

# Impact revisited - citation impact versus usage impact

## *OPEN ACCESS* – Die Digitalisierung der Wissenschaftskommunikation

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# Übersicht

---

- Warum ist Impact so wichtig?
- Klassische Impact-Messung: Die Referenzen
- Klassische Impact-Maße und Open Access
- Sind alternative Impact-Maße nötig und modellierbar?
- Impact-Maße: eine Kategorisierung
- Nutzungsbasierte Impact-Maße: Existieren Standards?
- Kurzdarstellung des Projekts „Open-Access-Statistik“



---

# Warum ist Impact so wichtig?

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# Bedeutung der Impact-Maße: Individuelle Sicht der Wissenschaftler

---

Karrierechancen der Wissenschaftler hängen von ihrer Reputation ab.

Reputation wird in aller Regel über *Impact* zu bestimmen versucht.

Impact wird ermittelt über Zitationen

- a) bezogen auf Zeitschriften, in denen Wissenschaftler publizieren
- b) bezogen auf Artikel eines Wissenschaftlers



# Bedeutung der Impact-Maße: Individuelle Sicht der Wissenschaftler

---

## Zeitschriftenzentrierte Zitationsmaße

- Qualität 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.



# Bedeutung der Impact-Maße: Individuelle Sicht der Wissenschaftler

---

## Autorenzentrierte Zitationsmaße

- Qualität der Wissenschaftler wird meist gleichgesetzt mit der Qualität ihrer Artikel.
- Qualität der Artikel wird gleichgesetzt mit deren Zitationshäufigkeit (Beispiel: Hirsch- oder h-Index).
- Je höher Hirsch-Index eines Autors ist, desto größer die Karrierechancen.



# Bedeutung der Impact-Maße: Die Organisationssicht

---

**Hochschulevaluation** und **Leistungsorientierte Mittelvergabe (LOM)** entscheiden über zukünftige materielle und personelle Ausstattung und durch Entzug oder Erhöhung von Mitteln auch über die Wahl von Forschungsschwerpunkten.

## Kriterien

1. Anzahl der Promotionen
2. Umfang der eingeworbenen Drittmittel
3. **Publikationsverhalten**  
-> bestimmt über Zitationsmaße, v.a. JIF, seltener h-Index



---

# Klassische Impact-Messung: Die Referenzen

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# Impact-Maße: Referenz I – Der Journal Impact Factor JIF

---

## Berechnung

**Zahl der Zitate im laufenden Jahr auf Artikel (eines Journals) der  
vergangenen zwei Jahre**

-----

**Zahl der Artikel des Journals der vergangenen zwei Jahre**



# Impact-Maße: Referenz I – Der Journal Impact Factor JIF

---

## Kritikpunkte Teil 1:

- Begrenzter Scope/ Ausschluss kompletter Dokumentarten: graue Literatur, Bücher, Großteil der Web-Publikationen.
- Berücksichtigt werden nur im Journal Citation Report JCR indizierte Journals.
- Sprachbias zugunsten englischsprachiger Journals: Zeitschriften in anderen Sprachen haben einen niedrigeren IF, da sie im Sample unterrepräsentiert sind.
- JIF bezieht sich auf Journale, nicht Artikel: I.d.R. führt eine geringe Anzahl sehr häufig zitierter Artikel zu einem hohen Wert für das Journal



# Impact-Maße: Referenz I – Der Journal Impact Factor JIF

---

## Kritikpunkte Teil 2:

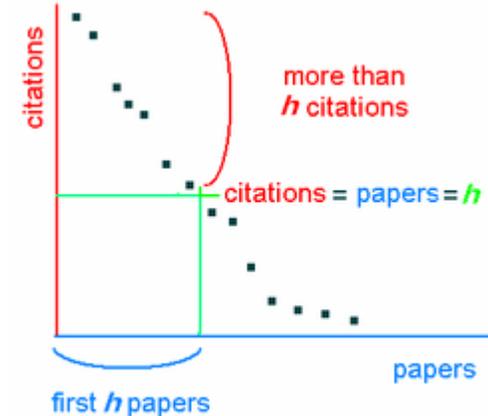
- Ignoranz der Verwertungszyklen in unterschiedlichen Disziplinen: Benachteiligung der Journale aus Disziplinen mit Verwertungszyklen  $> 2$  Jahren (z.B. Mathematik, Geisteswissenschaften)
- Ignoranz des Mehrautorenproblems
- Vernachlässigung kontextueller Aspekte (Gewichtung)
- Misst der JIF nicht eher Popularität als Qualität?



## Impact-Maße: Referenz II – Der Hirsch-Index (h-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>

Ein Autor einen h-Index von 8, wenn er 8 Schriften veröffentlicht hat, die jeweils mindestens 8 Mal zitiert worden sind. Ein h-Index von 12 setzt die Publikation von 12 Schriften voraus, die mindestens 12 Mal zitiert wurden.

# Impact-Maße: Referenz II – Der Hirsch-Index (h-index)

---

## Vergleich h-Index/JIF

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

Nachteil: innovative Ansätze werden nicht berücksichtigt.

## Kritik des h-Index:

- vgl. JIF-Kritik: Vernachlässigung von Dokumentengattungen und nicht-englischer Publikationen, Mehrautorenproblematik, Messung von Popularität oder Qualität? ...
- Trennscharfe Autorenidentifikation in der Datenbasis (Web of Science) ist nicht sichergestellt
- h-Index hängt von Alter des Autors und Disziplin ab



---

# Klassische Impact-Maße und Open-Access

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# Klassische Impact-Maße und Open Access

---

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 2004, Sietmann 2006).

Downloadhäufigkeit scheint mit Zitationshäufigkeit zu korrelieren und diese vorherzusagen (Brody, Harnad & Carr 2005).

Impact Messung des Open-Access-Angebots erfolgt meist mittelbar über Zitationsdatenbanken (Journal Citation Report, Web of Science) und mit Bezug zum Journal.



---

# Sind alternative Impact-Maße nötig und modellierbar?

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposion:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



## Sind alternative Impact-Maße nötig und modellierbar?

---

- JIF und h-Index sind mangelbehaftet, 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 und Sprache
  - Self-Archiving qua Scope des Journal Citation Report JCR
- Forderung nach Open Metrics (Suber 2007)



# Impact-Maße: eine Kategorisierung

---

## *Zitationsbasierte Maße*

- autorenzentriert
- Messung erst in den nachfolgenden Publikationsgenerationen möglich
- Messung erfolgt i.d.R. auf Journal- oder Autorenebene, Impact eines Textes/Objekts wird nicht abgebildet

## *Nutzungsbasierte Maße*

- leserzentriert
- Messung prinzipiell live möglich
- Impact eines Textes/Objekts wird abgebildet
- Stark automatisierte Messung möglich



---

# Impact-Maße: Eine Kategorisierung

---

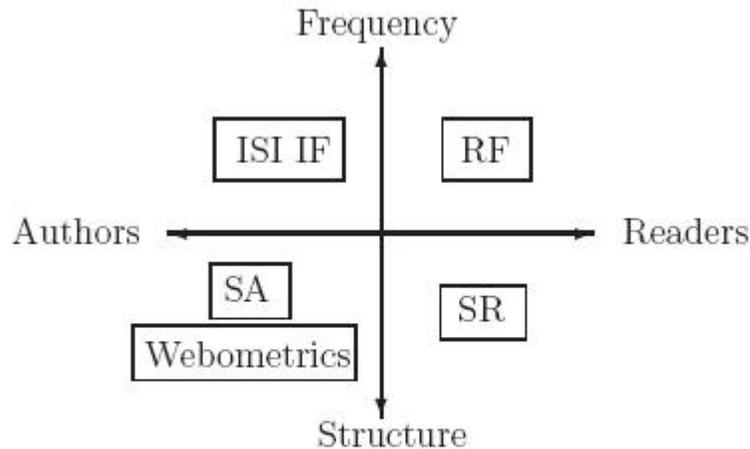
Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# Impact-Maße: eine Kategorisierung

## Zitationsbasierte und nutzungsbasierte Maße



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

Aus: Bollen et al. 2005



# Nutzungsbasierte Impact-Maße: Standards???

---

## Nutzungsbasierte Verfahren: Standardisierung?



# Nutzungsbasierte Impact-Maße: Ansätze

---

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

Dokumentnutzung auf aggregierter Ebene (Journals)

- Scope: Journals, die den COUNTER-Standard anwenden
- Messung der Nutzungshäufigkeit als Rentabilitätsprüfung
- Vollkommen automatisiertes Verfahren
- Interdisziplinär, keine Verzerrung durch intellektuelle Auswahl
- Probleme: Keine Aufschlüsselung auf Dokumentenebene, nur für TA-Journals existent, Definition der Timeframes für Doubleclicks, keine Elimination maschineller Zugriffe



# Nutzungsbasierte Impact-Maße: Standards???

---

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

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



# Nutzungsbasierte Impact-Maße: Standards???

---

**International Federation of Audit Bureaux of Circulations (IFABC),**  
<http://www.ifabc.org/>

Nutzung auf Artikelebene

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



## Nutzungsbasierte Impact-Maße: Standards???

---

**MESUR: MEtrics from Scholarly Usage of Resources,**  
<http://www.mesur.org/MESUR.html>

Nutzung auf Artikelebene

- Modellprojekt: 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



## Nutzungsbasierte Impact-Maße: Fazit

---

- Alternative Impact-Maße (abseits der Zitationsmaße h-Index, JIF) sind modellierbar
- Aber: bislang keine Standardisierung
- Vielversprechende, teils komplexe Modelle, v.a. im MeSUR-Vorgänger
- Aufwändige Infrastruktur zur Herstellung und zum Austausch interoperabler Daten zwischen Servern erforderlich
- Meta-Open-Access nötig? Sollen Nutzungsdaten unter CC-Lizenzen angeboten werden?



---

# Kurzdarstellung des Projekts „Open-Access-Statistik“

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# DFG-Projekt Open-Access-Statistik (OAS): Partner

---

Drittmittelgeber: Deutsche Forschungsgemeinschaft DFG  
<http://www.dfg.de>

Projektpartner:

- Niedersächsische Staats- und Universitätsbibliothek Göttingen
- Universitätsbibliothek Stuttgart
- Saarländische Universitäts- und Landesbibliothek
- Humboldt Universität zu Berlin (Computer- und Medienservice)

<http://www.dini.de/projekte/oa-statistik/>

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Symposium:  
„Open Access - Die Digitalisierung der  
Wissenschaftskommunikation“  
Universität Linz, 05.12.2008



# DFG-Projekt Open-Access-Statistik (OAS): Informationen

---

## Ziel:

Infrastruktur für standardisierte, interoperable Nutzungsstatistiken unter besonderer Berücksichtigung von Open-Access-Angeboten

## Motivation:

- Open-Access-Angebote werden von etablierten Impact-Modellen nicht/selten erfasst -> Impact = Anreiz, der OA attraktiv machen kann
- etablierte, zitationsbasierte Impact Maße weisen Mängel auf
- elektronisches Publizieren ermöglicht es nutzungsbasierte Maße darzustellen
- vielversprechende Evaluierungen nutzungsbasierter Impact-Modelle (Bollen et al. 2005)



# DFG-Projekt Open-Access-Statistik (OAS): Informationen

---

## Skizze:

- Aggregation von Nutzungsdaten verschiedener Open-Access-Angebote/ Server in einem Serviceprovider (v.a. Webserverlogs und Linkresolverlogs)
- Bereinigung von Verzerrungen (maschinelle Zugriffe durch Spider, Dublettenerkennung etc.)
- Aufbereitung der Nutzungsdaten anhand verschiedener Standards
- Rückspielen der ermittelten Nutzungshäufigkeiten in die verteilten Angebote und Verwendung als Metadatum (ggf. Ranking)



## Welche Verwendung können die Daten finden?

- Entwicklung aufbauender Services, z.B. Recommender
- Zusatzservice für wissenschaftliche Suchmaschinen und Zusatzmetadatum für Datenbankanbieter
- Verwendung als szientometrische Information in Szenarien unterschiedlicher Komplexität (reine Häufigkeiten als Vorhersage von Zitationswerten bis hin zum MESUR-Ansatz)
- Entwicklung und Evaluierung unterschiedlicher szientometrischer Verfahren

...



# 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>



# Diskussion

---

Fragen?

Anregungen?



Kommentare?



## Kontaktinformationen

---

Vielen Dank für Ihre Aufmerksamkeit!

**Ulrich Herb**

**Saarländische Universitäts- und Landesbibliothek  
Gebäude B1 1, Zi. 9.08,  
D-66123 Saarbrücken  
Telefon: 0049 681 302-2798  
u.herb@sulb.uni-saarland.de**



## Lizenzinformation

---

Dieses Werk ist unter einem Creative Commons Namensnennung 3.0 Deutschland Lizenzvertrag lizenziert. Um die Lizenz anzusehen, gehen Sie bitte zu <http://creativecommons.org/licenses/by/3.0/de/> oder schicken Sie einen Brief an

Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

