New Metrics, a Chance for Changing Scientometrics!
A Preliminary Discussion of Recent Approaches

Dr.-Ing. Bernd Markscheffel

Chair of Information and Knowledge Management
Dept. for Business Informatics
Ilmenau University of Technology
Agenda

- Introduction
- Motivation
- Methodical Approach
- Results
- Summary and Future Work
Scientometrics Timeline

Bibliometrics/Scientometrics

Documents
- Citation based metrics
- Content Analysis

Webometrics
- Web Pages
  - Link Analysis
  - Web Usage Analysis
  - Web Content Analysis

Scientometrics 2.0
- Tweets, Blogs
- Images, Videos
- Social Network Data
  - Tweetation
  - Blog Recognition
  - Likes/Plus Evaluation

Scientometrics 3.0
- Linked Open Data (LOD)
- Ontologies
  - Typed Link Analysis
  - Semantic Annotation
  - Semantic MetaData

1970  2000  2010  20??
Methotology

• Literature Analysis (Webster & Watson)
• Time Range: 3 years
• Sources:
  – the PLOS (Public Library of Science) collection,
    • http://www.plos.org
  – the altmetrics Mendeley group collection,
    • http://www.mendeley.com/groups/586171/altmetrics/papers/
  – the G+ altmetrics group
    • https://plus.google.com/communities/118115269559611745770?hl=de
• Research questions:
  – Which resources where used?
  – Which indicators were introduced in these articles?
  – Which tendencies can be observed?
Results

- Morphological Box
## Motivation

<table>
<thead>
<tr>
<th>Changes in Scholarly Environment</th>
<th>Changes in Technology</th>
<th>e-Infra-structure</th>
<th>Scholarship as Online Process</th>
<th>Complex Research Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changes in Publication Behaviour</td>
<td>Growth in the Size of Scholarly Literature</td>
<td>Critics on Peer Review Process</td>
<td>Trend Towards &quot;Enhanced&quot; Publication</td>
<td></td>
</tr>
<tr>
<td>Concerns about Citation Based Metrics</td>
<td>Citation Based Metrics Do Not Indicate the Nature of Its Semantically Connection</td>
<td>Citation Metrics are Slow</td>
<td>Lack of Transparency</td>
<td></td>
</tr>
</tbody>
</table>
### Resources

<table>
<thead>
<tr>
<th>Resources</th>
<th>Documents</th>
<th>Journals</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classical Resources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Web 1.0 based Resources</strong></td>
<td>Web page</td>
<td>Web pages' link structure</td>
<td>Web page usage information</td>
</tr>
<tr>
<td>News &amp; Recommendation Services</td>
<td>Reddit, Digg, Slashdot, Faculty of 1000...</td>
<td>Micro-blogging Services Twitter, Plurk, Tencent Weibo...</td>
<td>Social Bookmarking &amp; Reference Managing Services Cangeo, CiteULike, Mendeley, Zotero, JiaThis, bShare...</td>
</tr>
<tr>
<td><strong>Web 2.0 based Resources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Network Services</td>
<td>Youtube, Binkx LinkedIn, Xing, Sina, RenRen...</td>
<td>Media Based Services Flickr, Instagram, SlideShare...</td>
<td>Citation based Services WoS, Scopus Google Scholar, MS Academic Search</td>
</tr>
<tr>
<td><strong>Web 3.0 based Resources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ontologies</td>
<td>Cyc, Gene Ontology, WordNet...</td>
<td>Linked Open Data (LOD) DBpedia, FOAF...</td>
<td>Blogs Wordpress ecosystem, Blogger ecosystem, Sina ecosystem...</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Linkage Analysis</th>
<th>Co-Occurrence Similarities</th>
<th>Quantitative Methods</th>
<th>Dimensionality Reduction</th>
<th>Cluster Analysis (CA)</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct Citation Linkage</td>
<td>Co-Word-Analysis</td>
<td>Correlation Analysis</td>
<td>Eigenvector Decomposition</td>
<td>K-means Clustering</td>
<td>Social Network Analysis (SNA)</td>
</tr>
<tr>
<td></td>
<td>Co-Citation Analysis</td>
<td>Author Co-Citation Analysis</td>
<td>Regression Analysis</td>
<td>Multi-dimensional Scaling (MDS)</td>
<td>Hierarchical Clustering</td>
<td>Web Impact Assessment (WIA)</td>
</tr>
<tr>
<td></td>
<td>Bibliographic Coupling</td>
<td>Paper Co-Citation Analysis</td>
<td>Factor Analysis (FA)</td>
<td>Pathfinder Networks</td>
<td></td>
<td>Web-Link Analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Additional Information

- **Co-Classification Analysis**
# New Metrics

<table>
<thead>
<tr>
<th>New Metrics</th>
<th>Altmetrics</th>
<th>Plumanalytics</th>
<th>Article Level Metrics (PLOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Products, Tools and Institutions</td>
<td>ReaderMeter</td>
<td>CitEl</td>
<td>ScienceCard</td>
</tr>
<tr>
<td>Impactstory</td>
<td>PublisherPerish</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Altmetrics

- Altmetric was founded by Euan Adie in 2011 and grew out of the burgeoning altmetrics movement.
- Mr. Adie had previously worked on Postgenomic.com, an open source scientific blog aggregator founded in 2006.
- Interested in taking the ideas from Postgenomic forward they entered an altmetrics app into Elsevier's Apps for Science competition and ended up winning.
- The first standalone version of the Altmetric Explorer was released in February 2012.
Examples: Altmetric

- Altmetric collects article level metrics and the online conversations around research on behalf of publishers, institutions and funders,
- combines a selection of online indicators (both scholarly and non-scholarly) to give a measurement of digital impact and reach,
- tracks, collects and measures large amounts of data collected from all of the places where scientists, patient advocates, journalists, nurses, engineers and members of the public talk about science online
  - for example, blogs, Twitter, Facebook, Google+, message boards and mainstream newspapers and magazines

  > http://www.altmetric.com/article-level-metrics.php
## Altmetrics' Source Characteristics

<table>
<thead>
<tr>
<th>Activity</th>
<th>Twitter</th>
<th>Other Social Media</th>
<th>News</th>
<th>Blogs</th>
<th>Multimedia</th>
<th>Reference Managers</th>
</tr>
</thead>
<tbody>
<tr>
<td>How actively the source is used as a communication medium.</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Level of Insight</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>The depth of the insights that a typical mention in the source delivers.</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>None</td>
</tr>
<tr>
<td>Content Creators</td>
<td>Everyone</td>
<td>Everyone</td>
<td>Communicators</td>
<td>Communicators</td>
<td>Communicators</td>
<td>Researchers</td>
</tr>
<tr>
<td>The kinds of users who create content in this source.</td>
<td>Everyone</td>
<td>Everyone</td>
<td>Communicators</td>
<td>Communicators</td>
<td>Communicators</td>
<td>Researchers</td>
</tr>
<tr>
<td>Collection Time</td>
<td>Rapid</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Slow</td>
<td>Slow</td>
</tr>
<tr>
<td>The approximate length of time it takes Altmetric to collect and display mentions from the source.</td>
<td>Rapid</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Slow</td>
<td>Slow</td>
</tr>
</tbody>
</table>

http://www.altmetric.com/sources.php
Altmetrics: Value all research products
Heather Piwowar

A new funding policy by the US National Science Foundation represents a sea-change in how researchers are evaluated, says Heather Piwowar.

Subject terms: Publishing • Policy • Media formats

What a difference a word makes. For all new grant applications from 14 January, the US National Science Foundation (NSF) asks a principal investigator to list his or her research "products" rather than "publications" in the biographical sketch section. This means that, according to the NSF, a scientist's worth is not dependent solely on publications. Data sets, software and other non-traditional research products will count too.

There are more diverse research products now than ever before. Scientists are developing and releasing better tools to document their workflow, check each other's work and share information, from data repositories to post-publication discussion systems. As it gets easier to publish a wide variety of material online, it should also become easy to recognize the breadth of a scientist's intellectual contributions.
Altmetrics: Value all research products

265

Score in context
- Puts article in the top 5% of all articles ranked by attention
- show more...

Mentioned by
- 266 tweeters
- 8 Facebook users
- 8 science blogs
- 9 Google+ users

Readers on
- 65 Mendeley
- 28 Cit/Ulike

Track this article
- Get email updates when this article is shared

So far Altmetric has seen 12 posts from 8 blogs.

Process behind a Nature Comment

Research remix.
Publishing a Comment in Nature involved a process unlike any I've experienced to date, so I figured I'd document it...
11-Jan-2013

Let's Face It, No One Can Read Everything

Hack E-Science Librarianship
I read. A lot. Still I'm constantly amazed at how little of the massive amounts of information available to me I...
05-Feb-2013

The research impact agenda must translate measurement into learning

Impact of Social Sciences
Funders and the wider research community must avoid the temptation to reduce impact to just things that can be measured, says I...
14-Feb-2013

Opinion: Measure for measuring's sake?

Welcome Trust Blog
Can we assess the impact of research by simply counting outputs? We must avoid the temptation to reduce impact to things we can...
31-Jan-2013

Tipping Points Comment in Nature

Carl Baetig
On Wednesday my comment piece with Alan. Tipping points: from patterns to predictions appeared in Nature (next to a nice comment)
12-Jan-2013

Altmetrics: Nature commentary by Heather Piwowar

Altmetrics: Value all research products

The Altmetric score is one measure of the quality and quantity of online attention that this article has received. You can read about how Altmetric scores are calculated here.

This article scored 264.84

The context below was calculated when this article was last mentioned on 12th July 2013

Compared to all articles in Nature

So far Altmetric has tracked 19,879 articles from this journal. They typically receive a lot more attention than average, with a mean score of 28.9 vs the global average of 3.8. This article has done particularly well, scoring higher than 98% of its peers.

All articles of a similar age

Older articles will score higher simply because they’ve had more time to accumulate mentions. To account for age we can compare this score to the 70,653 tracked articles that were published within six weeks on either side of this one in any journal. This article has done particularly well, scoring higher than 99% of its contemporaries.

Other articles of a similar age in Nature

We’re also able to compare this article to 822 articles from the same journal and published within six weeks on either side of this one. This article has done very well, scoring higher than 95% of its contemporaries.

All articles

More generally, Altmetric has tracked 1,373,326 articles across all journals so far. Compared to these this article has done particularly well and is in the 99th percentile: it’s in the top 5% of all articles ever tracked by Altmetric.
Examples: PlumX

- **Usage**
  - Downloads,
  - Views,
  - BookHholdings,
  - ILL,
  - Document Delivery

- **Captures**
  - Favorites,
  - Bookmarks,
  - Saves,
  - Readers,
  - Groups,
  - Watchers

- **Mentions**
  - Blog Posts,
  - News Stories,
  - Wikipedia Articles,
  - Comments,
  - Reviews

- **Social Media**
  - Tweets,
  - +1's,
  - Likes,
  - Shares,
  - Ratings

- **Citations**
  - PubMed,
  - Scopus,
  - Patents

---

**SCIENTOMETRICS STATUS AND PROSPECTS FOR DEVELOPMENT**

24.10.2013  Page 20
<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>Example Source(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage</td>
<td>Abstract Views</td>
<td>dSpace, ePrints, PLoS</td>
<td>The number of times the abstract of an article has been viewed</td>
</tr>
<tr>
<td></td>
<td>Clicks</td>
<td>bit.ly, Facebook</td>
<td>The number of clicks of a URL</td>
</tr>
<tr>
<td></td>
<td>Collaborators</td>
<td>GitHub</td>
<td>The number of collaborators of an artifact</td>
</tr>
<tr>
<td></td>
<td>Downloads</td>
<td>Dryad, Figshare, Slideshare, Github</td>
<td>The number of times an artifact has been downloaded</td>
</tr>
<tr>
<td></td>
<td>Figure Views</td>
<td>figshare, PLoS</td>
<td>The number of times the figure of an article has been viewed</td>
</tr>
<tr>
<td></td>
<td>Full Text Views</td>
<td>PLoS</td>
<td>The number of times the full text of an article has been viewed</td>
</tr>
<tr>
<td></td>
<td>Holdings</td>
<td>WorldCat</td>
<td>The number of libraries that hold the book artifact</td>
</tr>
<tr>
<td></td>
<td>HTML Views</td>
<td>PLoS</td>
<td>The number of times the html of an article has been viewed</td>
</tr>
<tr>
<td></td>
<td>PDF Views</td>
<td>dSpace, ePrints, PLoS</td>
<td>The number of times the PDF of an article has been viewed</td>
</tr>
<tr>
<td></td>
<td>Views</td>
<td>Dryad</td>
<td>The number of times the dataset has been viewed</td>
</tr>
<tr>
<td></td>
<td>Supporting Data Views</td>
<td>PLoS</td>
<td>The number of times the supporting data of an article has been v.</td>
</tr>
<tr>
<td>Captures</td>
<td>Bookmarks</td>
<td>CiteULike, Delicious</td>
<td>Number of times an artifact has been bookmarked</td>
</tr>
<tr>
<td></td>
<td>Favorites</td>
<td>Slideshare, YouTube</td>
<td>The number of times the artifact has been marked as a favorite</td>
</tr>
<tr>
<td></td>
<td>Followers</td>
<td>GitHub</td>
<td>The number of times a person or artifact has been followed</td>
</tr>
<tr>
<td></td>
<td>Forks</td>
<td>GitHub</td>
<td>The number of times a repository has been forked</td>
</tr>
<tr>
<td></td>
<td>Groups</td>
<td>CiteULike, Mendeley</td>
<td>Number of times an artifact has been placed in a group’s library</td>
</tr>
<tr>
<td></td>
<td>Readers</td>
<td>Mendeley</td>
<td>The number of people who have added the artifact to their library</td>
</tr>
<tr>
<td></td>
<td>Subscribers</td>
<td>Vimeo, YouTube</td>
<td>The number of people who have subscribed for an update</td>
</tr>
<tr>
<td></td>
<td>Watcher</td>
<td>Github</td>
<td>The number of people watching the artifact for updates</td>
</tr>
<tr>
<td>Mentions</td>
<td>Comment count</td>
<td>Facebook, Reddit, Slideshare, Vimeo, YouTube</td>
<td>The number of comments made about an artifact</td>
</tr>
<tr>
<td></td>
<td>Forum Topic Count</td>
<td>Vimeo</td>
<td>The number of topics in a forum discussing the artifact</td>
</tr>
<tr>
<td></td>
<td>Gist count</td>
<td>GitHub</td>
<td>The number of gists in the source code repository</td>
</tr>
<tr>
<td></td>
<td>Links</td>
<td>Wikipedia</td>
<td>The number of links to the artifact</td>
</tr>
<tr>
<td></td>
<td>Review count</td>
<td>SourceForge</td>
<td>The number of user reviews of the artifact</td>
</tr>
<tr>
<td></td>
<td>Blog count</td>
<td>Research Blogging, Science Seeker</td>
<td>The number of blog posts written about the artifact</td>
</tr>
<tr>
<td>Social Media</td>
<td>Likes</td>
<td>Facebook, Vimeo, YouTube</td>
<td>The number of times an artifact has been liked</td>
</tr>
<tr>
<td></td>
<td>+1</td>
<td>Google</td>
<td>The number of times an artifact has gotten a +1</td>
</tr>
<tr>
<td></td>
<td>Ratings</td>
<td>SourceForge</td>
<td>The average user rating of the artifact</td>
</tr>
<tr>
<td></td>
<td>Recommendations</td>
<td>Figshare, SourceForge</td>
<td>The number of recommendations an artifact has received</td>
</tr>
<tr>
<td></td>
<td>Score</td>
<td>Reddit</td>
<td>The number of upvote minus downvotes on Reddit</td>
</tr>
<tr>
<td></td>
<td>Tweets</td>
<td>Topsy</td>
<td>The number of tweets that mention the artifact</td>
</tr>
<tr>
<td></td>
<td>Cited by</td>
<td>PubMed, Scopus</td>
<td>The number of PubMed Central articles that cite the artifact</td>
</tr>
</tbody>
</table>
Social Networks and Interest Similarity: The Case of CiteULike

by Danielle H. Lee, Peter Brusilovsky

Computer and Information Science : Miscellaneous Papers

Abstract

In collaborative filtering recommender systems, there is little room for users to get involved in the choice of their peer group. It leaves users defenseless against various spamming or shifting attacks. Other social Web-based systems, however, allow users to self-select peers and build a social network. We argue that users self-defined social networks could be valuable to increase the quality of recommendation in CF systems. To prove the feasibility of this idea we examined how similar are interests of users connected by self-defined relationships in a collaborative tagging system CiteULike. Interest similarity was measured by similarity of tags and meta-data they share and tags they use. Our study shows that users connected by social networks exhibit significantly higher similarity on all explored levels (items, meta-data, and tags) than non-connected users. This similarity is the highest for directly connected users and decreases with the increase of distance between users. Among other interesting properties of information sharing is the finding that between-user similarity in social connections on the level of meta-data and tags is much larger than similarity on the level of items. Overall, our findings support the feasibility of social network based recommender systems and offer some guidelines to the prospective authors of these systems.

Author-supplied keywords

- Social Networks
- Information Sharing
- Collaborative Filtering

Captures

Mendeley - Readers: 24
Mendeley - Groups: 2

Related

- A Survey on the Relationship between Trust and Interest Similarity in Online Social Networks
- ISGoDr: A framework for interest similarity-based information filtering in social networks

Readership Statistics

- 26 Reads on Mendeley
  - by Discipline: 66% Computer and Information Science, 4% Medicine, 4% Social Sciences
  - by Academic Status: 56% Ph.D. Student, 23% Student (Master), 12% Student (Postgraduate)
  - by Country: 51% United States, 12% Brazil, 12% United Kingdom
 Article-Level Metrics (PLOS)

**Usage**
- PLOS: views
- PDF downloads
- XML downloads
- PMC: views
- PDF downloads

**Citations**
- PubMed Central
- CrossRef
- Scopus
- Web of Science

**PLOS**
- Comments
- Notes
- Ratings

**Social Network**
- CiteULike
- Mendeley
- Twitter
- Facebook

**Blogs & Media**
- Nature Blogs
- ScienceSeeker
- Research Blogging
- Wikipedia
- Trackbacks
Summary and Future Work

• This paper presents a structured view on recent research in the field of alternative methods for scholarly impact.
• With the help of a morphological box we structured the several facets of these studies.
• Moreover, the morphological box as a creativity technique shows new paths for establishing further research.
• Especially the Semantic Web with its enhanced semantic capabilities can help to overcome one of the major points of criticism of citation based metrics – the lack of interpretation of the meaning of a citation.
• The Semantic Web with its typed link structure can be a solution for that.
• These opportunities combined with Web2.0 sensitive impact interpretation of scholarly results can open the horizon for a new, improved Scientometrics.
Thanx for your attention!!
Personal Invitation to Collnet 2014

10th International Conference on Webometrics, Informetrics & 15th Collnet Meeting
03.-05.09.2014
Ilmenau University of Technology
www.tu-ilmenau.de/collnet2014

Germany