

# Approaches to measuring scholarly impact

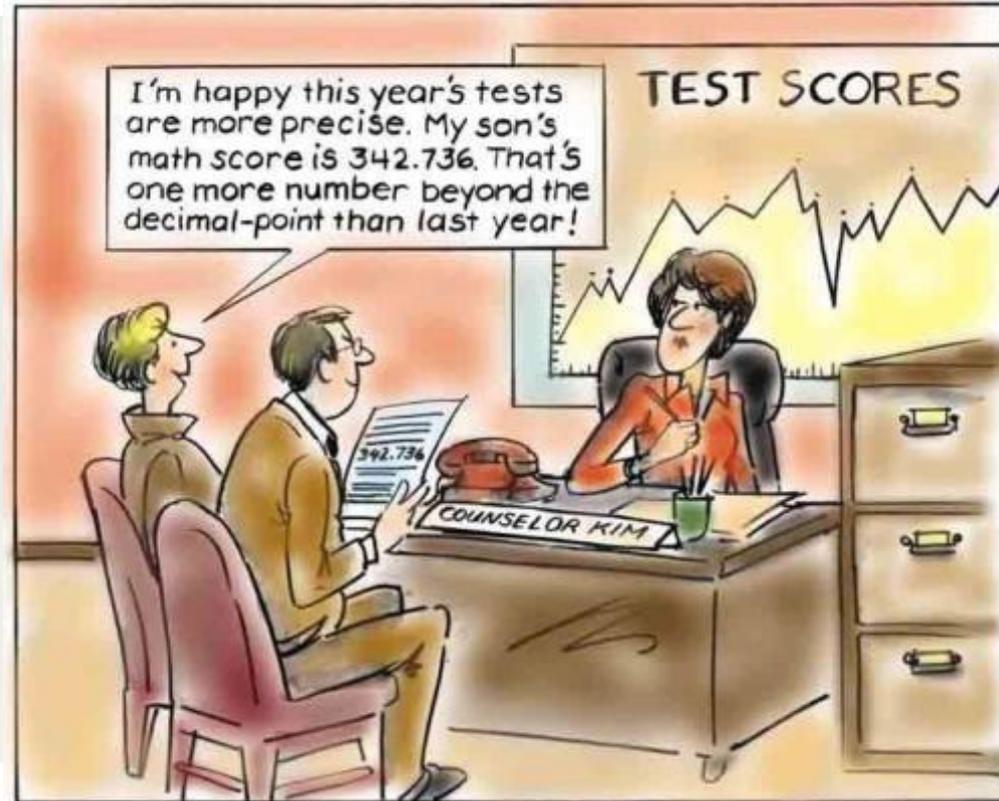
Johan Bollen

School of Informatics and Computing

Indiana University - Bloomington

# Scholarly evaluation

Limited resources: need for assessment? If so what kind of assessment?



# Science as a gift Economy

- Science is a gift-economy:
  - “Steal text, but not ideas” (loosely paraphrasing S. Harnad!)
  - Currency is acknowledgement of scholarly influence
- From print comes citation
  - Citation: ultimate acknowledgement of influence
  - More citations = more influence



# Counting citations

CiteSeer<sup>x</sup><sub>β</sub>

Most Cited Articles

Most Cited Citations

Most Cited Authors

## Most Cited Computer Science Authors

This is generated from documents in the CiteSeer<sup>x</sup> database as of May 20, 2012. An entry may contain errors. Citation counts may differ from search results because this list is generated from a subset of the database.

1. [D. Johnson](#)  
35557
2. [J. Smith](#)  
26367
3. [Y. Wang](#)  
24421
4. [J. Lee](#)  
24366
5. [S. Shenker](#)  
21963
6. [J. Wang](#)  
21937
7. [A. Gupta](#)  
21099
8. [H. Zhang](#)  
21047
9. [L. Zhang](#)  
20867
10. [J. Anderson](#)  
20702

Most Cited Articles

Most Cited Citations

Most Cited Authors

## Most Cited Computer Science Articles

This list is generated from documents in the CiteSeer<sup>x</sup> database as of May 20, 2012. This list is accurate and citation counts may differ from those currently in the CiteSeer<sup>x</sup> database, since the data is generated from a subset of the database.

1. A P Dempster, N M Laird, D B Rubin.  
[Maximum likelihood from incomplete data via the em algorithm](#). Journal of Royal Statistical Society, 5972
2. C A R Hoare.  
[Communicating sequential processes](#). ISSN 0001-0782. URL <http://doi.acm.org/10.1145/359576>, 3056
3. L Rabiner.  
[A tutorial on hidden Markov models and selected applications in speech recognition](#). Proceedings, 3014
4. I Stoica, R Morris, D Liben-Nowell, D Karger, M F Kaashoek, F Dabek, H Balakrishnan.  
[Chord: A scalable peer-to-peer lookup service for Internet applications](#). In SIGCOMM, 2001, 2971
5. J R QUINLAN.  
[Induction of Decision Trees](#). 0, 2845
6. D G Lowe.  
[Distinctive image features from scale-invariant keypoints](#). Int. J. Comput. Vision, 0, 2804
7. S Kirkpatrick, C D Gelatt, M P Vecchi.  
[Optimization by simulated annealing](#). Science, 1983, 2741

# Impact metrics: Entities, relations, and metric types

## **1) What are we evaluating?**

- Authors
- Articles
- Journals
- Domains
- Countries

## **2) On the basis of what kind of acknowledgement?**

- Citations, obviously
- Other expressions of influence, impact, attention, etc?
  - Online usage data?
  - Social Media data?
  - Explicit indicators: bookmarking, etc.

**3) What dimension of impact are we measuring:**  
counting indicators, network structure



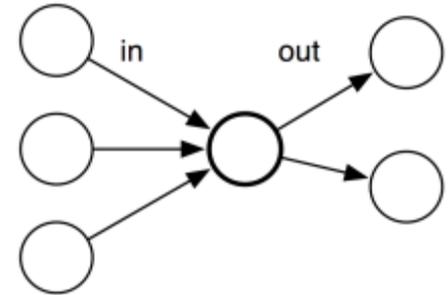
# Possible network metrics

Classes of metrics:

- Degree
- Shortest path
- Random walk
- Distribution

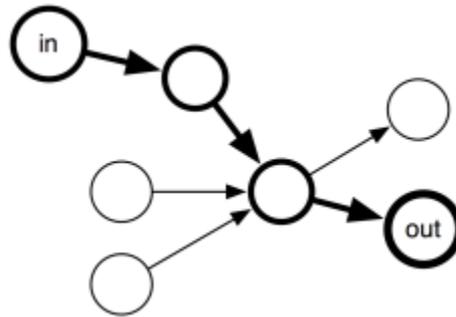
Degree

- In-degree
- Out-degree



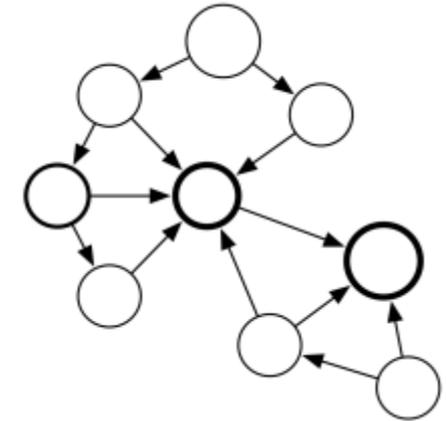
Shortest path

- Closeness
- Betweenness



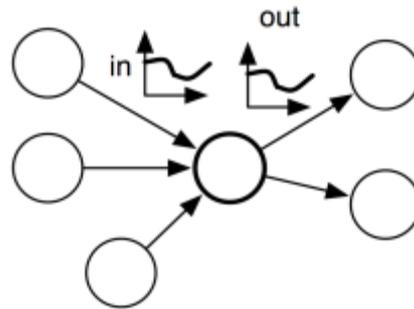
Random walk

- PageRank
- Eigenvector



Distribution

- In-degree entropy
- Out-degree entropy
- Bucket Entropy



Each can be defined to take into account weights by e.g. means of weighted shortest path definition

# A myriad of metrics

Many variations on the basis of what is ranked, how, and on the basis of what data.

Here I will focus on the most commonly used or referred to:

Impact Factor

Eigenfactor

COUNTER (not really a metric)

MESUR (shameless plug?)

PLoS ALM

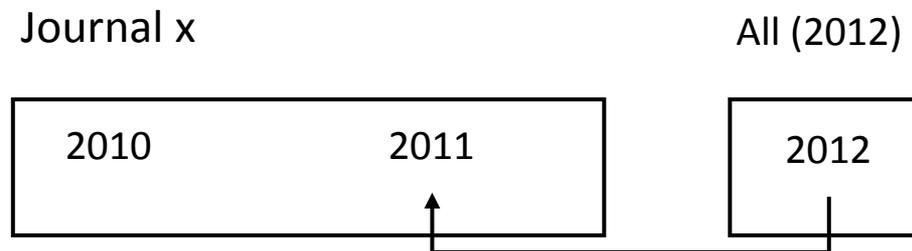
Eigenfactor

H-index

# Journal Impact Factor

## Definition:

Impact Factor of journal  $j$  in year  $t =$  all year  $t$  citations to articles published in  $j$  in  $t-1$  and  $t-2$ , divided by number of articles published in journal  $j$  in  $t-1$  and  $t-2$



IF is part of Thomson Reuter's Web of Science Journal Citation Reports (JCR)

- Commonly available tool
- Available in "science" and "social science" version

# Journal Impact Factor: network perspective

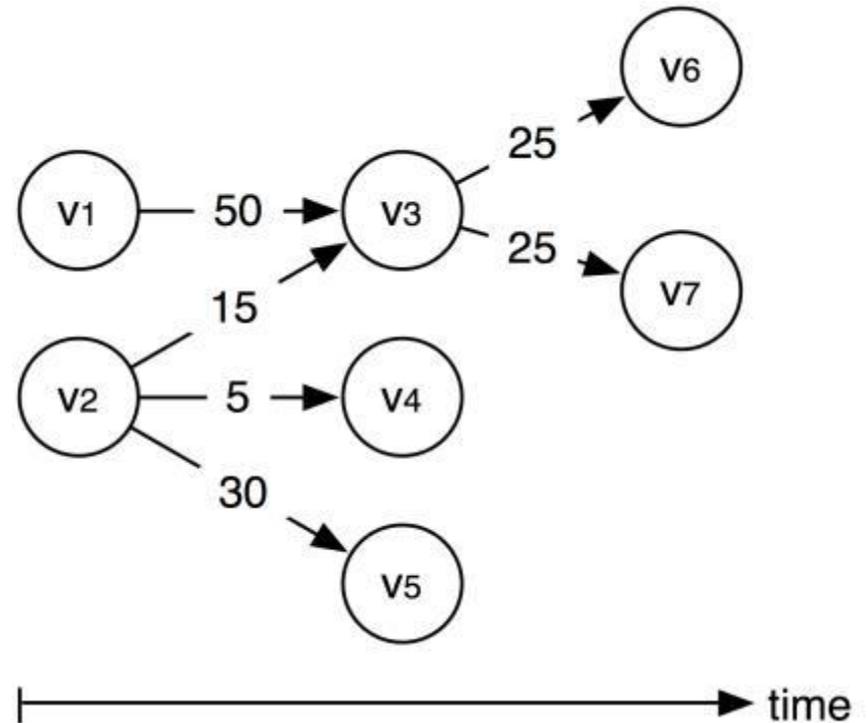
## JCR = citation graph

- +- 9k journals
- +-10M weighted citation edges
- A directed, weighed graph

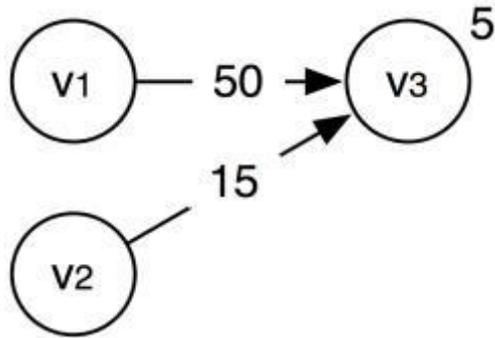
$$G = (V, E, W)$$

$$E \subseteq V^2$$

$$W : E \rightarrow \mathbb{N}^+$$



# Journal Impact Factor: normalized in-degree



$$IF_j = \frac{\sum_i w_{ij}}{N_j}$$

=Normalized in-degree

# Uses of the IF

- Frequently used to assess scholarly impact for authors, articles, journals, teams, countries by proxy
- Often by constructing “aggregate indicators”: average IF for author, IFs for top 5 papers, etc.
- Common in a variety of settings: promotion, tenure, funding decisions, and in many countries part of official requirements and science evaluation policies

Assumption is that the status of the journal somehow “rubs off” on those who published or that which was published.

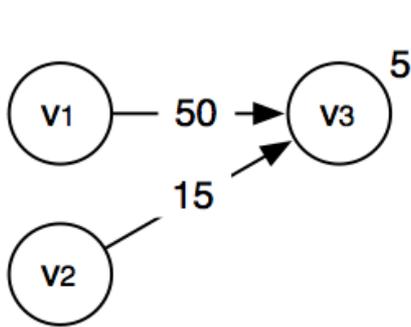
# Pros and cons

- Simple and easy to understand
- Based on article citation data (aggregated at journal level)
- Commonly used
- Does correspond to general idea of journal impact

BUT, Simple  $\neq$  valid

- Mean calculated from very skewed distribution: 80/20 rule
- Only 2 year period under consideration!
- Not appropriate for articles/author evaluation: journal impact  $\neq$  author impact  $\neq$  article impact
- Enumerator/Denominator manipulation:
  - Editors may choose to publish “opinion” pieces
  - Coerced citation

# Citation network: other metrics?



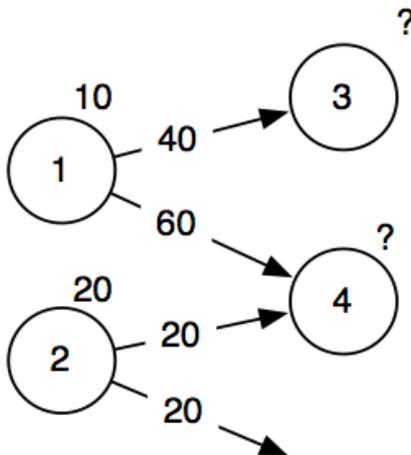
Impact Factor:  $IF_j = \frac{\sum_i w_{ij}}{N_j}$ :

normalized citation count

**origin of citation disregarded**

If it's good enough for Google...

A different route:



- $IF(v_i) \simeq \lambda \sum_j IF_j$

- $IF(v_i) \simeq \lambda \sum_j IF_j \times \frac{1}{O(v_j)}$

- $PR(v_i) \simeq \lambda \sum_j PR(v_j) \times \frac{1}{O(v_j)}$

- $PR(v_i) \simeq \frac{(1-\lambda)}{N} + \lambda \sum_j PR(v_j) \times \frac{1}{O(v_j)}$

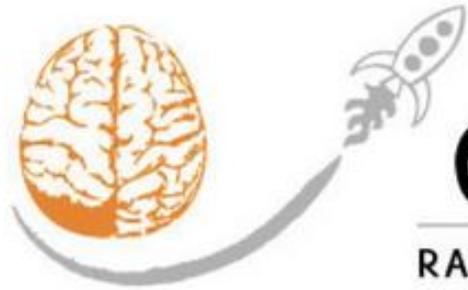
- $PR_w(v_i) = \frac{(1-\lambda)}{N} + \lambda \sum_j PR_w(v_j) \times w(v_j, v_i)$

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. *Information processing and management*, 12(5), 297-312.

Chen, P., Xie, H., Maslov, S., & Redner, S. (2007). Finding scientific gems with Google. *Journal of Informetrics*, 1(1), arxiv.org/abs/physics/0604130.

Johan Bollen, Marko A. Rodriguez, and Herbert Van de Sompel. *Journal status*. *Scientometrics*, 69(3), December 2006 (arxiv.org:cs.DL/0601030, DOI: 10.1007/s11192-006-0176-z)

# Eigenfactor.com



**eigenFACTOR.org**<sup>®</sup>  
RANKING AND MAPPING SCIENTIFIC KNOWLEDGE

[eigenfactor search](#) | [mapping](#) | [information](#) | [well-formed](#) | [cost effectiveness](#) | [recommendation](#) | [about](#)

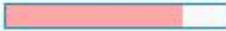
Order Journal

Percentile

EF ↓

AI ↓

1 J INFORMETR  
ISSN: 1751-1577

EF:  48  
AI:  80

0.001934

1.0069

“Article Influence score measures the average influence, per article, of the papers in a journal.”

# Eigenfactor

- Powerful metric that fixes many of the shortcomings of the IF
- Based on solid network science
- Longer citation period under consideration (5 years)
- Authors continuously work to produce innovative visualizations and update metrics

BUT (minor points):

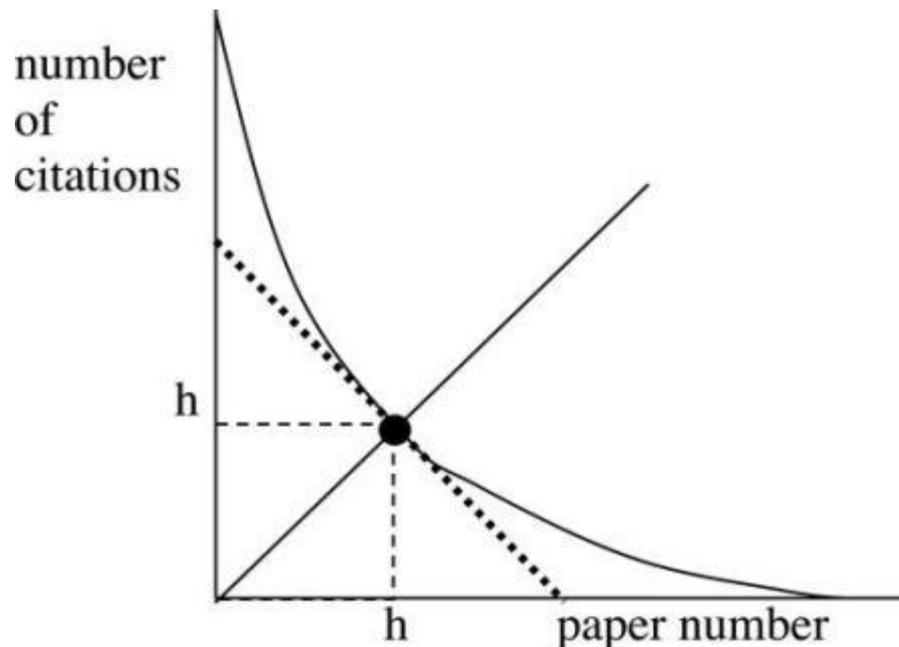
- Definition may not be intuitive for those not familiar with network science
- Applies mainly to journals, although Article Influence was added to lineup, and recent work with SSRN: authors, institution, and countries
- Parameters of calculation can vary and are to some degree arbitrary

# H-index

- Geared toward directly evaluating an author's productivity and impact
- Based on citation distribution of author's publication record
- Introduced by: Hirsch, J. E. (15 November 2005). "An index to quantify an individual's scientific research output". PNAS 102 (46): 16569–16572.

# H-index: calculation

- “A scientist has index  $h$  if  $h$  of his/her  $N_p$  papers have at least  $h$  citations each, and the other  $(N_p - h)$  papers have no more than  $h$  citations each.”



From: Hirsch, J. E. (15 November 2005). "An index to quantify an individual's scientific research output". PNAS 102 (46): 16569–16572.

# H-index: Pros and cons

- Simple and easy to understand
- Represents both productivity and impact of author's publication record
- Commonly used, commonly understood
- Can be used for journals, team, countries

BUT! 😊

- Does not take age of publication into account, sensitive to age of author or rather career length
- Does not take into account origin of citation
- Field dependent (but may be fixable: Radicchi, 2008)
- Why not simply calculate mean citations per paper?

# H-index: services

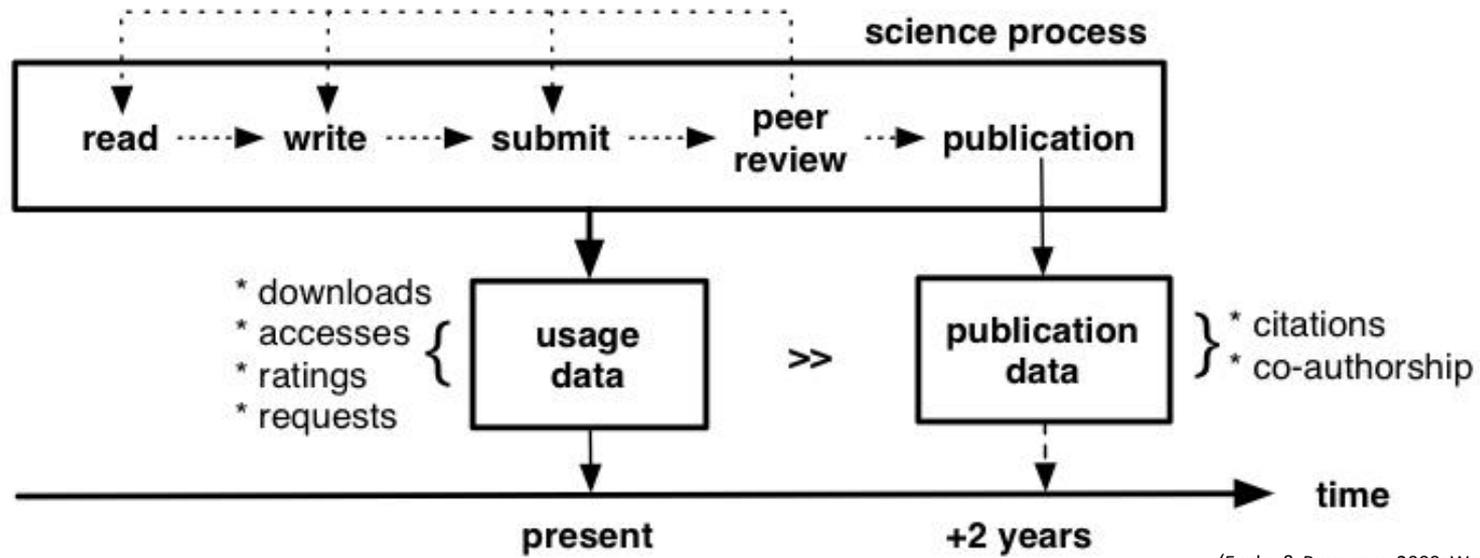
- Scholarometer: [scholarometer.indiana.edu](http://scholarometer.indiana.edu)
- Publish or perish: [www.harzing.com/pop.htm](http://www.harzing.com/pop.htm)
- Google Scholar: [scholar.google.com/citations](http://scholar.google.com/citations)
- Microsoft Academic Search:  
[academic.research.microsoft.com](http://academic.research.microsoft.com) (also offers a range of other citation-based indicators and very cool analytics)

Reliability varies with the quality and coverage of the citation data

# Lots of metrics, but underlying data?

- Are citations really the only valid way to count impact?
- Scholarship is now an online process
- Online indicators of scholarly impact?

# The scientific process: the importance of early indicators



(Egghe & Rousseau, 2000; Wouters, 1997)  
(Brody, Harnad, & Carr 2006),

## Usage data

- Scale, cf. Elsevier downloads (+1B) vs. Wos citations (650M)
- Immediate, early stages
- Variety of resources and actors

## Citation: final products

- Publication delays
- Focus on publications
- Focus on authors

# Usage data

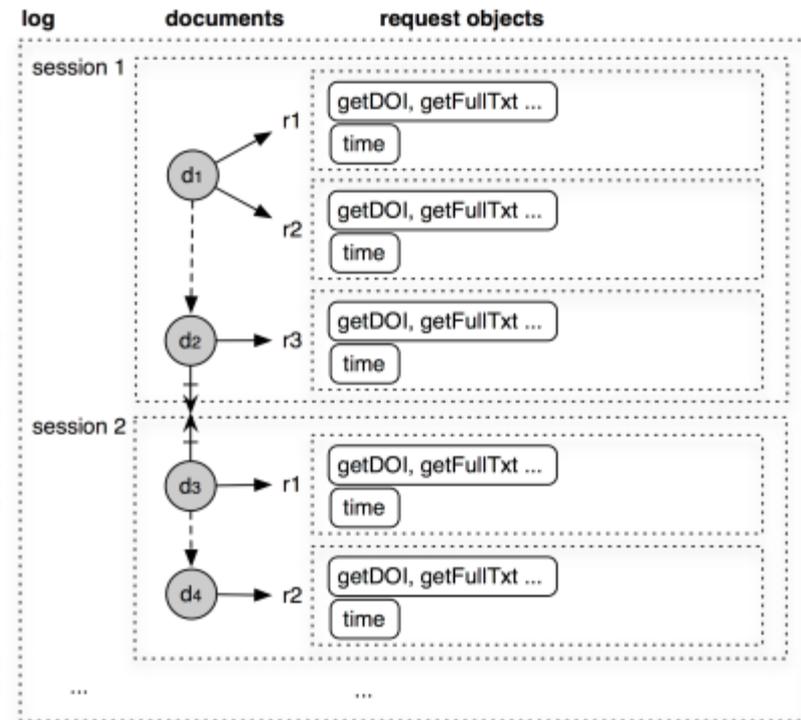
- **Detailed** records of user interactions with scholarly communication items
- Server logs: generally recorded as by-product of online access
- Does not measure actual reading or usage
  - Corresponds to more nebulous notions of “attention” or “behavior”
  - How to translate to indicators of actual impact?

# Usage data: fields

Fields and properties commonly found in usage data (and required by one of our own projects):

- Unique usage events (article level)
- Fields: unique session ID, date/time, unique document ID and/or metadata, request type

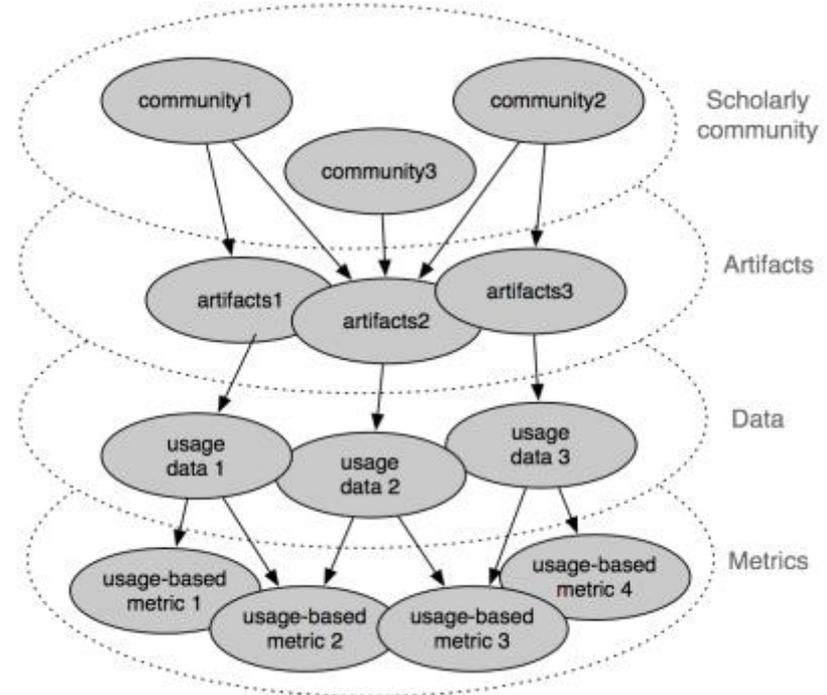
2007	9	1	0	0	1	CFA	cffoe	A172080.N1.Vanderbilt.Edu	unknown	AST	A	1996SF	
2007	9	1	0	0	1	CFA	cffoe	210.94.41.89	unknown	PHY	A	2007ApPhL.90a2120C	
2007	9	1	0	0	1	CFA	cffoe	24-196-228-125.dhcp.gwnt.ga.charter.com	unknown	AST	A		
2007	9	1	0	0	4	CFA	cffoe	163.152.35.114	4700387eae	PHY	A	1993WRR..29.1	
2007	9	1	0	0	6	CFA	cffoe	pd9e980fc.dip0.t-ipconnect.de	45f0c69881	AST	X	2i	
2007	9	1	0	0	1	CFA	cffoe	A172080.N1.Vanderbilt.Edu	unknown	AST	A	1996SF	
2007	9	1	0	0	1	CFA	cffoe	210.94.41.89	unknown	PHY	A	2007ApPhL.90a2120C	
2007	9	1	0	0	1	CFA	cffoe	24-196-228-125.dhcp.gwnt.ga.charter.com	unknown	AST	A		
2007	9	1	0	0	4	CFA	cffoe	163.152.35.114	4700387eae	PHY	A	1993WRR..29.1	
2007	9	1	0	0	6	CFA	cffoe	pd9e980fc.dip0.t-ipconnect.de	45f0c69881	AST	X	2i	
2007	9	1	0	0	6	CFA	cffoe	foel25144.4u.com.gh	47002f8eda	PHY	A	2002AGl	
2007	9	1	0	0	6	CFA	cffoe	66-215-171-214.dhcp.ccmn.ca.charter.com	4681d22a6f	AS			
2007	9	1	0	0	7	CFA	cffoe	nat-ptouser3.uspto.gov	unknown	PHY	A	2005ApPhL.86i	
2007	9	1	0	0	7	CFA	cffoe	cpe-71-65-25-115.ma.res.rr.com	unknown	PHY	A	1980S	
2007	9	1	0	0	7	CFA	cffoe	customer3491.pool1.unallocated-106-0.orangehomedsl.co.uk					
2007	9	1	0	0	8	CFA	cffoe	Uranus.seas.ucla.edu	46672d96b2	PHY	A	1966Phy	
2007	9	1	0	0	9	CFA	cffoe	75-121-173-37.dyn.centurytel.net	46cf1fd8a6	AST	D		
2007	9	1	0	0	13	CFA	cffoe	foel17-18.kln.forthnet.gr	unknown	AST	A	1987cosm.	
2007	9	1	0	0	15	CFA	cffoe	hades.astro.uiuc.edu	46f707564d	PRE	A	2007arXi	
2007	9	1	0	0	17	CFA	cffoe	ool-43554752.dyn.optonline.net	unknown	PHY	A	2000F	
2007	9	1	0	0	17	CFA	cffoe	c-68-33-176-222.hsd1.md.comcast.net	unknown	GEN	A		
2007	9	1	0	0	19	CFA	cffoe	74-36-139-46.dr02.brvl.mn.frontiernet.net	unknown	AST			
2007	9	1	0	0	19	CFA	cffoe	c-76-16-53-120.hsd1.il.comcast.net	46f667b71b	AST			
2007	9	1	0	0	20	CFA	cffoe	74-39-37-62.nas03.roch.ny.frontiernet.net	unknown	PHY			
2007	9	1	0	0	22	ANU	cffoe	bio-mirror	uatu-virtual1.anu.edu.au	46f9e8f87f	AST	A	
2007	9	1	0	0	22	CFA	cffoe	fw.hia.nrc.ca	46f1531d59	AST	A	2002P&SS..50.74E	
2007	9	1	0	0	22	CFA	cffoe	24-117-0-220.cpe.cableone.net	unknown	AST	A	1984E	



# From usage data to metrics

Usage data is generally dependent on:

- Community: user communities pertain to particular digital services.
- Artifacts: Vary according to institutional policies and subscriptions.
- Data: usage data therefore limited to particular sub-communities and collections of artifacts.
- Metrics: various metrics studied. Do differences result from sample, collection or metric definition? What do they mean? What type of impact/prestige do they express?



# Efforts to leverage usage data

Each approaching aforementioned issues in unique manner:

Project	Features
<b>COUNTER</b>	standard/compliance oriented, journal level, statistics (counts)
<b>MESUR</b>	format- and standards eclectic, article and journal level, statistics and network metrics
<b>PLoS</b>	Single, local server format, article level, statistics
<b>Eigenfactor</b>	format- and standards eclectic, article, journal, author level, network metrics

# COUNTER

Counting Online Usage of NeTworked Electronic Resources

[Home](#) | [About Counter](#) | [FAQs](#) | [Code of Practice](#) | [News and Activities](#) | [Compliant Vendors](#) | [Contact Us](#) | [Members](#) | [Usage Factor](#)

**Update**

**[Release 4 of the COUNTER Code of Practice now published](#)**

Click here to [apply for COUNTER Membership](#)

## About COUNTER

The use of online information resources is growing rapidly. It is widely agreed by producers and purchasers of information that the use of these resources should be measured in a more consistent way. Librarians want to understand better how the information they buy from a variety of sources is being used; publishers want to know how the information products they disseminate are being accessed. An essential requirement to meet these objectives is an agreed international set of standards and protocols governing the recording and exchange of online usage data. The COUNTER Codes of Practice provide these standards and protocols and are published in full on this website. Currently available are:

[Release 4 of the COUNTER Code of Practice for e-Resources](#) (Published April 2012)

[Release 3 of the COUNTER Code of Practice for Journals and Databases](#) (published August 2008)

[Release 1 of the COUNTER Code of Practice for Books and Reference Works](#) (published in March 2006)

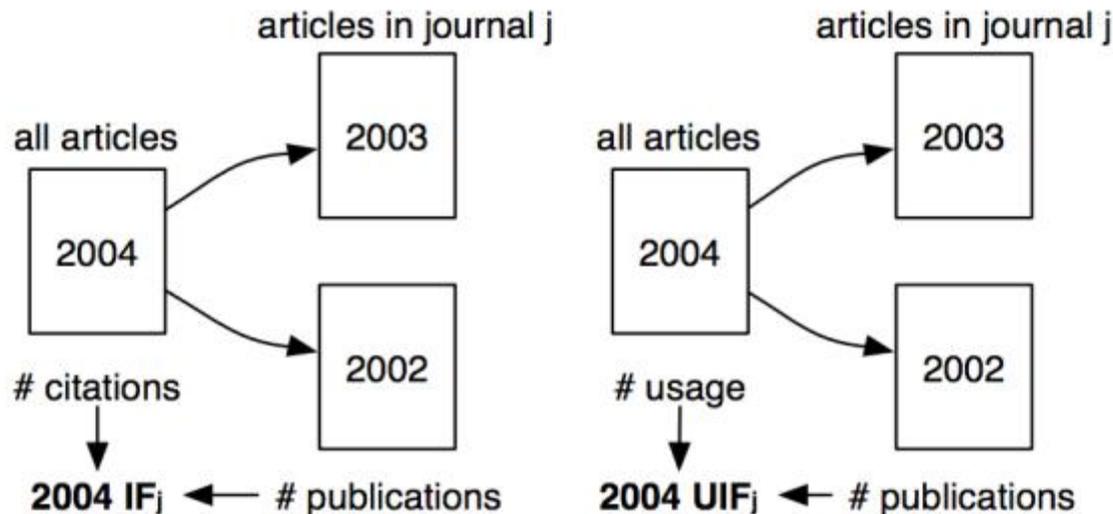
## COUNTER compliance - a step-by-step Guide for vendors

The purpose of this Guide is to take Vendors who are interested in becoming COUNTER compliant through the compliance process, to enable them to select the COUNTER reports that are relevant to their products, and to describe what happens once a Vendor has become compliant. [COUNTER compliance: a step-by-step Guide for Vendors](#)

# Usage Impact Factor?

COUNTER statistics:

- Monthly usage counts per journal (and potentially articles)
- Use citation-based Impact Factor definition



# PLoS Article Level Metrics

RESEARCH ARTICLE

OPEN ACCESS

## A Principal Component Analysis of 39 Scientific Impact Measures

Article

Metrics

Related Content

Comments: 3

### Article Usage ⓘ

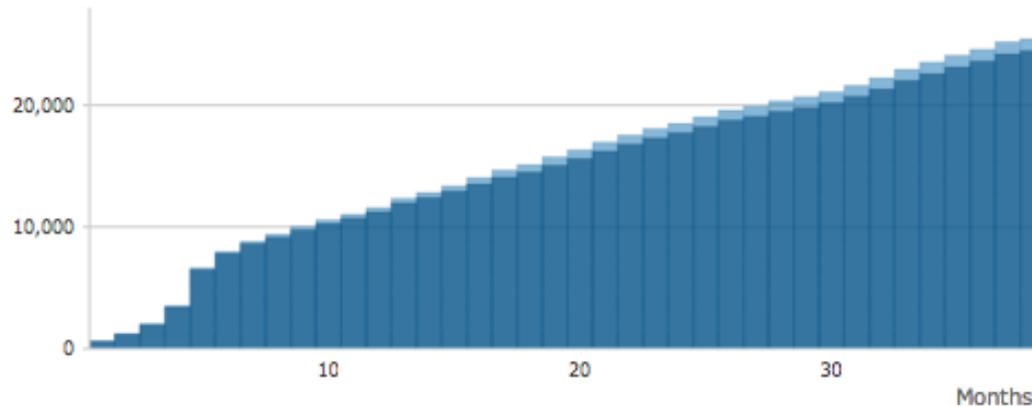
Total Article Views

**25,459**

Jun 29, 2009 (publication date)  
through Jul 17, 2012\*

	HTML Page Views	PDF Downloads	XML Downloads	Totals
PLoS	20,961	3,429	67	24,457
PMC	672	330	n.a.	1,002
Totals	21,633	3,759	67	25,459

Cumulative Views



# MESUR

studying science from large-scale usage data

[FRONT PAGE](#)

[ABOUT](#)

[PEOPLE](#)

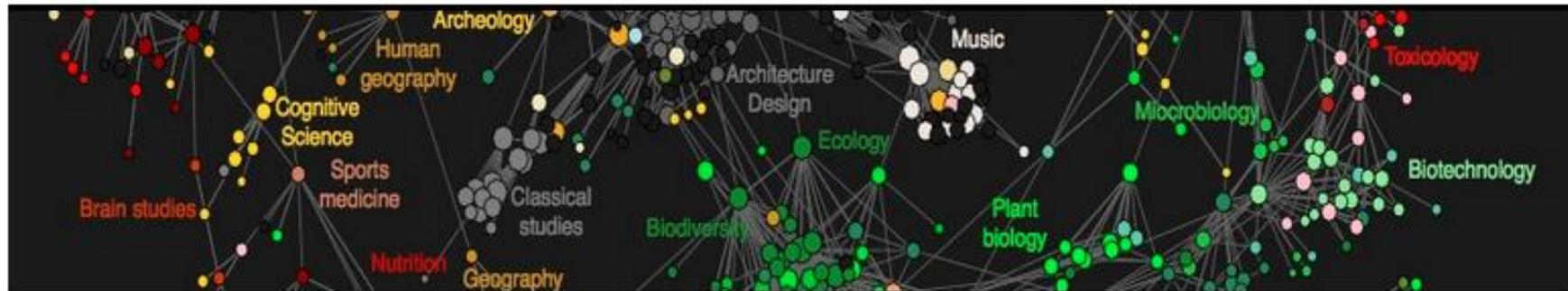
[PUBLICATIONS](#)

[EVENTS](#)

[DEMOS](#)

[DATA](#)

RSS 



## MESUR is now operating at Indiana University – School of Informatics and Computing.

February 23rd, 2010 · Uncategorized

MESUR has continued its activities at the Indiana University, School of Informatics and Computing. Agreements have been closed with major providers of usage data. We are funded by:

1) the National Science Foundation( #0914939: Tracking Scientific Innovation from Usage Data: Models and Tools to Support a Science of Science), a collaboration with Carl Bergstrom of University of Washington.

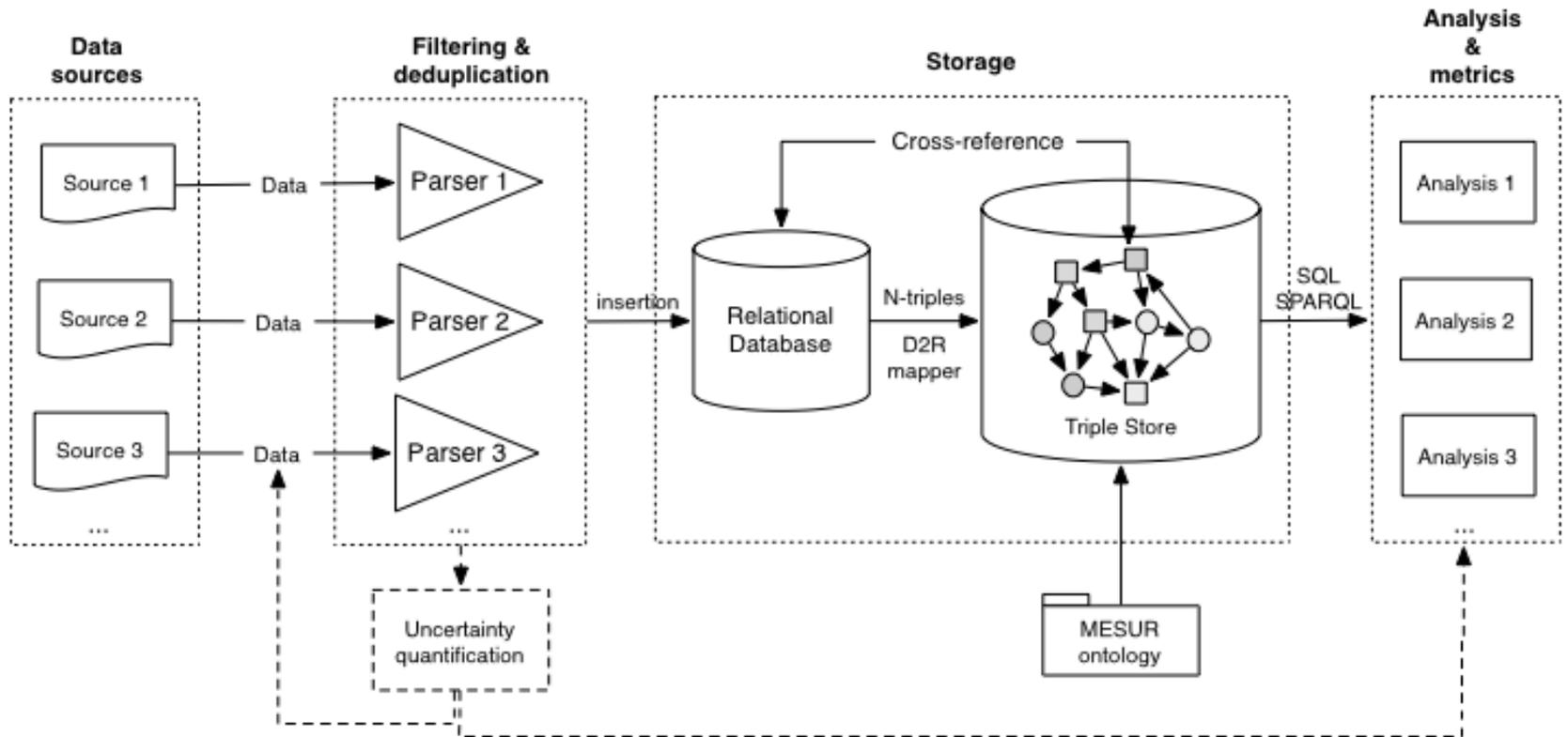
### SEARCH IT!

To search, type and hit enter

### RECENT ENTRIES

- [MESUR is now operating at Indiana University – School of Informatics and Computing.](#) 2.23
- [Visit the archives for more!](#)

# MESUR



<http://www.mesur.org/schemas/2007-01/mesur/>

# MESUR Mapping and ranking



- directed unweighted in-degree centrality
- directed unweighted out-degree centrality
- directed unweighted pagerank centrality
- directed weighted in-degree centrality
- directed weighted out-degree centrality
- directed weighted pagerank centrality
- journal use probability
- undirected unweighted betweenness centrality
- undirected unweighted closeness centrality
- undirected unweighted out-degree centrality
- undirected unweighted pagerank centrality
- undirected weighted betweenness centrality
- undirected weighted closeness centrality
- undirected weighted out-degree centrality
- undirected weighted pagerank centrality
- usage impact factor**
- uses-duha
- uses-duhh
- uses-dwha
- uses-dwhh

Show journal ranking of subject:  year:

by metric:  usage-based  citation-based

with title containing:

Among 76 in the subject, 76 title(s) matched 1 - 76

Rank	Title/ISSN	Percentile	Metric
1	austral ecology 1442-9985	<div style="width: 98.68%;"></div> 98.68	37.3759
2	international journal of sustainable development and wo... 1350-4509	<div style="width: 97.37%;"></div> 97.37	31.0286
3	ecography 0906-7590	<div style="width: 96.05%;"></div> 96.05	29.766
4	journal of applied ecology 0021-8901	<div style="width: 94.74%;"></div> 94.74	26.8812
5	restoration ecology 1061-2971	<div style="width: 93.42%;"></div> 93.42	26.1185
6	global ecology and biogeography 1466-822X	<div style="width: 92.11%;"></div> 92.11	24.6489

# MESUR Mapping and ranking

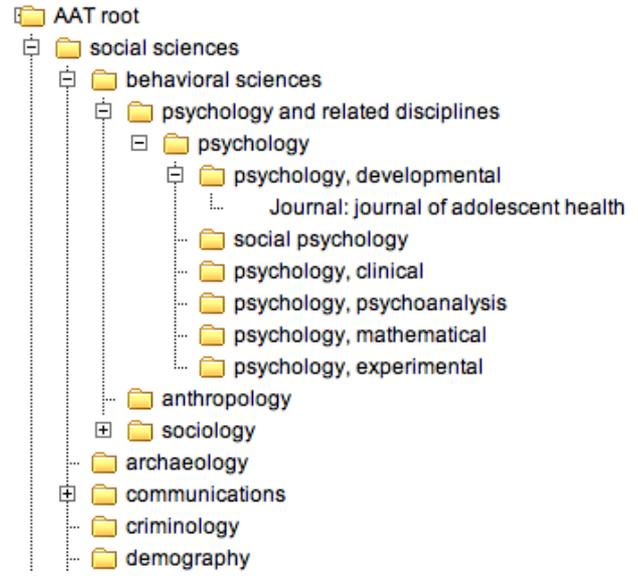


Show profile of journal with title or ISSN:

## Identifiers

- Title: journal of adolescent health
- ISSN: 1054-139x
- Alternative title: j adoles h
- Alternative title 2: j adolesc health
- Alternative title 3: j adolescent health

## Subject branch



## Metric percentile by subject depth

Metric:  usage-based  citation-based  
  
Year:

social sciences	98.82
behavioral sciences	97.35
psychology and related disciplines	96.49
psychology	96.49
psychology, developmental	92.86

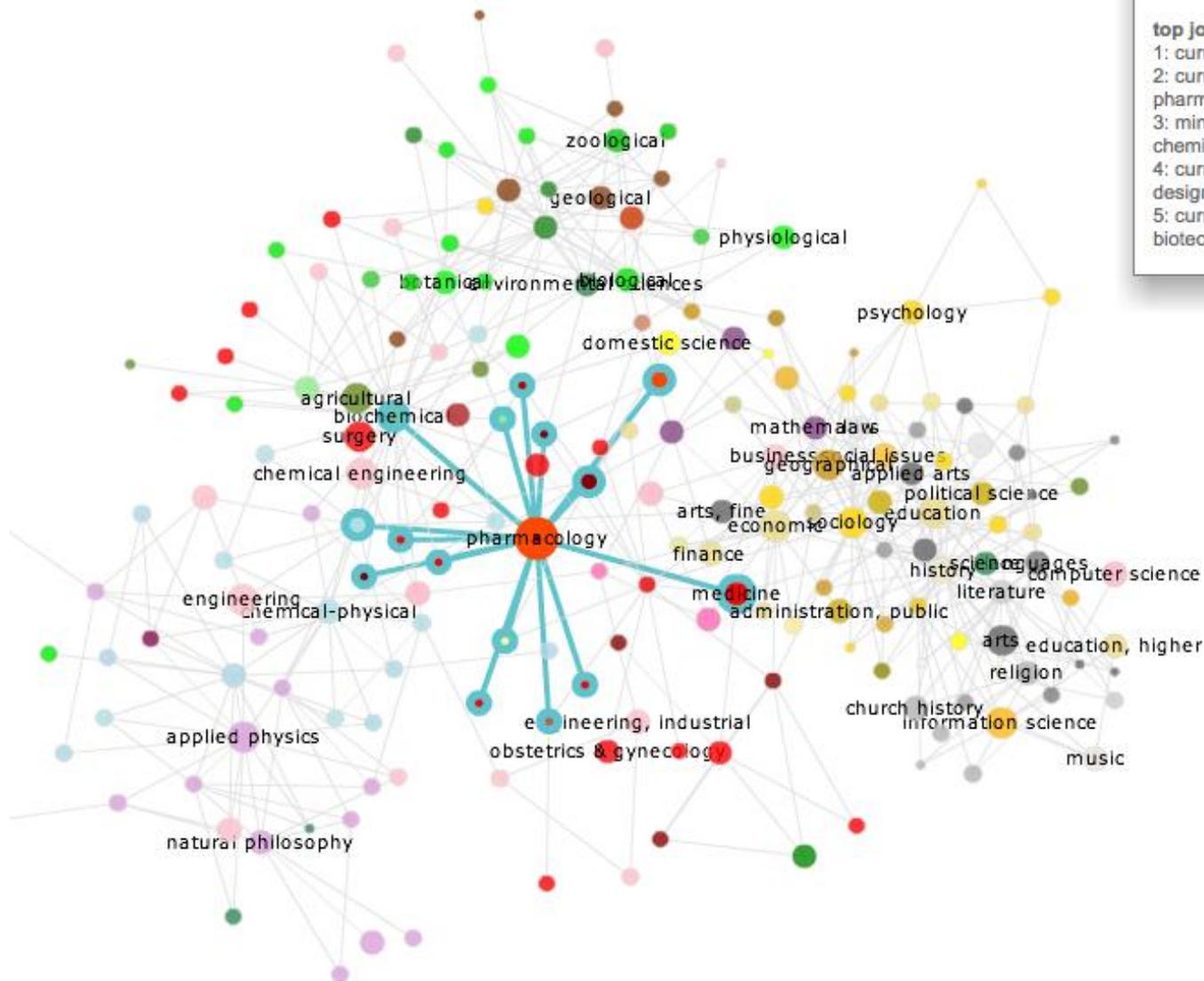
# MESUR Mapping and ranking

**pharmacology**

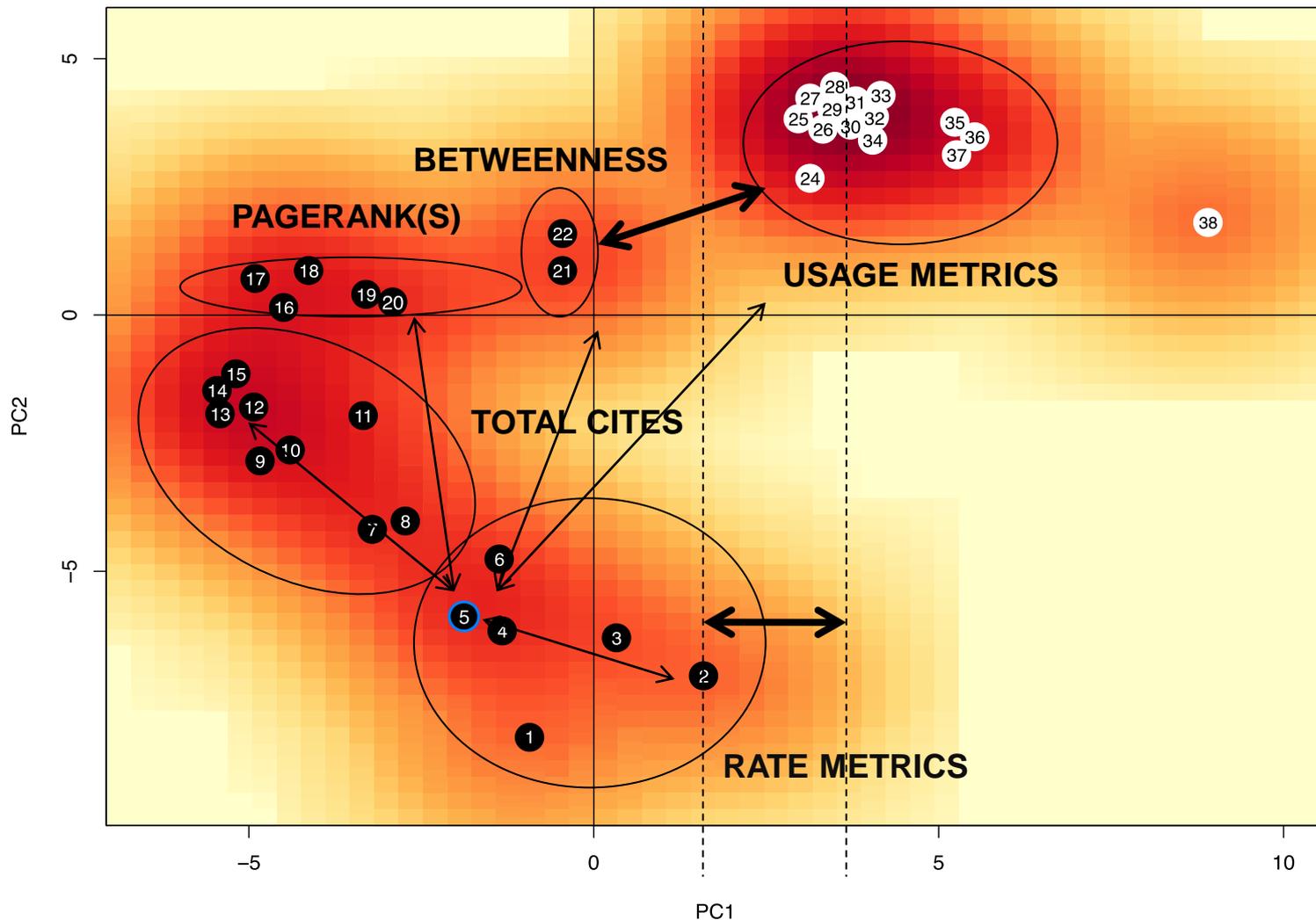
Journal count: 999

**top journals:**

- 1: current drug targets
- 2: current vascular pharmacology
- 3: mini reviews in medicinal chemistry
- 4: current pharmaceutical design
- 5: current pharmaceutical biotechnology



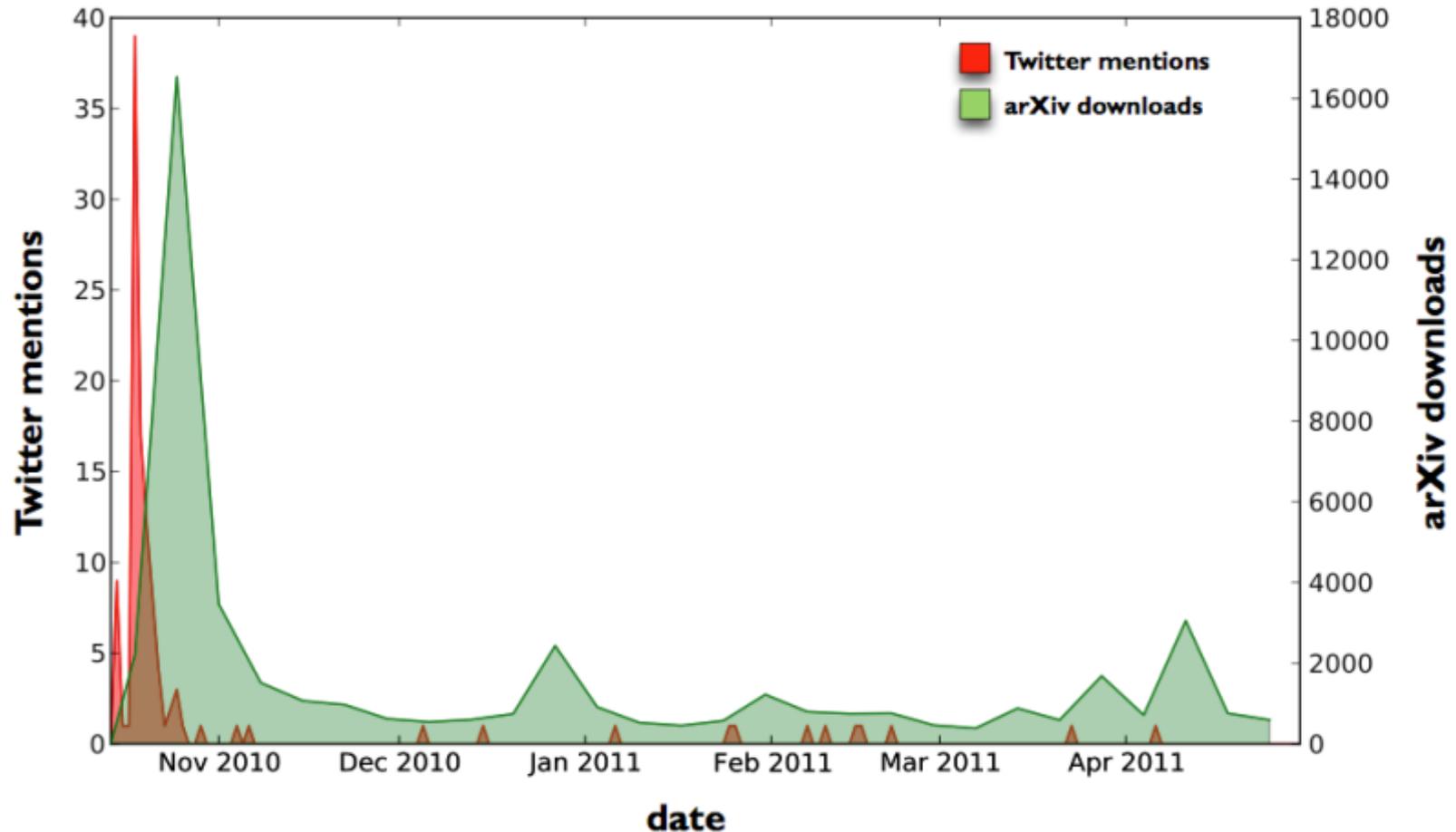
# The MESUR Metrics Map



ID	Type	Measure
1	Citation	Scimago Journal Rank
2	Citation	Immediacy Index
3	Citation	Closeness
4	Citation	Cites per doc
5	Citation	Journal Impact Factor
6	Citation	Closeness centrality
7	Citation	Out-degree centrality
8	Citation	Out-degree centrality
9	Citation	Degree Centrality
10	Citation	Degree Centrality
11	Citation	H-Index
12	Citation	Scimago Total cites
13	Citation	Journal Cite Probability
14	Citation	In-degree centrality
15	Citation	In-degree centrality
16	Citation	PageRank
17	Citation	PageRank
18	Citation	PageRank
19	Citation	PageRank
20	Citation	Y-factor
21	Citation	Betweenness centrality
22	Citation	Betweenness centrality
23	Citation	<i>Citation Half-Life</i>
24	Usage	Closeness centrality
25	Usage	Closeness centrality
26	Usage	Degree centrality
27	Usage	PageRank
28	Usage	PageRank
29	Usage	In-degree centrality
30	Usage	Out-degree centrality
31	Usage	PageRank
32	Usage	PageRank
33	Usage	Betweenness centrality
34	Usage	Betweenness centrality
35	Usage	Degree centrality
36	Usage	Out-degree centrality
37	Usage	In-degree centrality
38	Usage	Journal Use Probability
39	Usage	<i>Usage Impact Factor</i>

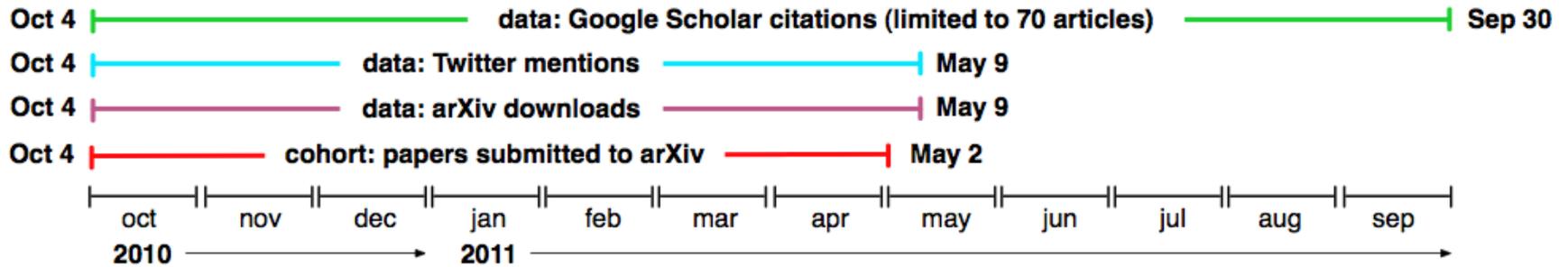
Johan Bollen, Herbert Van de Sompel, Aric Hagberg and Ryan Chute. A Principal Component Analysis of 39 Scientific Impact Measures. PLoS ONE, June 2009. URL: <http://dx.plos.org/10.1371/journal.pone.0006022>.

# And now for something completely different...



Usage data represents only a small fraction of online activity related to scholarly communication items

# Twitter -> arxiv downloads -> citations?



**Figure 1.** Timeline of data collection. Our cohort consists of all papers submitted to arXiv between October 4, 2010 and May 2, 2011. Weekly article downloads and daily Twitter mentions were recorded after the article’s submission date, up to May 9, 2011. Early citation counts for each article were manually recorded from Google Scholar on September 30th, 2011.

**Table 3.** Multi variant linear regression analysis of article citations  $C$  vs. twitter mentions  $T$ , article arXiv downloads  $A$ , and time in days elapsed between beginning of our test period and submission of article,  $P$ .

model	$\beta_1$ (st. error)	$\beta_2$ (st. error)	$\beta_3$ (st. error)
$C = \beta_1 T + \beta_2 P + \varepsilon_1$	0.150*** (0.035)	0.044** (0.019)	-
$C = \beta_1 A + \beta_2 P + \varepsilon_2$	2e-04*** (7e-05)	0.038* (0.020)	-
$C = \beta_1 T + \beta_2 A + \beta_3 P + \varepsilon_3$	0.120*** (0.040)	1e-04 (8e-05)	0.041** (0.019)
*: $p < 0.1$ , **: $p < 0.05$ , ***: $p < 0.01$ , ****: $p < 0.001$			

# Social media indicators

← → ↻ 🏠 total-impact.org

## total·impact

create about api bl

### Uncover the invisible impacts of your research.

total-impact aggregates altmetrics: diverse impacts from your articles, datasets, blog posts, and more.

**get my impact!** | or, show me a sample collection

**My Collection**

- articles**  
Shotton, Forwitz, Kijne, Miles (2008) Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article. PLoS Comput Biol. 11652
- datasets**  
(2011) Data from: Data archiving is a good investment. Dryad Digital Repository. 1273
- software**  
(2011) Sunb-Twitter-study Study: how many scholars use Twitter, and how do they use it? GitHub. 5 4

# Indicators based on bookmarking, “likes”

## ReaderMeter ALPHA

Forename(s)

Surname

Search

*...research impact, crowdsourced*

[Example of author readership analysis](#)

[NEW Act now: require free access to publicly funded research](#)

[Read the blog post of the official launch](#)

[Read the alt-metrics manifesto](#)

[About](#)

[FAQ](#)

[News](#)

[Contact](#)

Management software



MENDELEY

Welcome

Dashboard My Library Papers Groups People

How to choose a good scientific problem.

by Uri Alon

Biological Sciences | Miscellaneous Papers

Overview

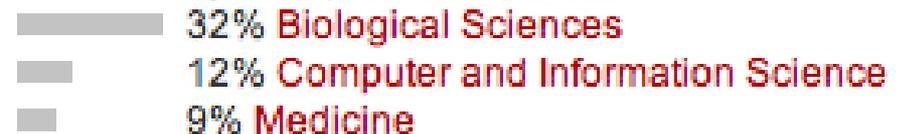
Mendeley Post (2008)

## Readership Statistics

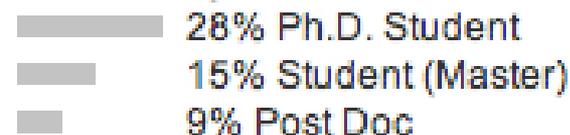


3627 Readers on Mendeley

by Discipline



by Academic Status



by Country



# Conclusions

- Evolving notions of scholarship: evolving notions of impact
- Existing impact metrics offer some degree of quantitative analytics, but are limited by the data that they rely on, where they can be validly applied, and by not taking into account “social” structure of scholarly community
- I think the latter is a major deficiency: the gift economy of science is based on a social network of scholars
- As scholarship goes online, expect increasing progress in the area of online impact indicators: usage data, and even social media indicators (Twitter, blogs, etc)

Challenge: community understanding of which metric best suits a particular community, particular set of entities, and validly represents which aspect of scholarly impact.

# Some relevant publications.

- Xin Shuai, Alberto Pepe, and Johan Bollen. How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations, <http://arxiv.org/abs/1202.2461>, 2012 (submitted to PLoS ONE)
- Johan Bollen, Herbert Van de Sompel, Aric Hagberg and Ryan Chute. A Principal Component Analysis of 39 Scientific Impact Measures. PLoS ONE, June 2009. URL: <http://dx.plos.org/10.1371/journal.pone.0006022>.
- Johan Bollen, Herbert Van de Sompel, Aric Hagberg, Luis Bettencourt, Ryan Chute, Marko A. Rodriguez, Lyudmila Balakireva. Clickstream data yields high-resolution maps of science. Clickstream data yields high-resolution maps of science. PLoS One, February 2009.
- Johan Bollen, Marko A. Rodriguez, and Herbert Van de Sompel. **Journal status**. Scientometrics, 69(3), December 2006 (arxiv.org:cs.DL/0601030)
- Johan Bollen, Herbert Van de Sompel, and Marko A. Rodriguez. **Towards usage-based impact metrics: first results from the MESUR project**. In Proceedings of the Joint Conference on Digital Libraries, Pittsburgh, June 2008
- Marko A. Rodriguez, Johan Bollen and Herbert Van de Sompel. **A Practical Ontology for the Large-Scale Modeling of Scholarly Artifacts and their Usage**, In Proceedings of the Joint Conference on Digital Libraries, Vancouver, June 2007
- Johan Bollen and Herbert Van de Sompel. **Usage Impact Factor: the effects of sample characteristics on usage-based impact metrics**. (cs.DL/0610154)
- Johan Bollen and Herbert Van de Sompel. **An architecture for the aggregation and analysis of scholarly usage data**. In Joint Conference on Digital Libraries (JCDL2006), pages 298-307, June 2006.
- Johan Bollen and Herbert Van de Sompel. **Mapping the structure of science through usage**. Scientometrics, 69(2), 2006.
- Johan Bollen, Herbert Van de Sompel, Joan Smith, and Rick Luce. **Toward alternative metrics of journal impact: a comparison of download and citation data**. Information Processing and Management, 41(6):1419-1440, 2005.