei Jahre



JIF revisited

Vernachlässigung kontextueller Aspekte des Impacts (Gewichtung):
Zitationen aus Zeitschriften mit hohem IF erhalten kein höheres Gewicht
als Zitationen aus Zeitschriften mit niedrigem Ranking

Begrenzter Scope: ausgeschlossen sind komplette Dokumentarten wie
graue Literatur und Bücher, ein Großteil der Web-Publikationen.
Berücksichtigt werden im Journal Citation Report JCR indizierte Journals.

Messung bezieht sich auf Journale, nicht Artikel: I.d.R. führt eine geringe
Anzahl sehr häufig zitierter Artikel zu einem hohen IF für das Journal

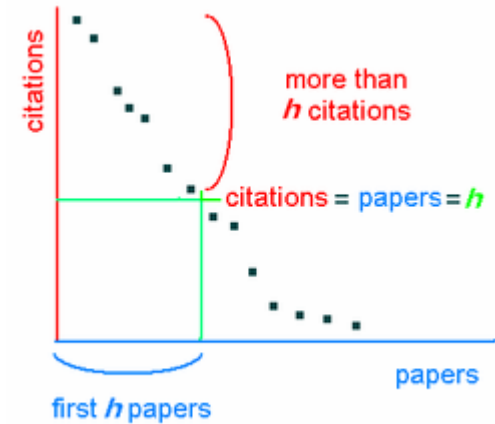
Sprachbias: Englischsprachige Journals werden bevorzugt. Zeitschriften in
anderen Sprachen haben einen niedrigeren IF, da sie im Sample
unterrepräsentiert sind.

JIF revisited (continued)

- Keine Berücksichtigung der unterschiedlichen Verwertungszyklen in unterschiedlichen Disziplinen: in sich rasch entfaltenden Wissenschaften werden verstärkt kürzlich publizierte Texte zitiert. Journals dieser Disziplinen erhalten einen hohen IF. Die Konstruktion des IF benachteiligt Texte aus Disziplinen (z.B. Mathematik, Geisteswissenschaften) mit längeren Verwertungszyklen.
- Der JIF erlaubt die Impact Messung der Publikation erst durch die folgende Auswertung der folgenden Publikationsgeneration.

Impact Maße: Der echte Hirsch (Index)

- Hirsch (h-) Index bezieht sich auf Autoren und nicht auf Journals
- Berechnung: Ein Autor hat einen Index h , wenn h von seinen insgesamt N Veröffentlichungen mindestens jeweils h Zitierungen haben und die anderen $(N-h)$ Publikationen weniger als h Zitierungen.



<http://de.wikipedia.org/wiki/H-Index>

Hat ein Forscher 12 Artikel publiziert, die 12 mal zitiert wurden, hat er einen h-Index von 12; wurde der am wenigsten zitierte Artikel nur neun mal zitiert und die restlichen elf 12 mal, hat er einen h-Index von neun.

H-Index revisited

Der h-Index teilt einige Mängel mit dem JIF:

- Vernachlässigung ganzer Dokumentengattungen
- Vernachlässigung nicht-englischer Publikationen
- Ignoranz des Mehrautorenproblems
- Trennscharfe Autorenidentifikation in der Datenbasis (Web of Science) ist nicht sichergestellt
- Misst der h-Index nicht eher Popularität als Qualität?

Vorteil gegenüber JIF: Zitationen einer einzigen, vielzitierten Veröffentlichung schlagen sich nicht nieder.

Nachteil: innovative Ansätze werden nicht berücksichtigt.

Traditionelle Impact Maße und Self-Archiving

Traditionelle, zitationsbasierte Impact Maße liefern Argumente pro Open Access:

Open-Access-Dokumente werden im Vergleich zu lizenzpflichtigen signifikant häufiger heruntergeladen, genutzt und zitiert (Lawrence 2001, Brody, Harnad & Carr 2004, Brody & Harnad 2005, Sietmann 2006)

Allerdings: Impact Messung des IR/Self-Archiving erfolgt meist mittelbar über Zitationsdatenbanken und mit Bezug zum Journal

Traditionelle Impact Maße und Self-Archiving

- JIF ist *buggy*, zentrales Manko: begrenzter Scope
- Elektrifizierung der wissenschaftlichen Fachinformation und des wissenschaftlichen Publizierens ermöglichen neue Techniken
- OA-Publikationen waren lange vom JIF ausgeschlossen und werden es tendenziell bleiben
 - Self-Publishing wegen fehlender Zitationshistorie
 - Self-Archiving qua Scope des Journal Citation Report JCR
- Zusätzlich: Forderung nach Open Metrics (Suber 2007)

Modellierung eines eigenen Metric-Modells?

Wie lässt sich Impact messen?

Zitationsbasierte Maße

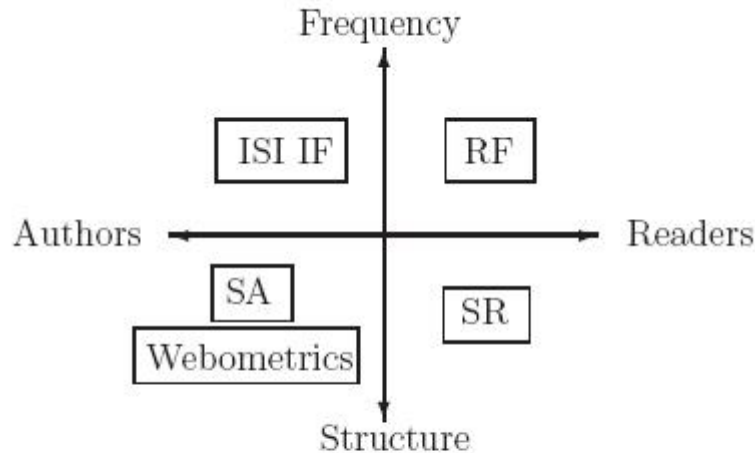
- autorenzentriert
- Impact-Messung erst in den nachfolgenden Publikationsgenerationen möglich

Nutzungsbasierte Maße

- leserzentriert
- Impact-Messung prinzipiell live möglich

Kategorisierung

Zitationsbasierte und Nutzungsbasierte Maße



Aus: Bollen et al. 2005

ISI IF = Journal Impact Factor

RF = Reading Factor

SA = Structure Author

- basiert auf Netzwerk, das durch Autorenhandlung gebildet wird

- Bsp: Google PageRank, Zitationsgraphen, Webometrics

SR = Structure Reader

- basiert auf Kontextinformationen der Dokumentnutzung

- Recommendersysteme

- Bsp: Downloadgraphen

Standardisierung?



Mögliche Ansätze

COUNTER

Dokumentnutzung auf aggregierter Ebene (Journals)

- Scope: Alle Verlagsjournals, die den COUNTER-Standard anwenden
<http://www.projectcounter.org/articles.html>, TA-Journals
- Messung der Nutzungshäufigkeit als Rentabilitätsprüfung
- Vollkommen automatisiertes Verfahren
- Interdisziplinär, keine Verzerrung durch intellektuelle Auswahl
- Problematisch: Keine Aufschlüsselung auf Dokumentenebene, nur für TA-Journals existent, Definition der Timeframes für Doubleclicks, keine Elimination maschineller Zugriffe

<http://www.projectcounter.org/>

Mögliche Ansätze

LogEC

Dokumentnutzung auf Artikelebene

- Statistikmodul des wirtschaftswissenschaftlichen Netzwerks RepEC
<http://repec.org/>
- Messung der Nutzungshäufigkeit als Impact Bestimmung
- Vollkommen automatisiertes Verfahren
- Maschinelle Zugriffe werden eliminiert
- Problematisch: Definition der Timeframes für Doubleclicks, keine Kontextinformationen

<http://logec.repec.org/>

Mögliche Ansätze

International Federation of Audit Bureaux of Circulations (IFABC)

Nutzung auf Artikelebene

- Grundlage des Zählpixel-Mechanismus der VG Wort
- Messung der Nutzungshäufigkeit zur Bestimmung der Vergütungshöhe
- Vollkommen automatisiertes Verfahren
- Maschinelle Zugriffe werden eliminiert
- Problematisch: Definition der Timeframes für Doubleclicks, keine Kontextinformationen

<http://www.ifabc.org/>

Mögliche Ansätze

MESUR: MEtrics from Scholarly Usage of Resources

Nutzung auf Artikelebene

- Modellprojekt des Los Alamos National Laboratory
- Messung der Nutzungshäufigkeit und -struktur zur Ermittlung szientometrischer Kennziffern
- Rankings auf Basis von Kontextinformationen und Verwandtschaftsbeziehungen zwischen Dokumenten, ermittelt über Nutzung/ Zitation
- Technik: Soziale Netzwerkanalyse, automatisiertes Verfahren
- Problematisch: Maschinelle Zugriffe werden nicht eliminiert, Projekt nutzt Linkresolverdaten und ideale experimentelle Bedingungen

<http://www.mesur.org/MESUR.html>

Fazit

Repository Metrics?

- Repositories fördern den Impact für Publikationen in Form klassischer Zitationsmaße
- Repositories können eigene Impact Maße anbieten
- Aber: bislang keine Standardisierung
- Erste Ansätze: Verwendung COUNTER-kompatibler Nutzungsdaten als Empfehlung im DINI-Zertifikat 2007
- Vielversprechende, teils komplexe Modelle, v.a. im MeSUR-Vorgänger
- Aufwändige Infrastruktur zur Herstellung und zum Austausch interoperabler Daten zwischen Repositories erforderlich
- Meta-Open-Access nötig? Sollen Nutzungsdaten unter CC-Lizenzen angeboten werden?

Fragen?

Anregungen?



Kommentare?



Literatur

Bollen, Johan et al. (2005): Toward alternative metrics of journal impact: A comparison of download and citation data. In: Information Processing and Management 41(6): S. 1419-1440. Preprint Online: <http://arxiv.org/abs/cs.DL/0503007>

Brody, Tim and Harnad, Stevan (2004). Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. Dlib Magazine, 10, Nr. 6, <http://www.dlib.org/dlib/june04/harnad/06harnad.html>

Brody, Tim, Harnad, Stevan and Carr, Les (2005). Earlier Web Usage Statistics as Predictors of Later Citation Impact. Journal of the American Association for Information Science and Technology (JASIST). <http://eprints.ecs.soton.ac.uk/10713/>

Lawrence, Steve (2001). Free online availability substantially increases a paper's impact. Nature, 411, S. 521-522 <http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html>

Sietmann, Richard (2006): Über die Ketten der Wissensgesellschaft. In: c't Magazin für Computer und Technik(12): S. 190-199.

Suber, Peter (2007): Why we need OA to citation data. In: Open Access News. News from the open access movement. Online: <http://www.earlham.edu/~peters/fos/2007/12/why-we-need-oa-to-citation-data.html>

Vielen Dank für ihre Aufmerksamkeit

Ulrich Herb

Saarländische Universitäts- und Landesbibliothek
Gebäude B1 1, Zi. 9.08,
D-66123 Saarbrücken
Telefon: 0681 302-2798
u.herb@sulb.uni-saarland.de