Altmetrics and Revolutions:

New Products, New Metrics, and the New Era of Web-native Scholarship.

bit.ly/sla-altmetrics
@jasonpriem

June 10, Special Libraries Association Mtg,
San Diego, CA
What you do matters so much.
Communication is the soul of science.
<table>
<thead>
<tr>
<th>tech</th>
<th>pen and paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>products</td>
<td>letter</td>
</tr>
<tr>
<td>filters</td>
<td>personal</td>
</tr>
</tbody>
</table>
In 1665, the first revolution:

Oldenburg publishes Phil. Trans; applies the best available technology (printing press) to vastly improve dissemination.
A step forward: the journal

<table>
<thead>
<tr>
<th>tech</th>
<th>The Letter</th>
<th>The Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>products</td>
<td>pen and paper</td>
<td>printing press</td>
</tr>
<tr>
<td>filters</td>
<td>letter</td>
<td>article</td>
</tr>
<tr>
<td></td>
<td>personal</td>
<td>peer-review</td>
</tr>
</tbody>
</table>
A second revolution is coming

<table>
<thead>
<tr>
<th></th>
<th>The Letter</th>
<th>The Journal</th>
<th>The DcJ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>tech</strong></td>
<td>pen and paper</td>
<td>printing press</td>
<td>the web</td>
</tr>
<tr>
<td><strong>products</strong></td>
<td>letter</td>
<td>article</td>
<td>everything</td>
</tr>
<tr>
<td><strong>filters</strong></td>
<td>personal</td>
<td>peer-review</td>
<td>altmetrics</td>
</tr>
</tbody>
</table>
But journals are already online!

Your revolution is over, sir!

"The Digital Publishing Revolution Is Over"
Online journals are paper journals delivered by faster horses.
The First Revolution promoted homogeneity of outputs.
The standardized article was born of the need for industrial-scale replication and interchangeability.
The Second Revolution will promote diversity of outputs.

With publication nearly free, it becomes trivial to capture the missing pieces of the scholarly record.
Instead of moving paper products faster, we can create web-native science.
data
Quantifying the Detection of Early Warning Signals


- Author: Carl Boettiger
- License: CC0
- Project navigation
stories

fiction:
short story, novella, novel, series, play, film, comic book, etc, etc...

scholarship:
paper, monograph, video, blog posts, notebooks, infographics, slides, etc, etc
How many proofs that $\pi_n(S^n) = \mathbb{Z}$ are there?

Offhand I can think of two ways in classical homotopy theory:

1. Show that $\pi_k(S^n) = 0$ for $k < n$ by deforming a map $S^k \to S^n$ to be non-surjective, then contracting it away from a point not in its image. Now use the Hurewicz theorem to show $\pi_n(S^n) = H_n(S^n) = \mathbb{Z}$, which is easy to calculate with cellular homology.

2. Use the Freudenthal suspension theorem to induct up from $\pi_1(S^1) = \mathbb{Z}$, which you can prove using (say) the universal covering space $\mathbb{R} \to S^1$.

What other ways are there to prove $\pi_n(S^n) = \mathbb{Z}$?
Examples: Twitter

In one month, over 58k citations from Twitter to scholarly articles (citwaitions?)

It is like having a jury preselect what will probably interest you…. Occasionally there will be something that people will link to, and it will change what I think, or what I’m doing, or what I’m interested in.

-study participant
(Priem and Costello, 2010)
Examples: Twitter

Cumulative growth in the number of scholarly Twitter accounts.

*faculty*

*nonfaculty*

Each dot represents one tweet; each horizontal line of tweets is from one user. The left edge of each line is when the user joined Twitter.

(Priem, Costello, and Dzuba 2011)
Web-native science means we can start making public, not merely "Publishing."
Here's my journal:
Here's how I publish:

The Decoupled Journal article: a case study.

This is a pre-publication copy, submitted to the Frontiers in Computational Neuroscience special issue “Beyond open access: visions for open evaluation of scientific papers by post-publication peer review” and posted here with permission. It is still under peer review. Comments, ideas, and arguments are welcome.
But how do we filter? How do we measure?
Don't turn off the taps,
Build boats.
The old way: countin' citations
And that's awesome!
But citations only tell part of the story

Spotting emerging research fronts will require tracking "formal and informal communication" (Kuhn, 1977)

Heart of scholarly communication is "visits, personal contacts, and letters." (Bernal, 1944)

"...there are undoubtedly highly useful journals that are not cited frequently." (Garfield, 1972)
Impact has multiple dimensions:

**Audience:** scholars, public

**Engagement type:** views, discussion, saves, citation, recommendation
Impact has multiple dimensions:

<table>
<thead>
<tr>
<th>scholarly</th>
<th>public</th>
</tr>
</thead>
<tbody>
<tr>
<td>recommended</td>
<td></td>
</tr>
<tr>
<td>cited</td>
<td></td>
</tr>
<tr>
<td>discussed</td>
<td></td>
</tr>
<tr>
<td>saved</td>
<td></td>
</tr>
<tr>
<td>scholarly</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scholarly</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>recommended</strong></td>
<td></td>
</tr>
<tr>
<td><strong>cited</strong></td>
<td>traditional citation</td>
</tr>
<tr>
<td><strong>discussed</strong></td>
<td></td>
</tr>
<tr>
<td><strong>saved</strong></td>
<td></td>
</tr>
<tr>
<td><strong>scholarly</strong></td>
<td></td>
</tr>
</tbody>
</table>
Altmetrics measures impact:

<table>
<thead>
<tr>
<th></th>
<th>scholarly</th>
<th>public</th>
</tr>
</thead>
<tbody>
<tr>
<td>recommended</td>
<td>faculty of 1000</td>
<td>popular press</td>
</tr>
<tr>
<td>cited</td>
<td>traditional citation</td>
<td>wikipedia</td>
</tr>
<tr>
<td>discussed</td>
<td>scholarly blogs</td>
<td>blogs, twitter</td>
</tr>
<tr>
<td>saved</td>
<td>mendeley, citeulike</td>
<td>delicious</td>
</tr>
<tr>
<td>scholarly</td>
<td>pdf views</td>
<td>html views</td>
</tr>
</tbody>
</table>
Bibliometrics mined impact on the first scholarly Web.

altmetrics mines impact on the next one.
ImpactStory.

An open-source, nonprofit startup to gather and share altmetrics.

Heather Piwowar
Jason Priem
article

Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches
(2006) Smith, Beaulieu, Donoghue  *BMC Evol Biol*

Computational toxicology using the OpenTox application programming interface and Bioclipse.
(2011) Willighagen, Jeliazkova, Hardy et al.  *BMC research notes*

Multistep correlations via covariance processing of COSY/GCOSY spectra: opportunities and artifacts

Automated dielectric single cell spectroscopy - temperature dependence of electrorotation

Pyramid symmetry transforms: From local to global symmetry

dataset

Data from: Data archiving is a good investment
Highly saved by scholars

This item has 6 Mendeley readers. That's better than 76% of items indexed by Web of Science in 2008, suggesting it's highly saved by scholars. Click to learn more.
Why altmetrics?

- Fairer evaluation.
- Faster evaluation.
- Etc, etc.
- Ultimately, most important: we can fundamentally change the way we filter research.

This changes the game. And the game needs to be changed.
Today’s journals are the best scholarly communication system possible using 17th-century technology.
We don’t use the Web.

Berners-Lee created the Web as a scholarly communication tool.

Today the Web has revolutionized everything but scholarly communication.
The network is the key

At web scale, the value isn't in manual curation...

(ask these guys)
It's in mining the network

(ask these guys)
What does the system do now?

Any modernization will have to do everything the current system does, better.

Journals have four “traditional functions.”
How does this system work?

Every journal does every function itself. Each produces the same product. Little variety, little choice:

publishing as a fixed-price menu.
The decoupled journal (DcJ)

Functions are offered as individual services; authors pick which ones they want:

Publishing a la carte.
A DcJ example

Peer-review stamp + aggregated comments
Author’s Twitter feed + mailing lists
Google + disciplinary search service
Typesetting

The product is stored and published from an institutional repository, and given a DOI as a unique ID.
The second revolution has started.

Once we have altmetric data, it’s too useful to ignore; alternative filters and even certification paths based on this data will open.

As Peter Vinkler says, citation graph data is like Chekhov’s gun: once on stage, it has to be fired.
A wise man, that Chekov.
Thanks!

Advisors:
- Brad Hemminger,
- Todd Vision

Funders:
- Alfred P. Sloan Foundation
- DataONE
- Dryad
- National Science Foundation
- Open Society Foundations
- Royster Society of Fellows
Questions?

Jason Priem  @jasonpriem, http://jasonpriem.org

This is a living document; if a particular version is important to you, make sure you link to that revision.