What is Impact? Bibliometrics for Editors

ISMTE

11 August 2017 • Denver, CO

Jeno Mavzer
Coronis Group
orcid.org/0000-0002-0212-0773

Roadmap

Key Concepts & History

Databases

- Web of Science
- Scopus
- Google Scholar

Types of Metrics

- Journal-level (Impact Factor, CiteScore, SJR)
- Article-level (Citation counts, RCR, Altmetrics)
- Author-level (h-index, i10-index, m-para)

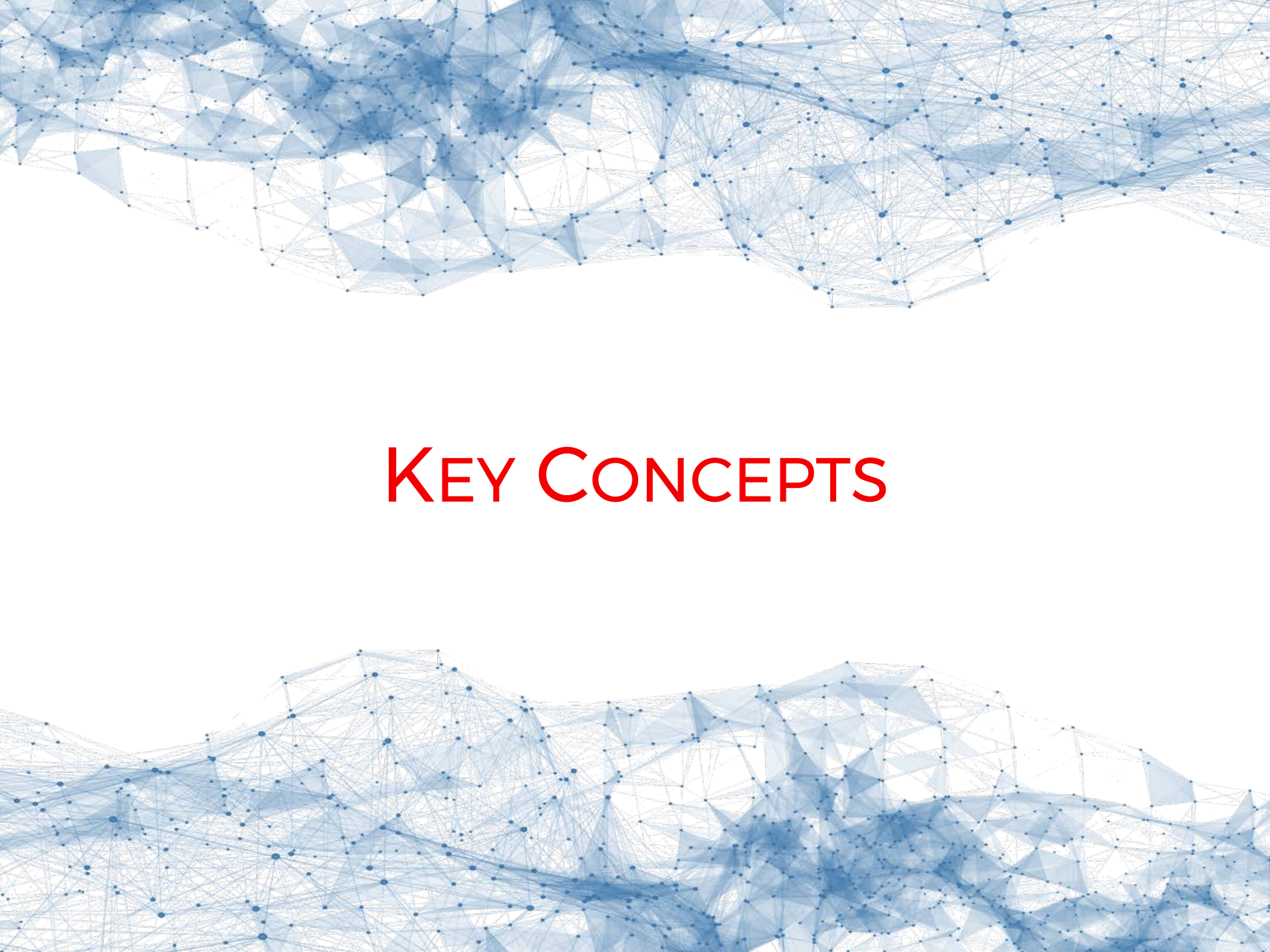
Citation Analysis

- Citation Peaks
- Networks of Influence

Session Objectives

Following this session, you will:

- Be familiar with a range of metrics available for evaluating journal and content performance.
- Emerge with some guidelines to help you select appropriate sets of metrics for different cases.
- Know how to translate bibliometric data findings to actionable editorial strategy.


KEY CONCEPTS

CITATIONS

Inflated, conflated, and often manipulated.

'Game of Thrones': The major battle scenes, ranked

Washington Post · 20h ago

RELATED COVERAGE

Game of Thrones: Daenerys' Stand-In Is the Spitting Image of the Khaleesi Herself

Highly Cited · People Magazine · Apr 22, 2015

Giant dinosaur slims down... a bit

BBC News · 2h ago

RELATED COVERAGE

New giant titanosaur from Patagonia | Proceedings of the Royal Society of London B: Biological Sciences

Most Referenced · Proceedings of the Royal Society B - Journals · 7h ago

How do you define impact?

Number of publications?
Number of citations?
Rate of citation growth?
Recognition of peers? Prizes?
Career progression? Early tenure?
Practical application of findings?

$$F = \frac{m(v-u)}{\Delta t}$$

?

Citations and Impact

Why it has become more difficult to predict Nobel Prize winners: a bibliometric analysis of Nominees and Winners of the Chemistry and Physics Prizes (1901–2007)

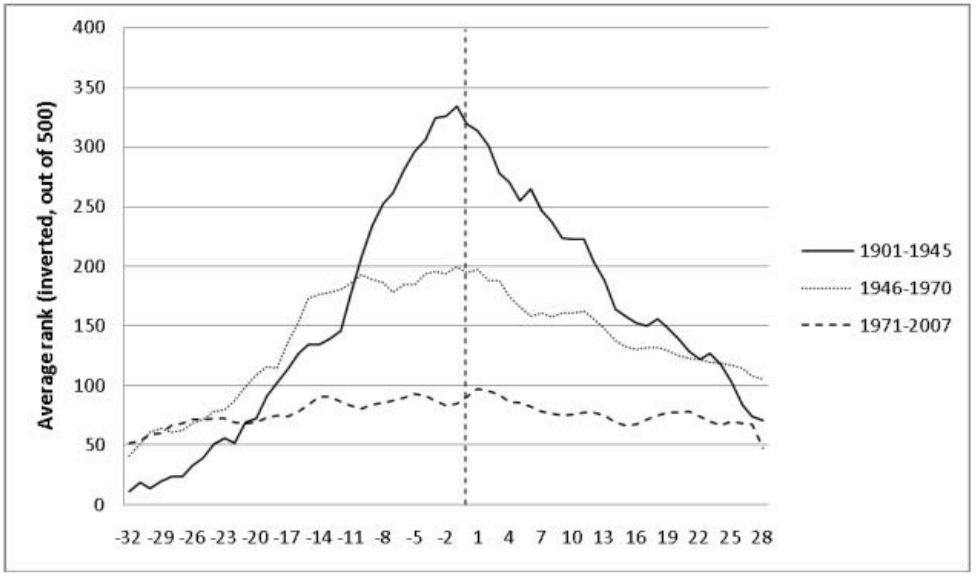


Figure 3 : Physics prizewinners' citation rankings, averaged over all years for three different periods. Once again, the vertical dashed line represents the year "0", when the Prize is won.

Source: arXiv:0808.2517v1 [physics.soc-ph]

Metrics

Bibliometrics

- Origin in print
- Citation-based
- Lagging indicators

Altmetrics

- Digital genesis
- Data derived from social media
- Immediate

What are Bibliometrics?

"...the application of mathematics and statistical methods to books and other media of communication"

— Alan Pritchard (1969)

Source: *Journal of Documentation*. 1969 Dec;25(4):348-349.

Bibliometrics Provide

Tools to track articles, authors, organizations, funders

A statistical approach to normalize publication performance against that of its cohort

Analysis of structure and dynamics of the field

Rationale for design of editorial processes and policies

Adapted from:
NIH Library https://nihlibrary.nih.gov

Citation Index

- Citation indexes for science; a new dimension in documentation through association of ideas. (*Science*. 1955;122(3159):108-11)
- Institute for Scientific Information (ISI) established in 1961 – began to publish Science Citation Index.
- Journal Citation Reports by Clarivate (formerly Thomson Reuters)

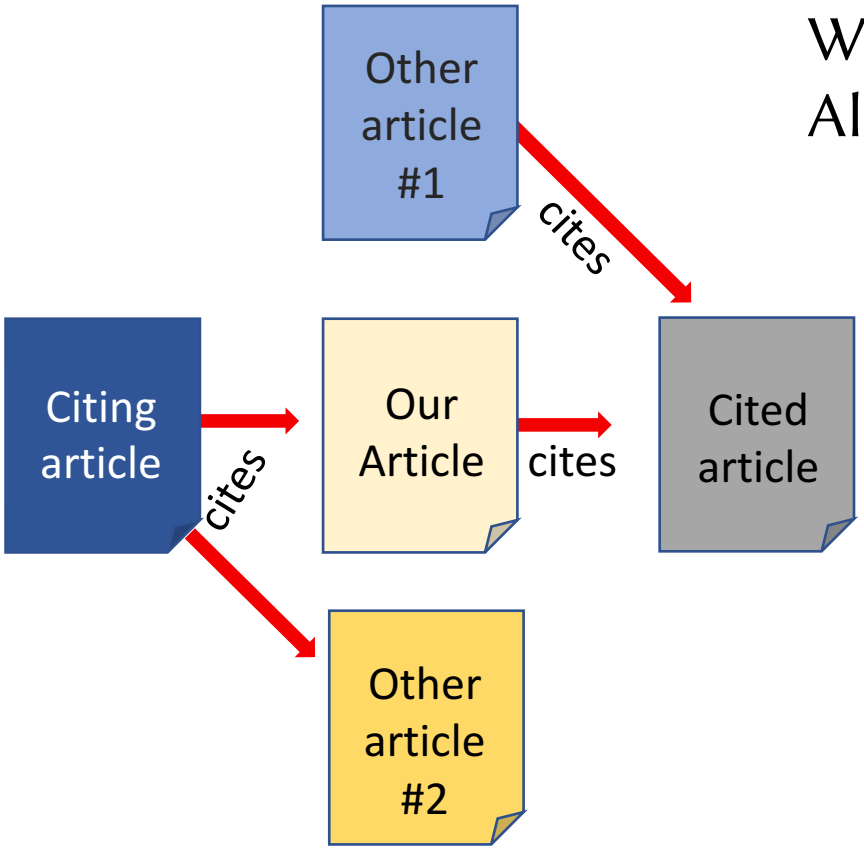
Impact Factor

"Like nuclear energy, the impact factor is a mixed blessing. I expected it to be used constructively while recognizing that in the wrong hands it might be abused."

—Eugene Garfield (1999)

Source: *CMAJ.* 1999;161(8): 979-980

What is Citation Analysis?



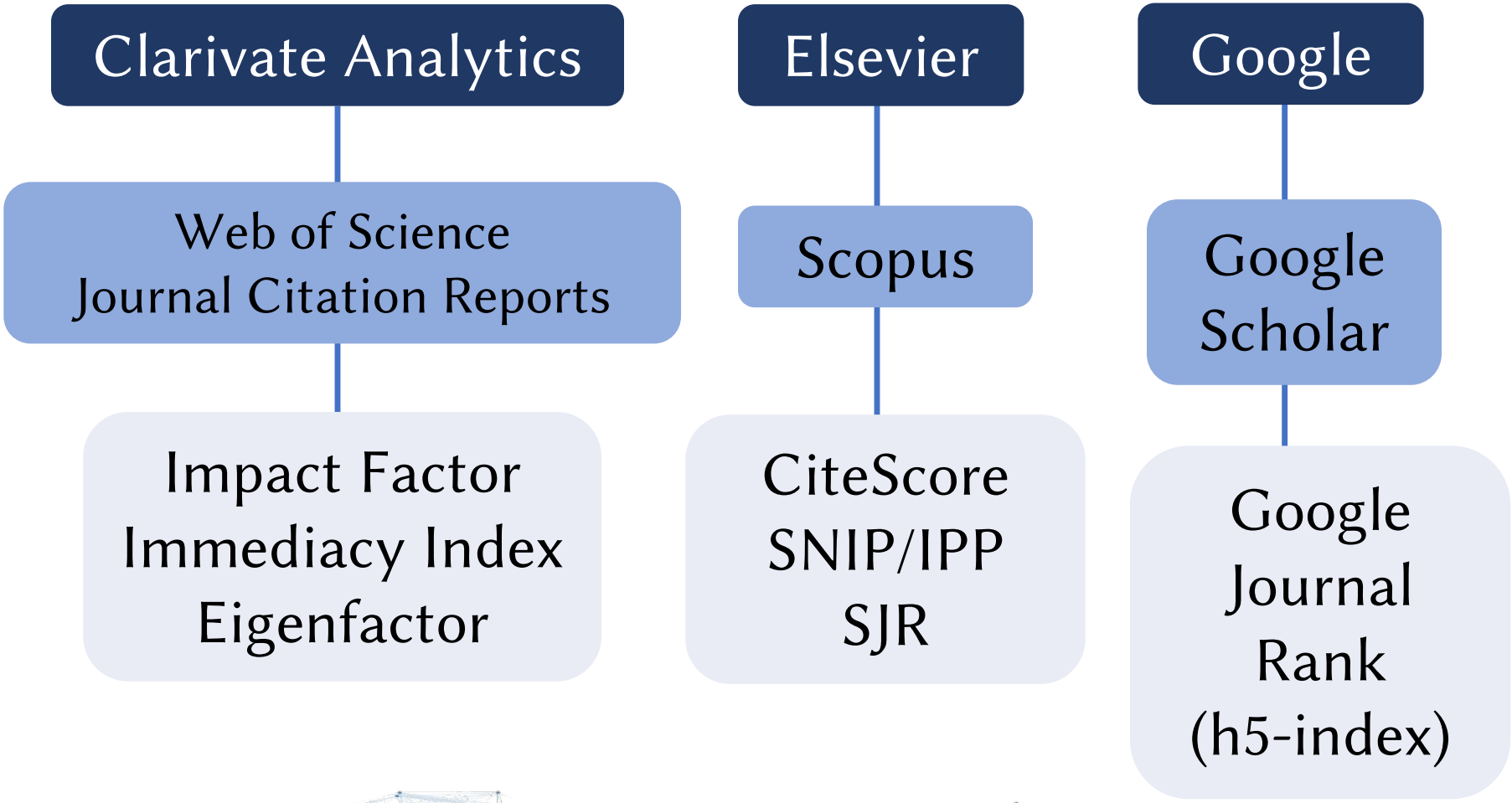
Who is citing what and when? Also, why?

Our Article & Other Article #1 are **bibliographically coupled** (i.e. by virtue of the article we both cite)

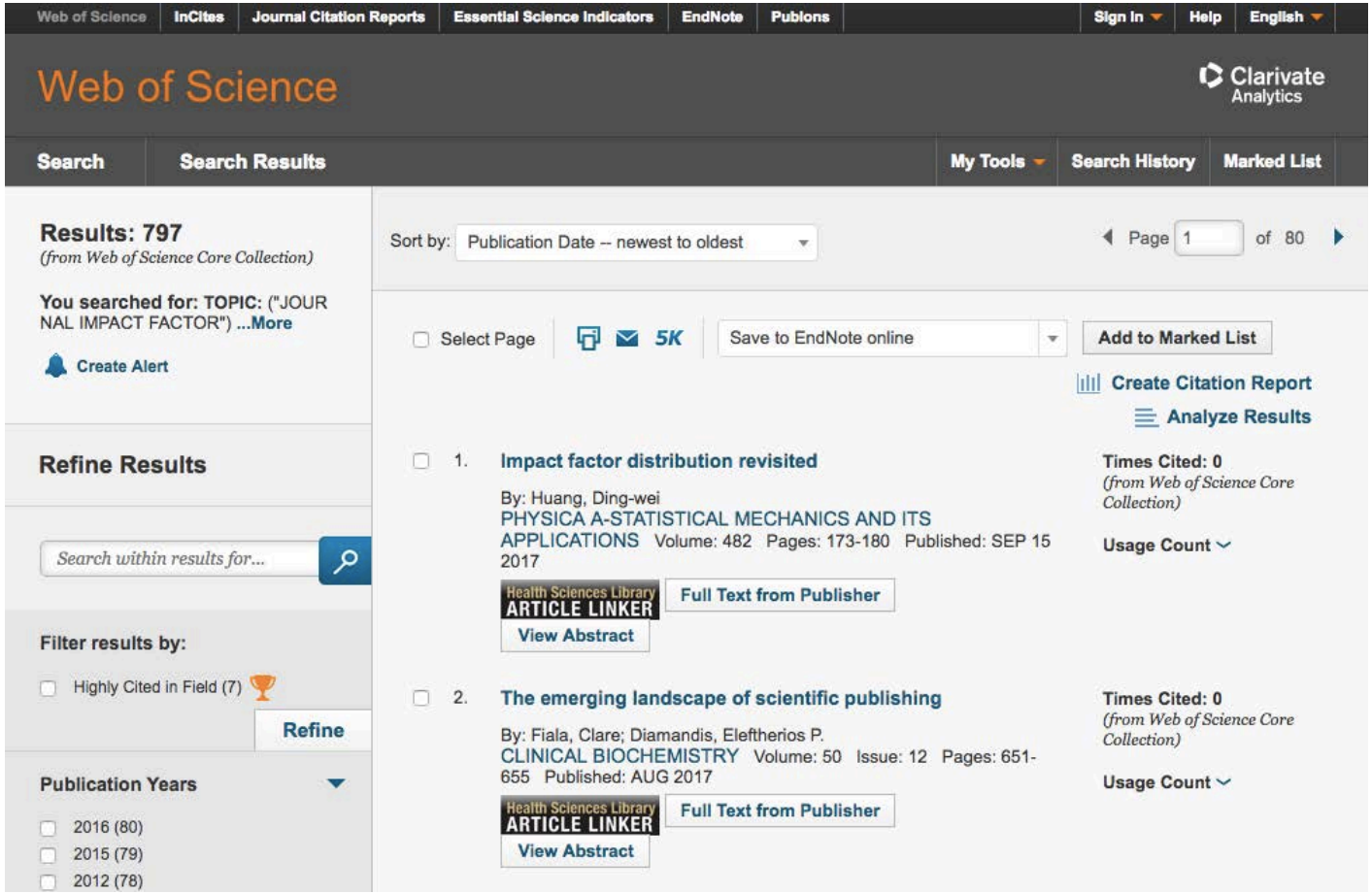
Our Article & Other Article #2 are **co-cited** (i.e. in the article that cited us both)

Data Sources

Databases

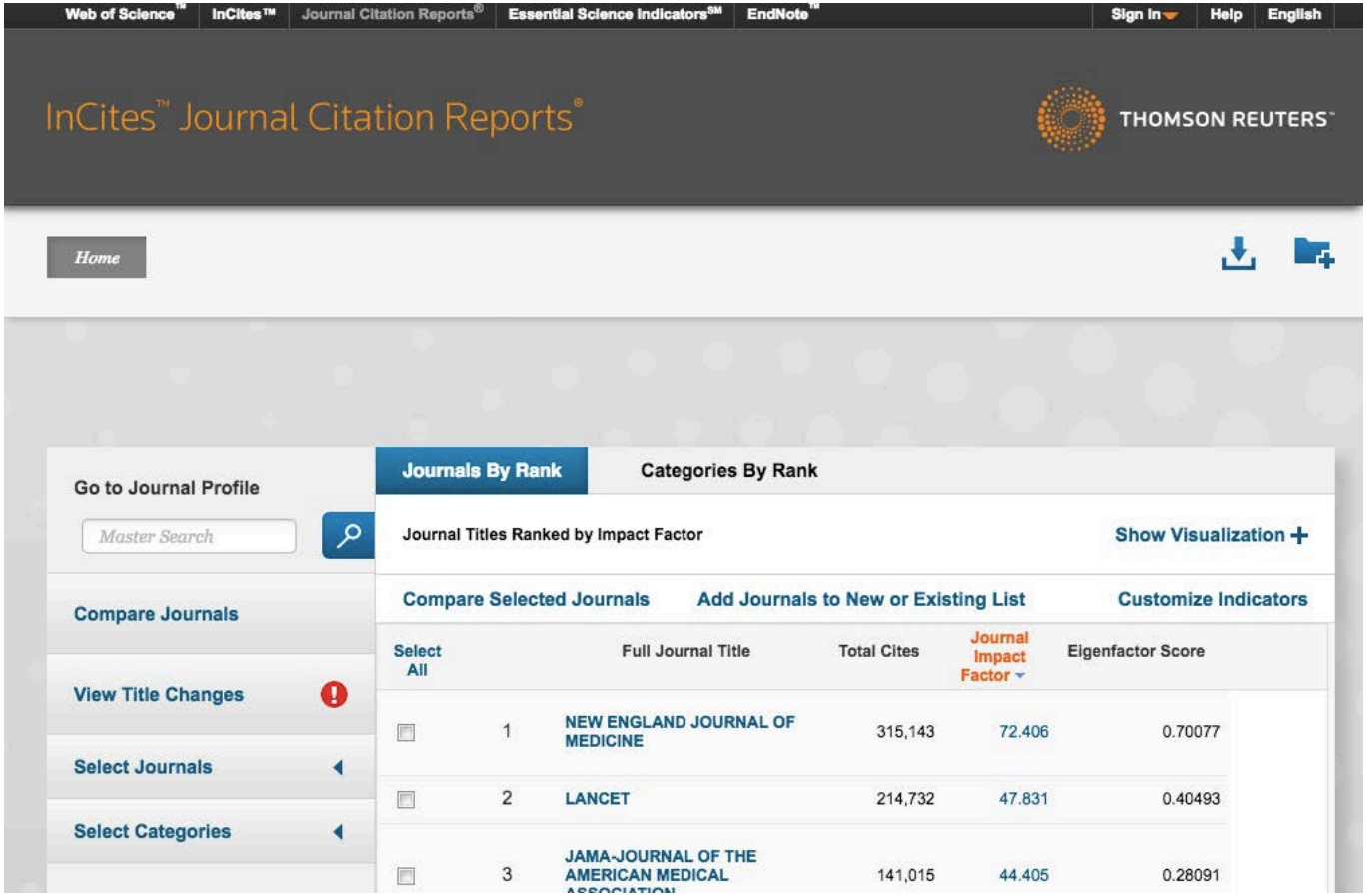

Clarivate Analytics
Web of Science
Journal Citation Reports
Impact Factor
Immediacy Index
Eigenfactor
Elsevier
Scopus
CiteScore
SNIP/IPP
SJR
Google
Google
Scholar
Google
Journal
Rank
(h5-index)

Web of Science



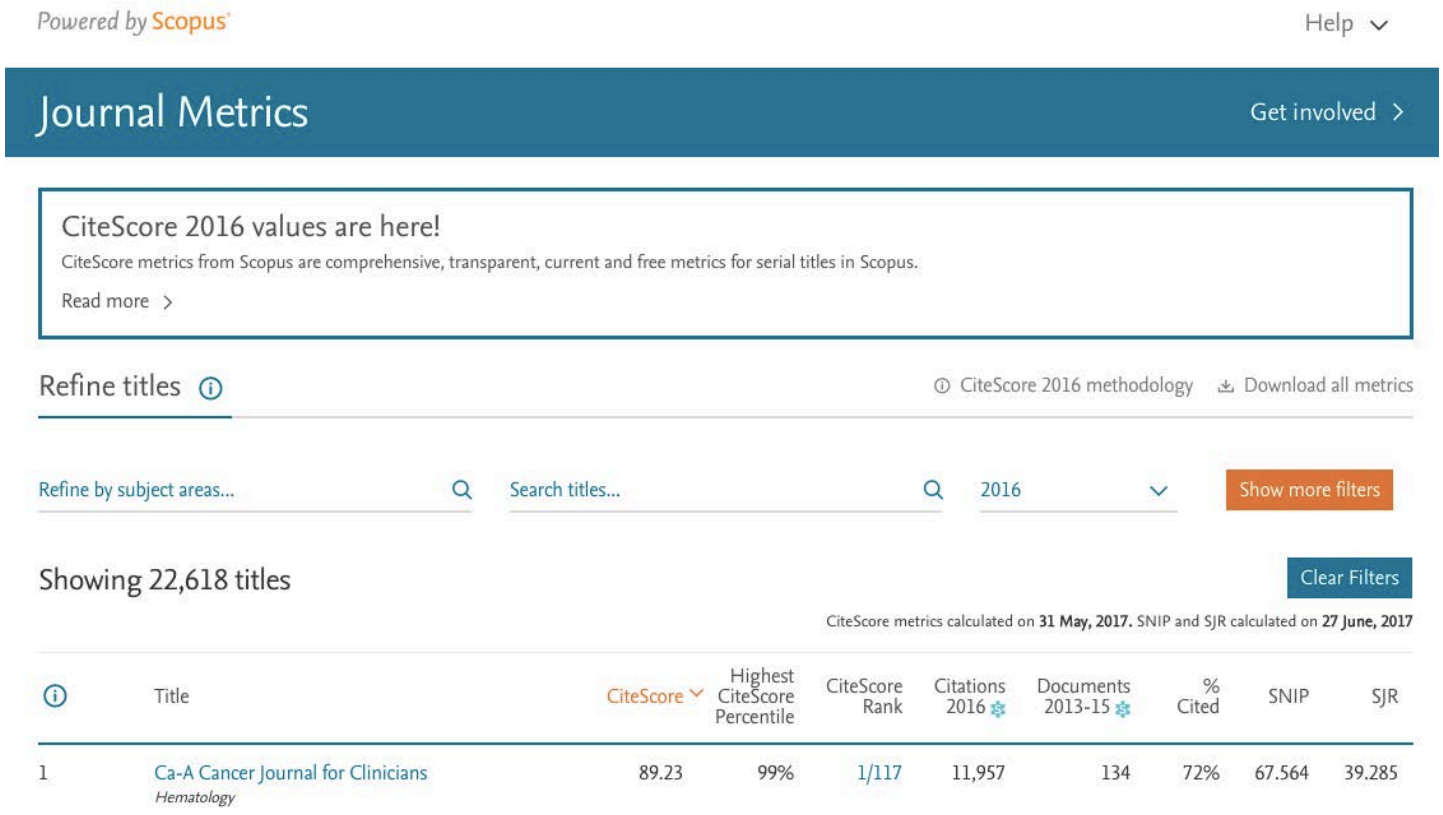
https://www.webofknowledge.com/

Journal Citation Reports



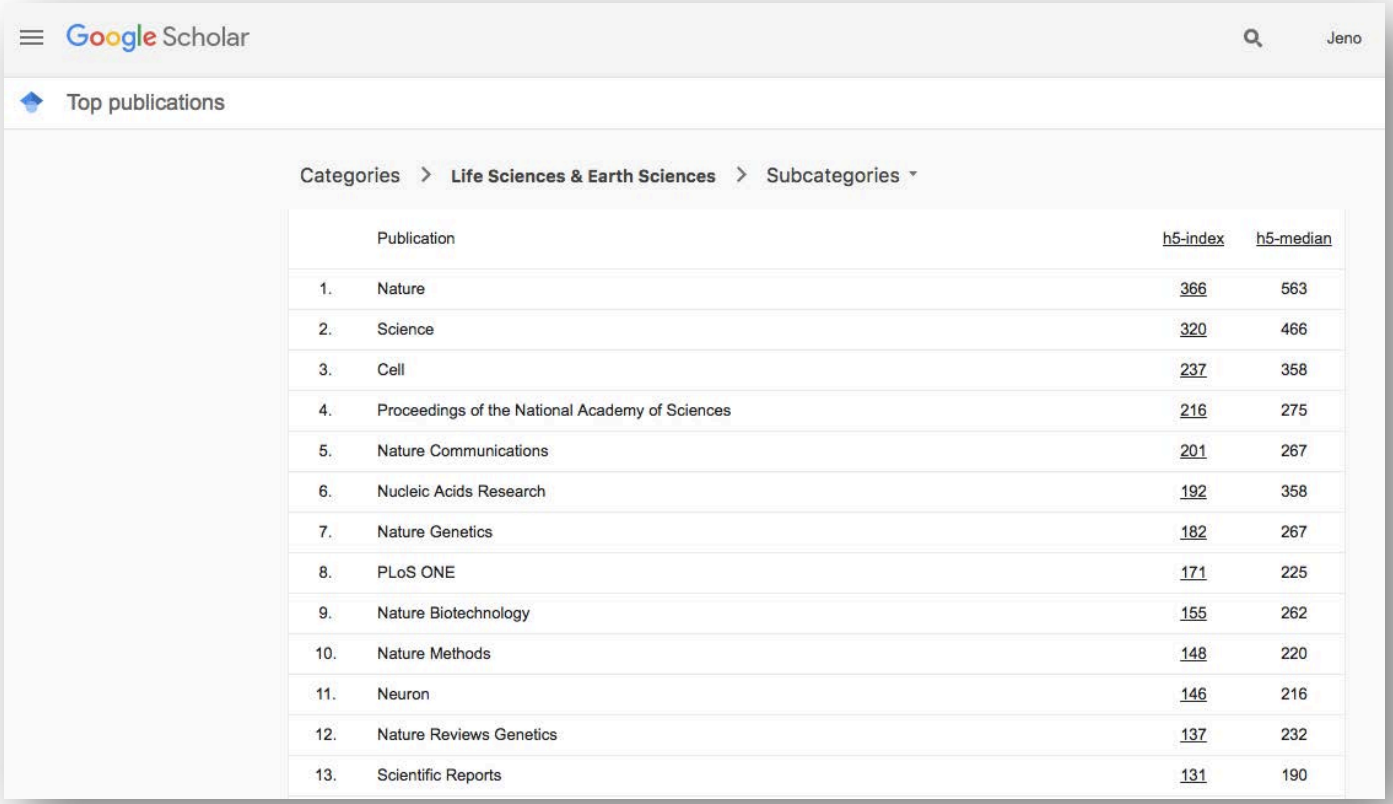
https://jcr.incites.thomsonreuters.com/

SCOPUS



https://journalmetrics.scopus.com/

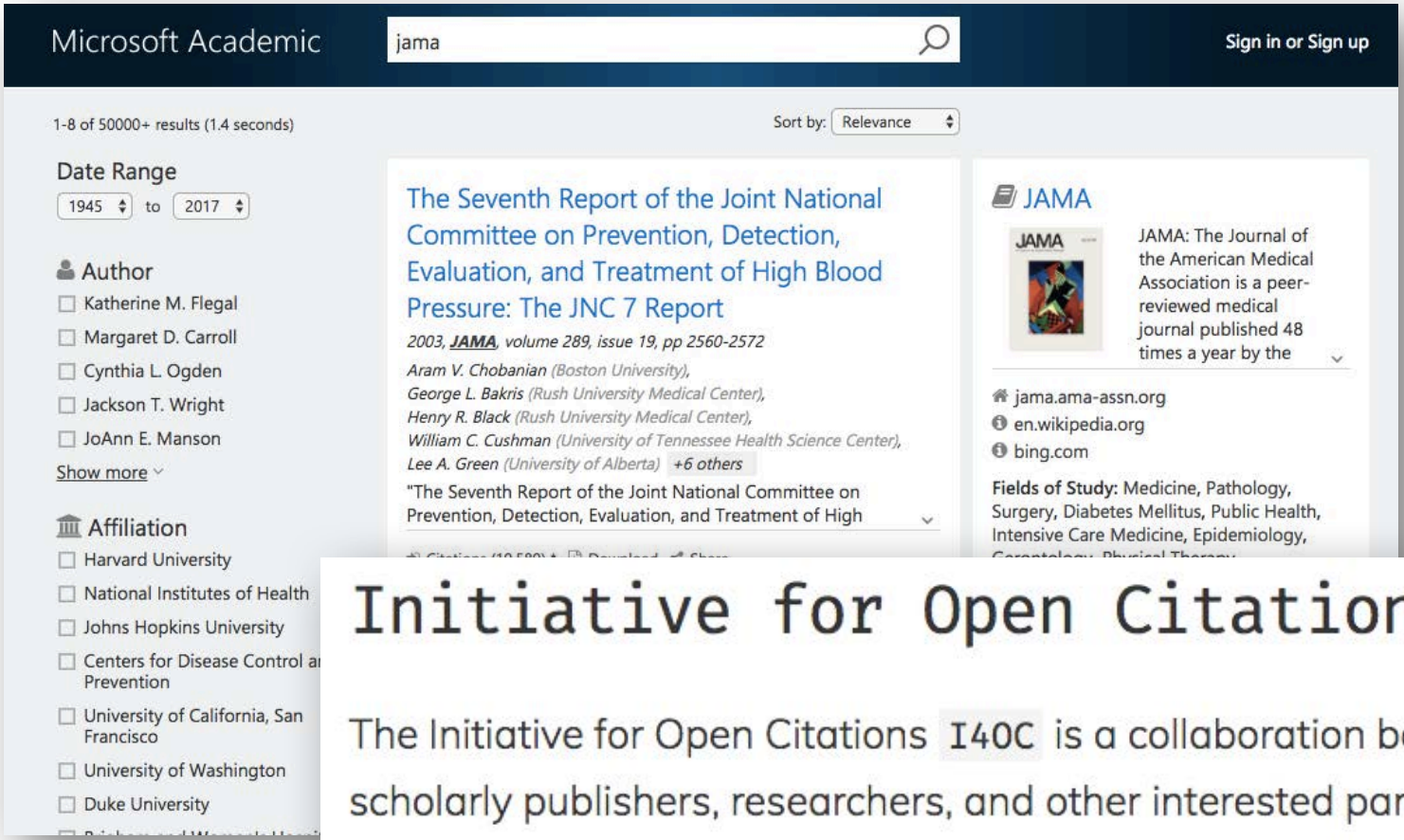
Google Scholar



	Publication	h5-index	h5-median
1.	Nature	366	563
2.	Science	320	466
3.	Cell	237	358
4.	Proceedings of the National Academy of Sciences	216	275
5.	Nature Communications	201	267
6.	Nucleic Acids Research	192	358
7.	Nature Genetics	182	267
8.	PLoS ONE	171	225
9.	Nature Biotechnology	155	262
10.	Nature Methods	148	220
11.	Neuron	146	216
12.	Nature Reviews Genetics	137	232
13.	Scientific Reports	131	190

https://scholar.google.com/citations?view_op=top_venues /

EMERGING…



Initiative for Open Citations

The Initiative for Open Citations `I4OC` is a collaboration between scholarly publishers, researchers, and other interested parties to promote the unrestricted availability of scholarly citation data.


Metrics

“Journal performance is a complex, multidimensional concept that cannot be fully captured in one single metric.”

—Henk F. Moed et al. Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way. *Scientometrics*. 2012;92(2):367–376

Metrics for All

- **Researchers:** Metrics focused on individual scholarly contribution
- **Editors & Publishers:** Metrics focused on the journals that produce individual scholarly contributions.
- **Administrators:** Metrics focused on research output over time.
- **Various:** Metrics focused on group, institutional, or national output over time.

Metrics for Editors

- **Field-level:** Aggregate Impact Factor, manual cohort clustering
- **Journal-level:** Impact Factor, Immediacy, Eigenfactor; CiteScore, SJR, SNIP, IPP; h5-index
- **Article-level:** Citation counts, RCR, Altmetrics
- **Author-level:** Citation counts, h-Index, i10-Index, m-Index

Impact Factor

How is it calculated?

2016 Impact Factor =

$$\frac{\text{Number of cites in 2016 to all papers published in 2014 \& 2015}}{\text{Total number of } \textbf{citable} \text{ articles published in 2014 \& 2015}}$$

When is it calculated?

Annually, typically released mid-June

What does it mean?

Average frequency of citations to recent articles.

What is a "good" impact factor?

Difficult to say. IFs are highly field specific – one must compare a journal to its citation cohort to determine performance.

Source of data?

Web of Science citation data

IF Considerations

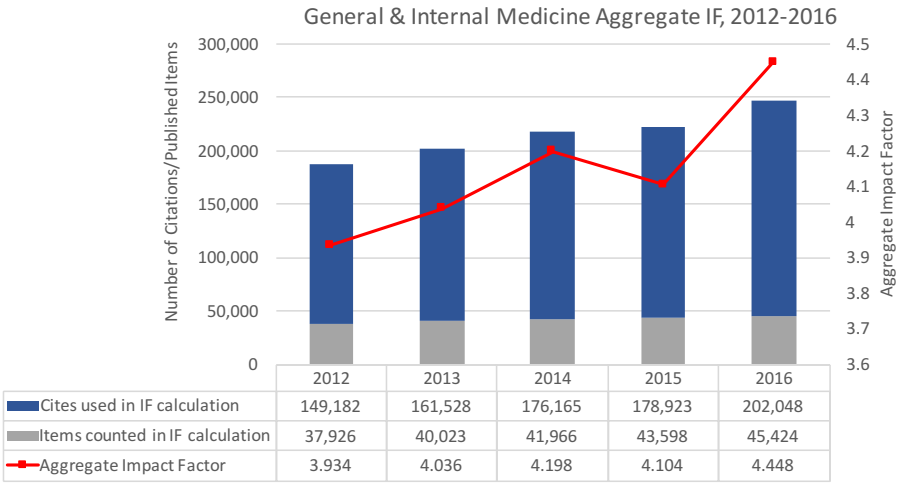
- Impact Factor can be affected significantly by a small number of papers in a journal (skewness).
- Editors can also manipulate the score – through both legal (frontloading, early online) and improper (coercive citation) methods.
- IFs should not be used to compare journals across disciplines.

"...impact factors don't tell us as much as some people may think about the respective quality of the science that journals are publishing."

"Not-so-deep impact"
Nature. 435:1003-1004
(23 June 2005)

Consider Field-Level Context

Aggregate Impact Factor (JCR Subject Category)

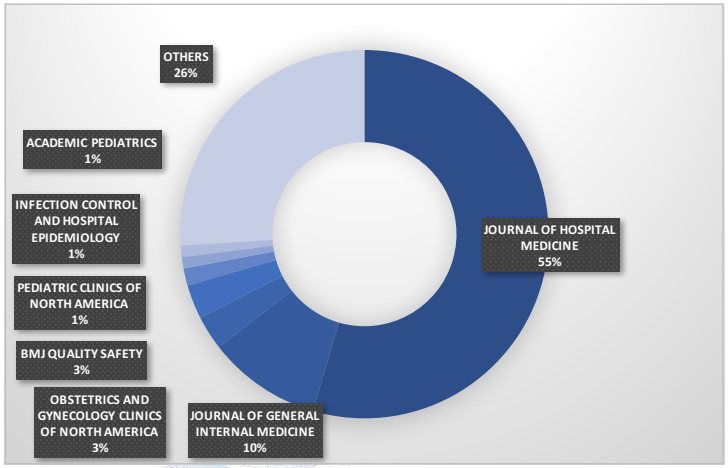


	2012	2013	2014	2015	2016
Cites used in IF calculation	149,182	161,528	176,165	178,923	202,048
Items counted in IF calculation	37,926	40,023	41,966	43,598	45,424
Aggregate Impact Factor	3.934	4.036	4.198	4.104	4.448

Individual journal impact factors should be considered in light of JCR subject category...

...as well as that of titles publishing within the same citation space.

Subfield Norms (Web of Science)



Immediacy Index

How is it calculated?

2016 Immediacy Index =

$$\frac{\text{Number of citations received in 2016}}{\text{Total number of articles published in 2016}}$$

When is it calculated?

Annually, alongside IF in mid-June

What does it mean?

The average number of times an article is cited in the year it is published.

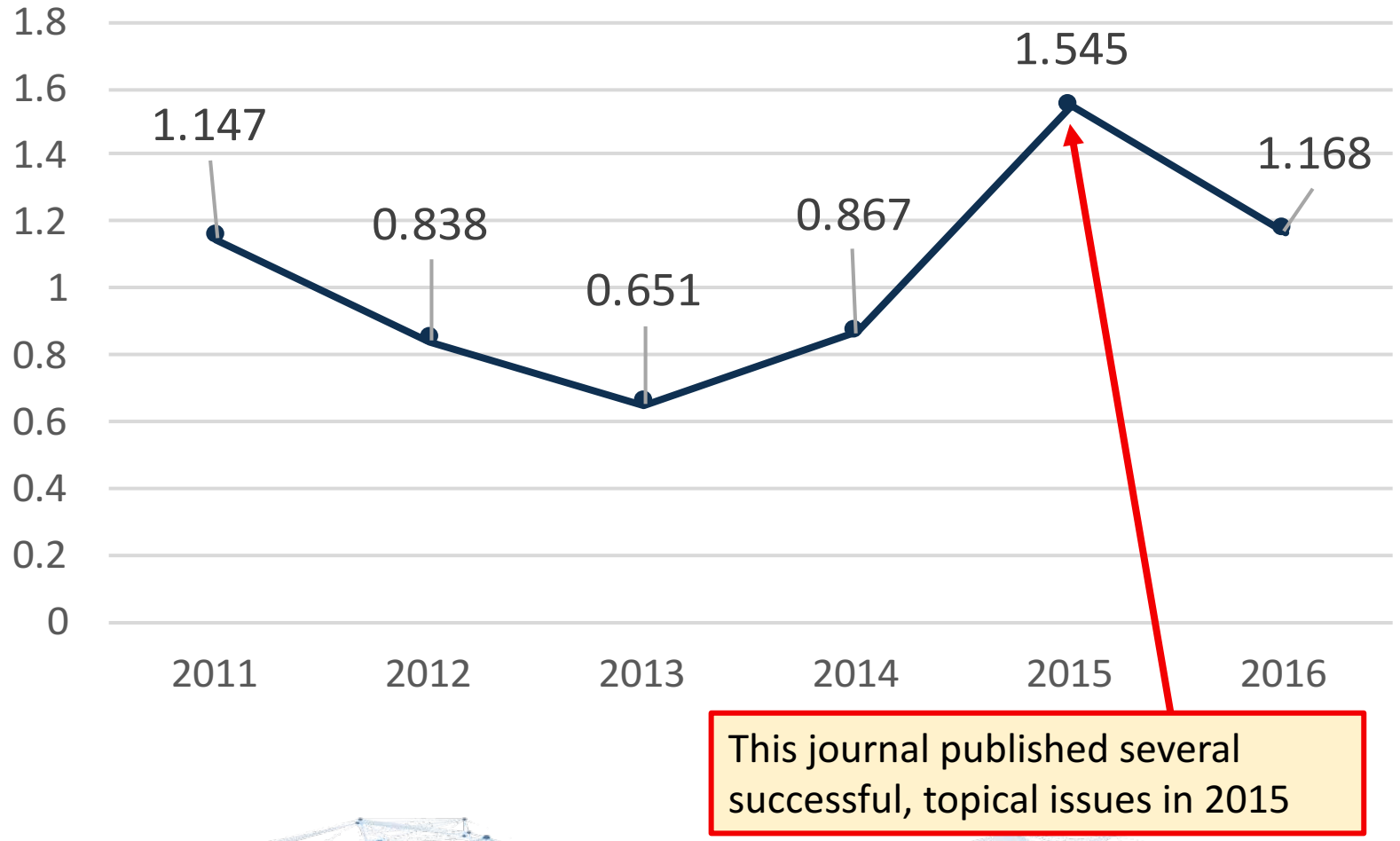
What editorial practices affect it?

Special issues published early in the year. Publishing content online ahead of print and engaging online readers.

Why does it really matter?

Spikes in immediacy can presage rise in IF one year later.

Immediacy Example



Eigenfactor & Article Influence

How are they calculated?

2016 Eigenfactor =

a) Constructs citation matrix consisting of 2016 citations *from* each journal in JCR *to* each journal (using articles from 2011-2015, not including self-citations).
b) Scale so that the sum of all journal scores is 100.

2016 Article Influence = $\frac{\text{Eigenfactor score for 2016}}{\text{Total number of articles published in 2016}}$

What do they measure?

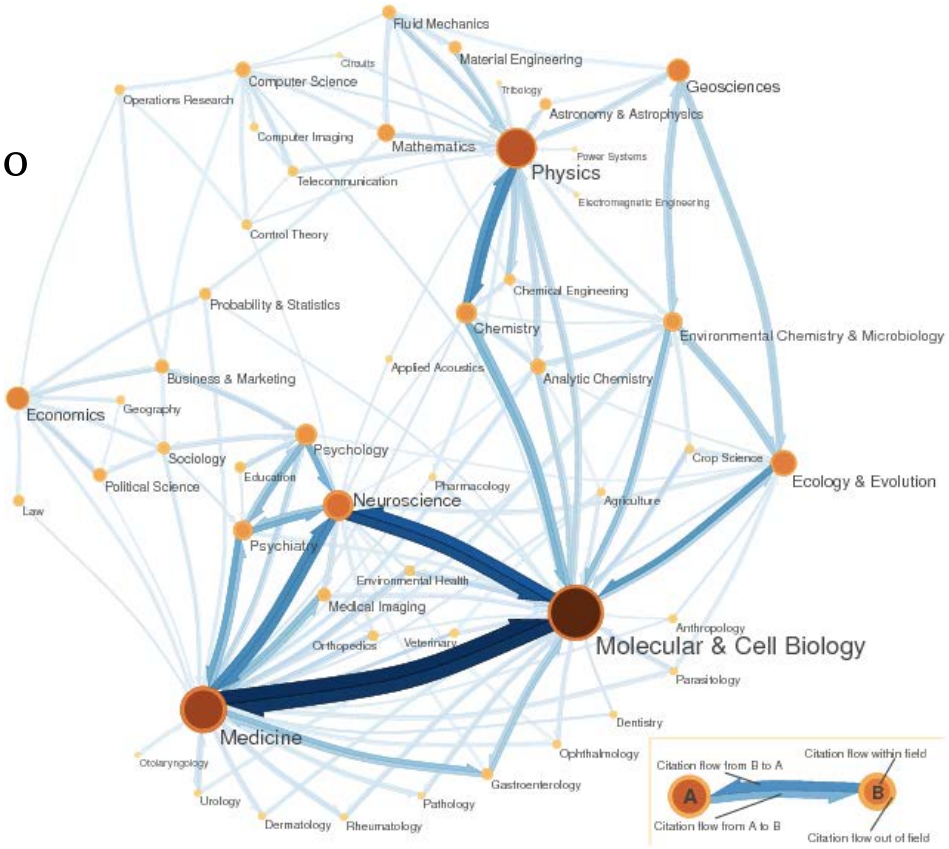
Eigenfactor measures relationships between journals within the scientific community (defined as all journals in JCR). Intended to reflect the influence and prestige of journals.

Article influence measures the average influence, per article, of the papers published in a journal. An Article Influence score greater than 1.00 indicates that the articles in a journal have an above-average influence.

Eigenfactor Relationships

- Eigenfactor's algorithm uses the structure of the entire network to evaluate the importance of each journal, regardless of discipline. Self-citations are excluded.

- This corresponds to a mathematical model mimicking the action of a reader following chains of citations as they move from journal to journal.



Demo: WoS/JCR Metrics

Web of Science: http://login.webofknowledge.com
Eigenfactor: http://eigenfactor.org/

CiteScore

How is it calculated?

2016 CiteScore =

$$\frac{\text{Number of cites in 2016 to articles published 2013-2015}}{\text{Total number of } \textbf{all} \text{ articles published 2013-2015}}$$

When is it calculated?

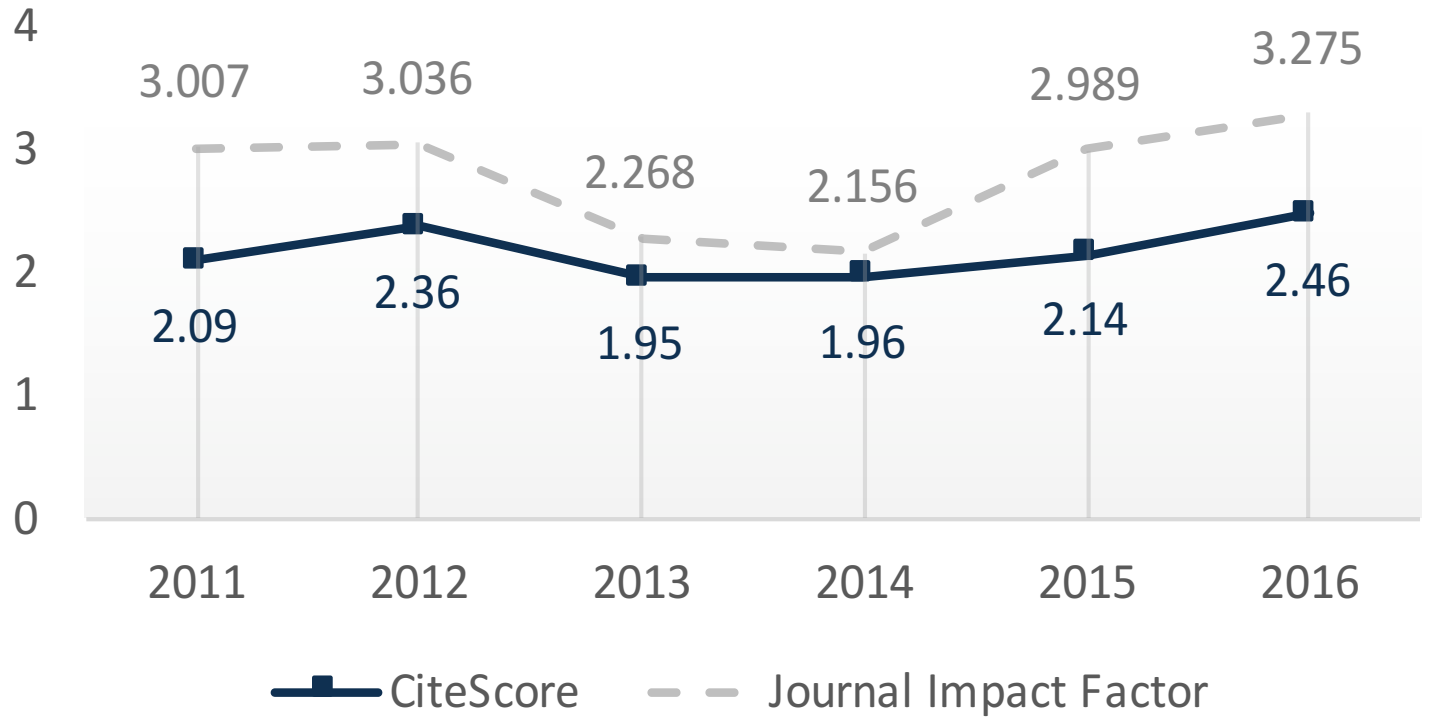
Monthly updates until annual final score (released in late May).

How does it differ from IF?

- Calculated from Scopus journal list, which includes more social sciences and humanities journals.
- Uses 3-year citation window, rather than the 2-year window of the Impact Factor.
- Does not differentiate between article types in denominator – eliminates advantage of publishing large number of editorials, news items, etc
- Free

CiteScore vs IF

This journal publishes a fair number of “noncitable” items (editorials, commentary, letters). The boost seen in IF is not captured by CiteScore. Likewise, pure research journals tend to have CiteScores higher than their IF.



SciMago Journal Rank

How is it calculated?

2016 SJR =

Measures weighted citations received in 2016 to documents published in a journal in years 2013-2015.

Cool. What does that mean?

SJR is a prestige metric that assigns relative scores to all of the sources in a citation network. Similar to Article Influence, a citation from a journal with a high SJR is "worth" more.

How does it differ from Eigenfactor?

- **Size of source network:** SJR is based on Scopus, which contains more journals than WoS, especially in social sciences, engineering, and arts.
- **Citation window:** SJR uses three years of citation data, vs Eigenfactor's five-year window.
- **Self-citation:** SJR allows self-citation, within limits of 1/3 all incoming cites. Eigenfactor eliminates self-citation.

SNIP & IPP

What are these?

2016 Impact per Publication (IPP) =

Number of citations in 2016 to papers published 2013-2015
Number of papers published 2013-2015

2016 Source Normalized Impact per Publication (SNIP) =

SNIP measures actual 2016 citations received relative to citations expected for the serial's subject field

Where are these located?

SNIP is included alongside CiteScore. It is also calculated annually by Leiden University's Centre for Science and Technology Studies, which also curates IPP.

How are these useful?

IPP is essentially CiteScore, minus a few article types. SNIP accounts for differences in citing behavior across fields (the longer the reference list of a citing publication, the lower the value of a citation from it).

Demo: Scopus Metrics

CiteScore: https://journalmetrics.scopus.com/
SJR: http://www.scimagojr.com/index.php
SNIP/IPP: http://www.journalindicators.com/indicators

Google Scholar Journal Ranking

What metric(s) does Google use to rank journals?

Google Scholar relies on h5-index and h5-median to rank journals. Top-20 lists are provided by subject category.

How are these metrics calculated?

Algorithmically (citation counts provided by computer program).

h5-index is the h-index for articles published in the last 5 complete years. It is the largest number h such that h articles published in 2012-2016 have at least h citations each

h5-median for a journal is the median number of citations for the articles that make up its h5-index.

Google Scholar Example

Categories > Chemical & Material Sciences > **Chemical & Material Sciences (general)**

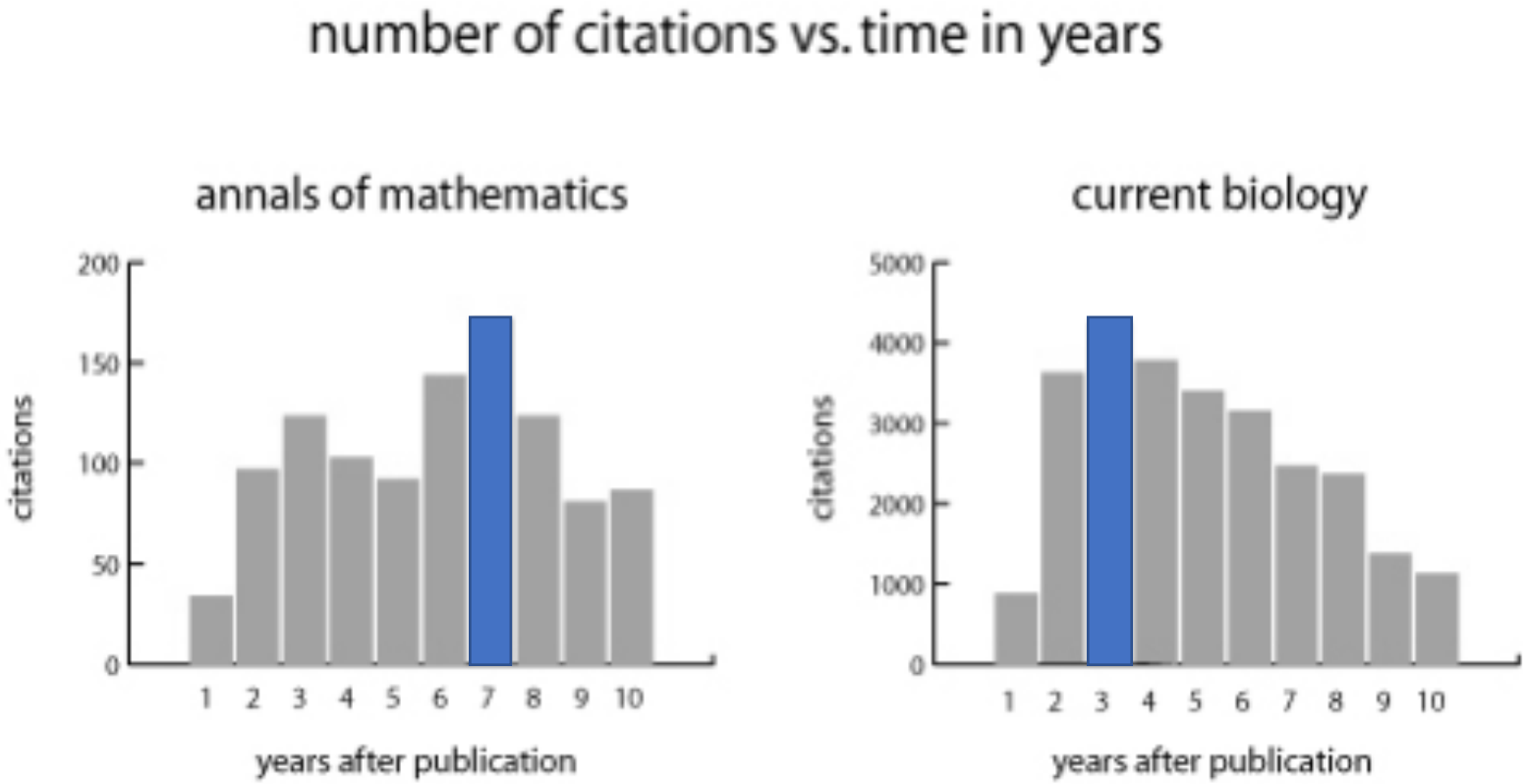
	Publication	h5-index	h5-median
1.	Chemical Society reviews	241	346
2.	Journal of the American Chemical Society	219	288
3.	Chemical Reviews	211	339
4.	Accounts of Chemical Research	150	206
5.	Chemical communications (Cambridge, England)	137	171
6.	Nature Chemistry	114	169
7.	The Journal of Physical Chemistry Letters	109	157
8.	Physical chemistry chemical physics: PCCP	106	136
9.	Chemical engineering journal	93	118
10.	Chemistry-A European Journal	92	115

Citations

"In the construction and interpretation of journal citation measures it is crucial to take into account differences in communication and citation practices between research fields."

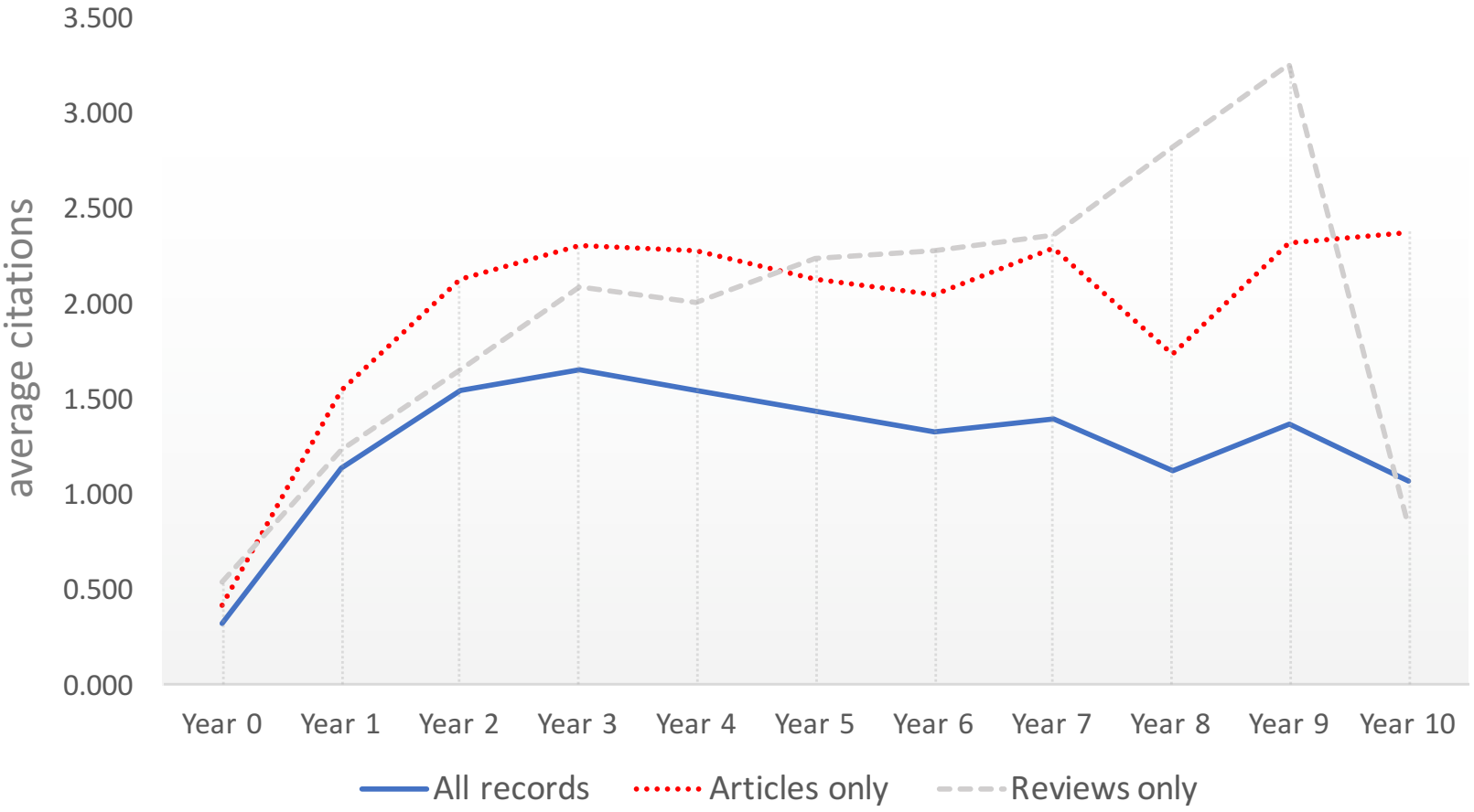
—Moed et al. *Scientometrics*. 2012;92(2):367–376

Citation Peaks – Journals



Source: Eigenfactor.org

Citation Peaks – Article Types

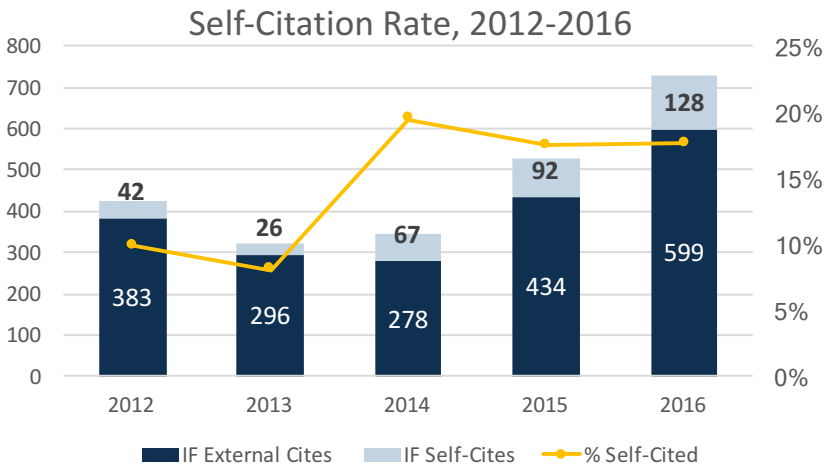


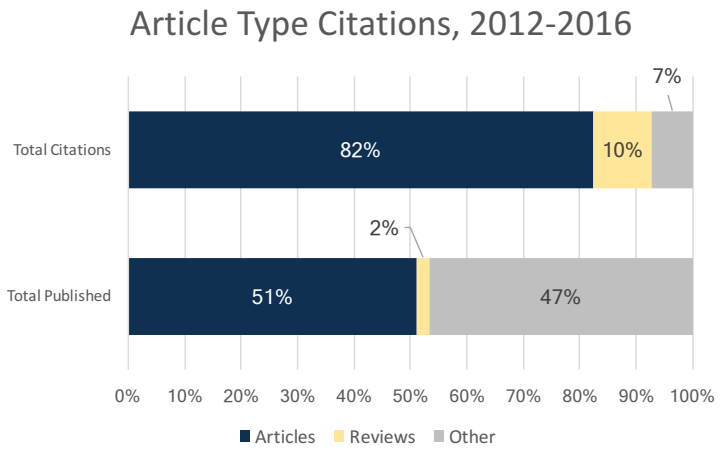
Article-Level Citations

Web of Science & Scopus:

What? Examine citation rates through time. Partition by article type to gain resolution on the breakdown of citations. Identify highly- or never-cited content.

How? Download citation sets and sort by parameter of interest. Can also retrieve full bibliographic records for visualization of larger patterns.





Article-Level Citations

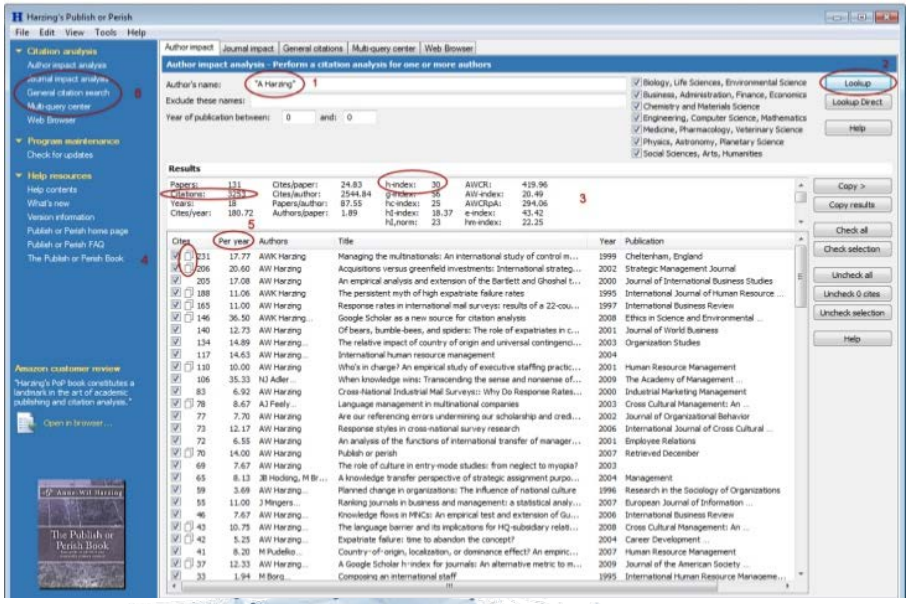
Google Scholar:

What? Google Scholar contains high citation counts, as it includes references from diverse sources. For papers not indexed in WoS or Scopus (i.e. historical classics), can look at accrued citations to date.

How?

Use Publish or Perish software to retrieve and analyze citation information. Includes h-index and related parameters.

Publish or Perish Software

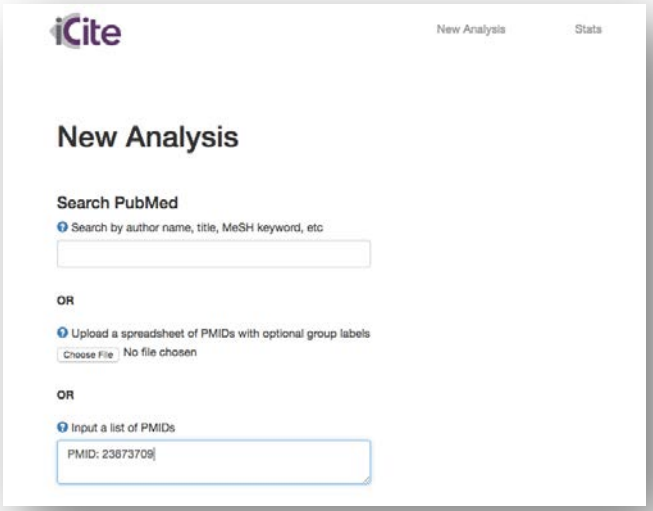


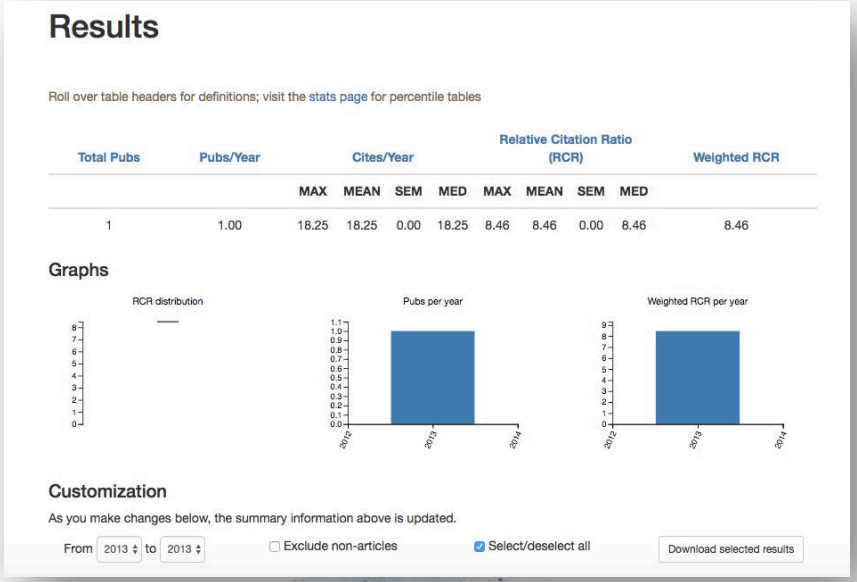
Article-Level Citations

Relative Citation Ratio:

What? RCR is the only metric specifically designed to measure the influence of individual articles. Field-independent measure that shows influence relative to the average NIH-funded paper.

How? Capture PMIDs in PubMed and run through NIH's iCite Tool.





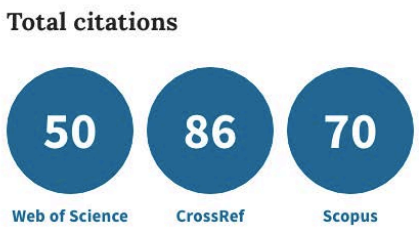
Article- or Author-Level

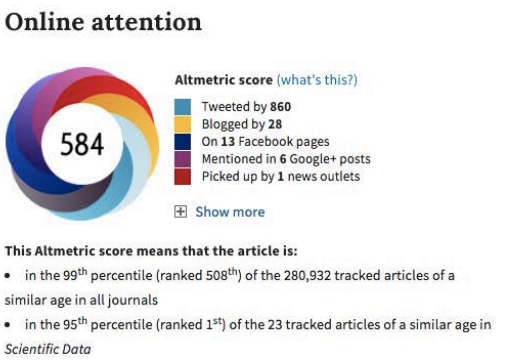
Altmetrics:

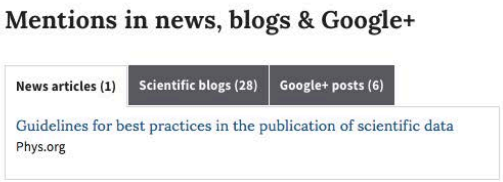
What? Altmetrics track usage stats and social media/press attention. These immediate measures complement traditional bibliometrics and are increasingly important to stakeholders.

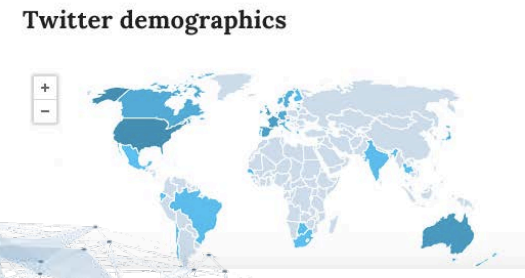
How? Use what's available (i.e. Twitter analytics, social influence metrics, etc).

Contact your publisher to see whether Altmetric or similar API in effect. If yes, request the data.

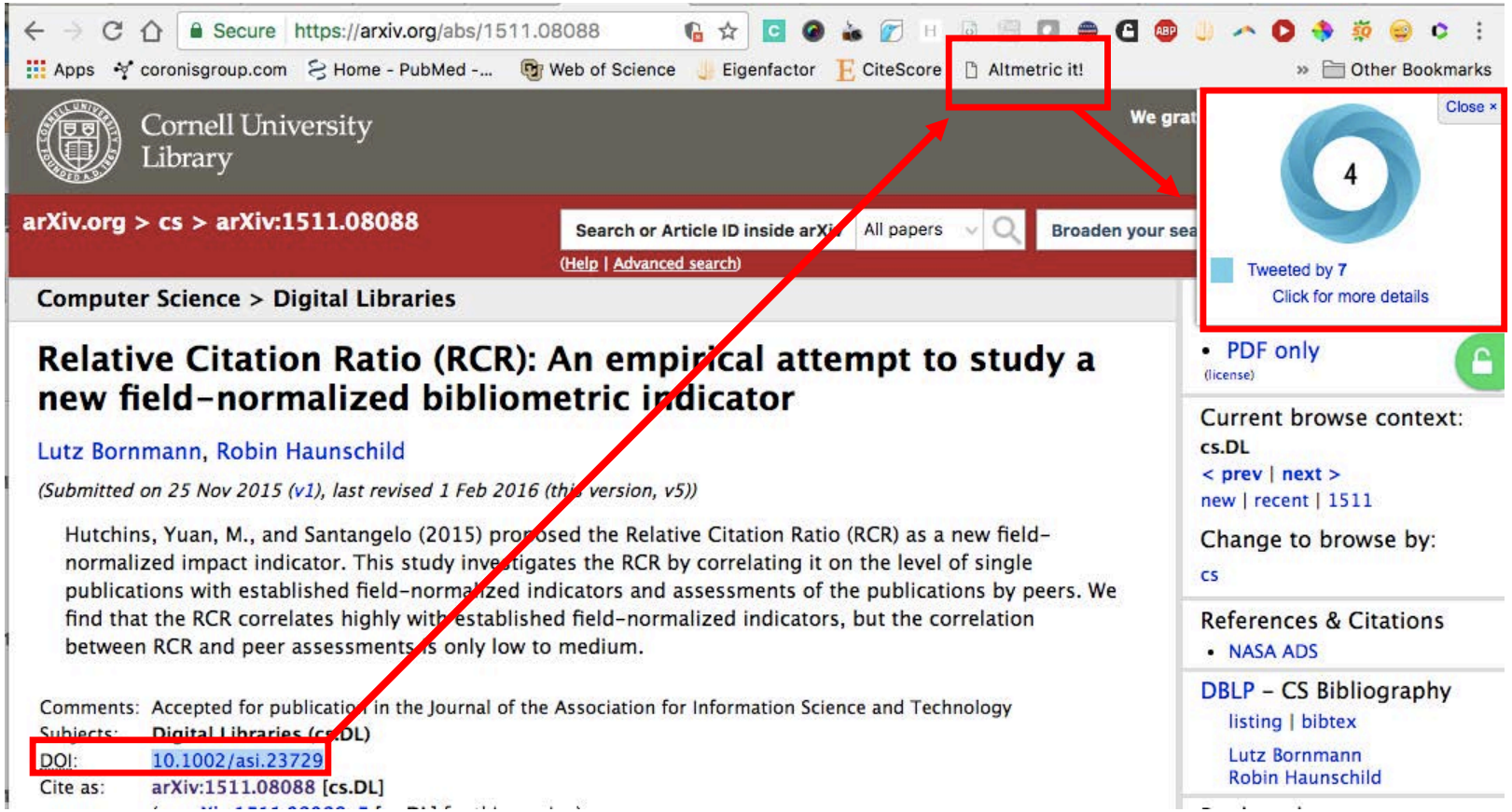








Tracking Altmetrics by Hand



Altmetric bookmarklet can be downloaded from https://www.altmetric.com/products/free-tools/bookmarklet/

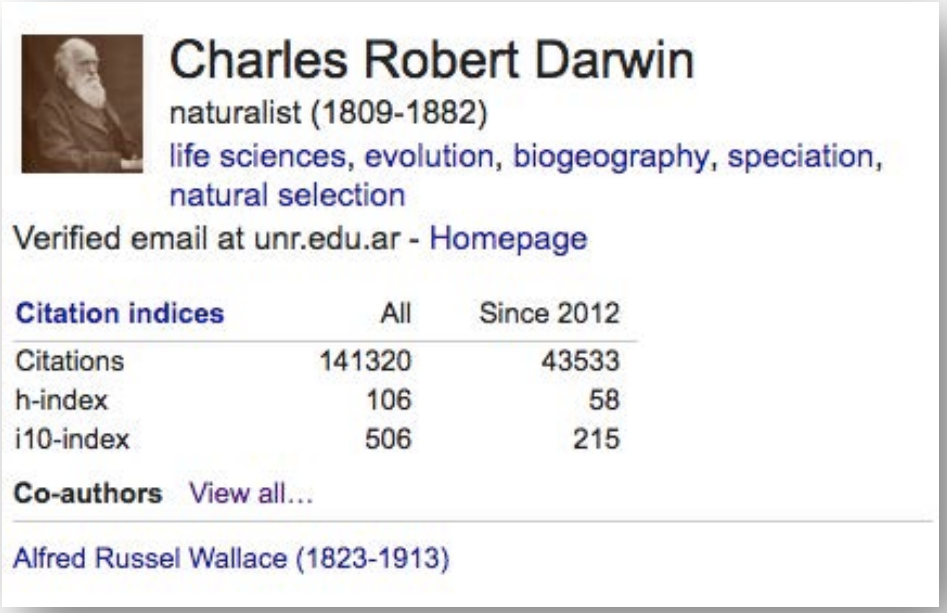
Increasing Altmetrics

- ❑ If press team available, use them! Funnel papers flagged as interesting during review. No press team? Serve as conduit for journalists directly.
- ❑ Encourage board members to Tweet, post, blog about articles articles of interest to them in each issue.
- ❑ Work with societies or institutional groups to start online journal clubs.
- ❑ Encourage board members to follow the journal's social media accounts and RT/share.
- ❑ Board members with large social media followings could give a presentation at ed board meeting with tips for other members.
- ❑ When article is accepted, acceptance letter should encourage authors promote the link through their social media accounts.
    - ❑ Provide template for visual abstracts.

Author-Level Metrics

Citation Counts

Straight counting. Identify productive or impactful authors by publication or citation counts. Web of Science generally returns lower counts than Google Scholar and Scopus.



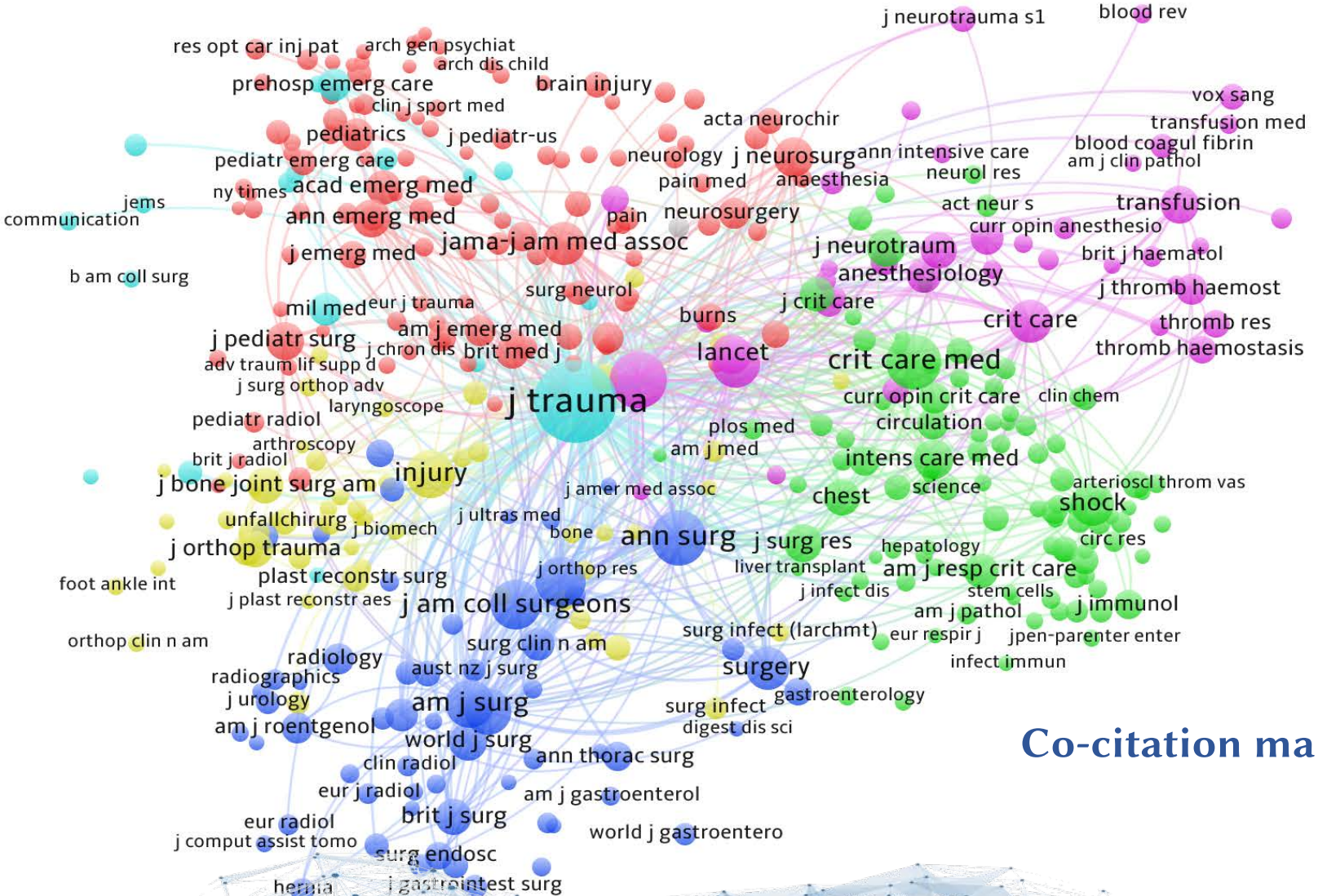
H-index

Author with an index of **h** has published **h** papers, each of which has been cited at least **h** times. Benefits longer publication histories. Normalize by dividing h-index by years active (m-parameter).

I10-index

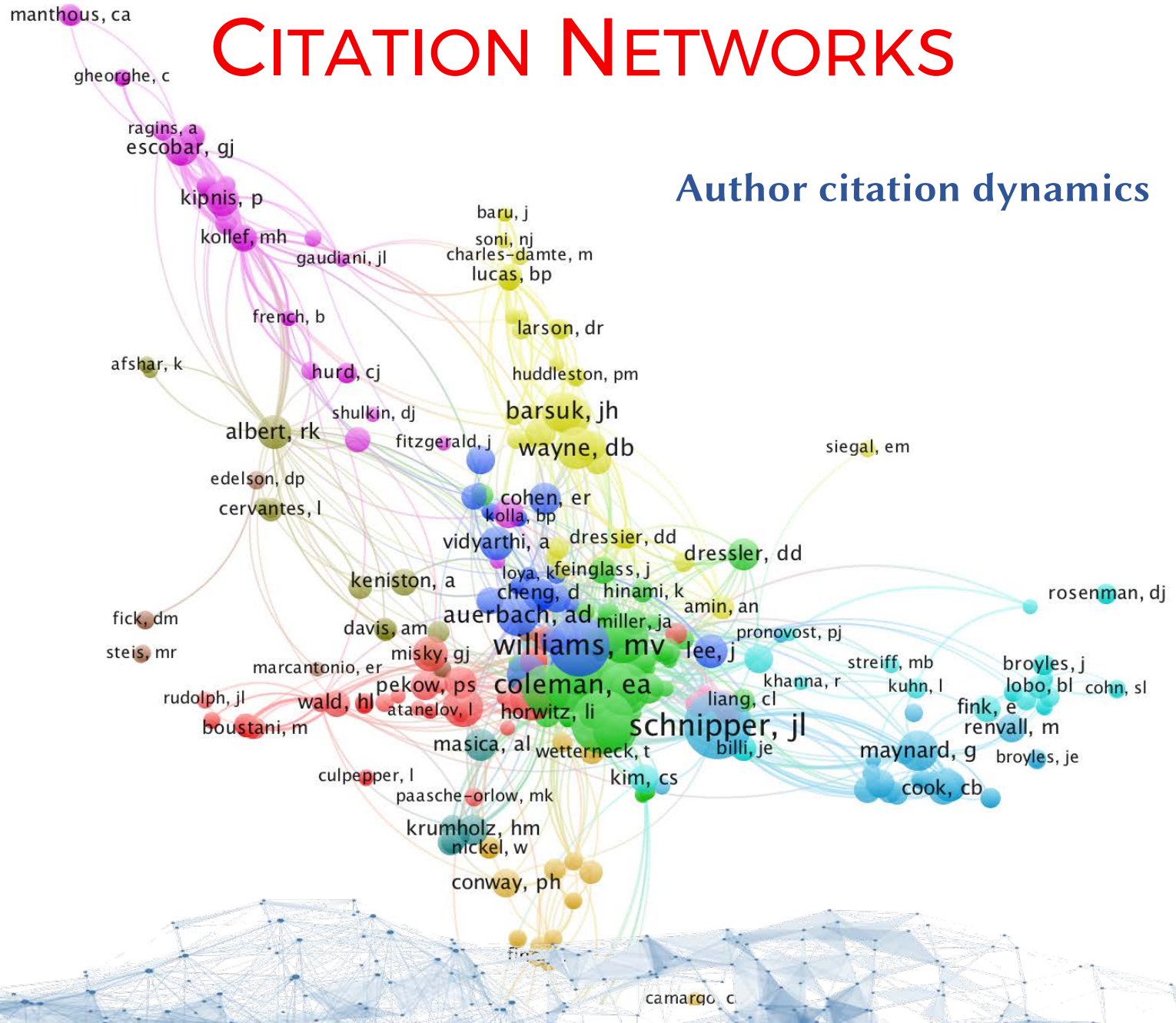
The number of publications with at least 10 citations. Used only by Google Scholar.

Citation Networks



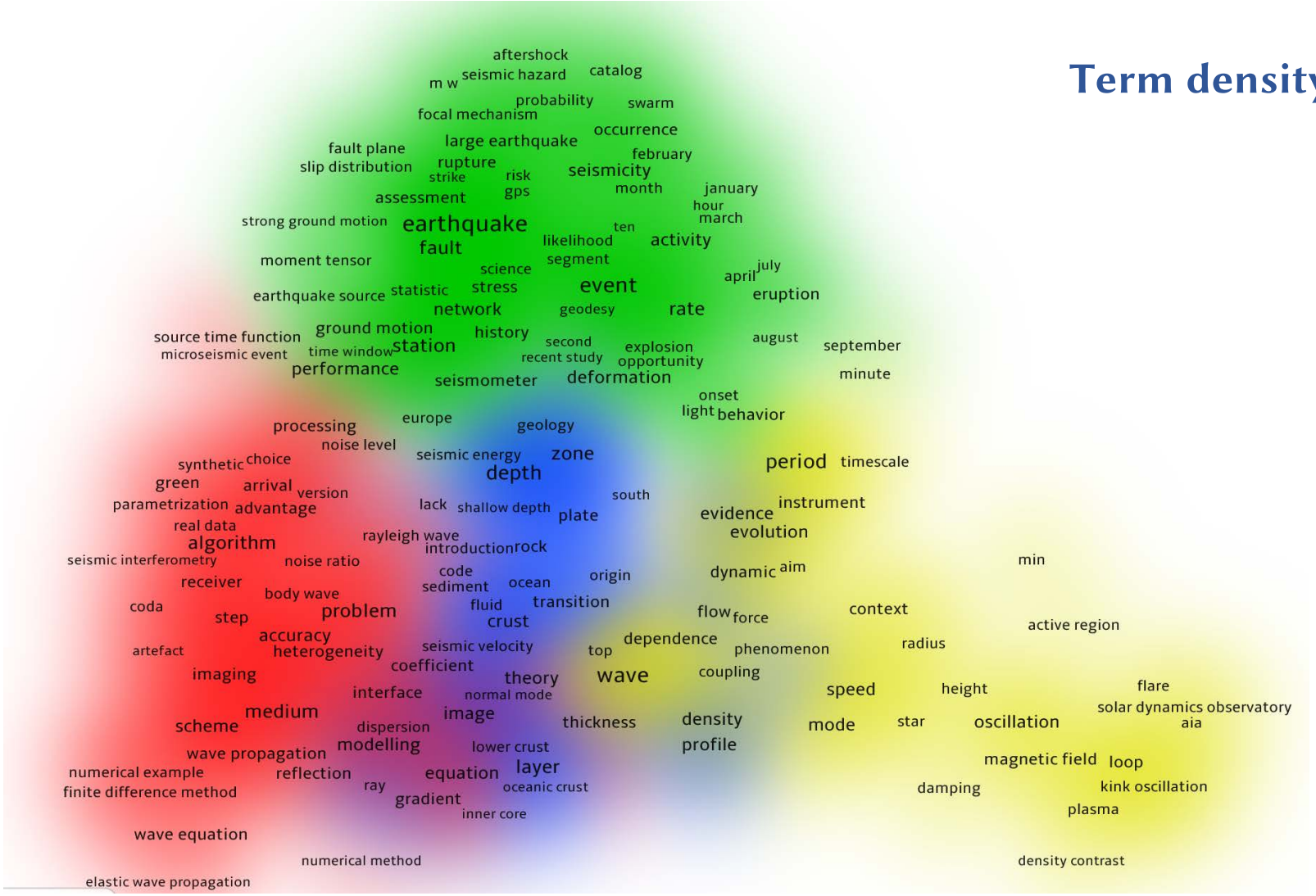
Co-citation map

Citation Networks



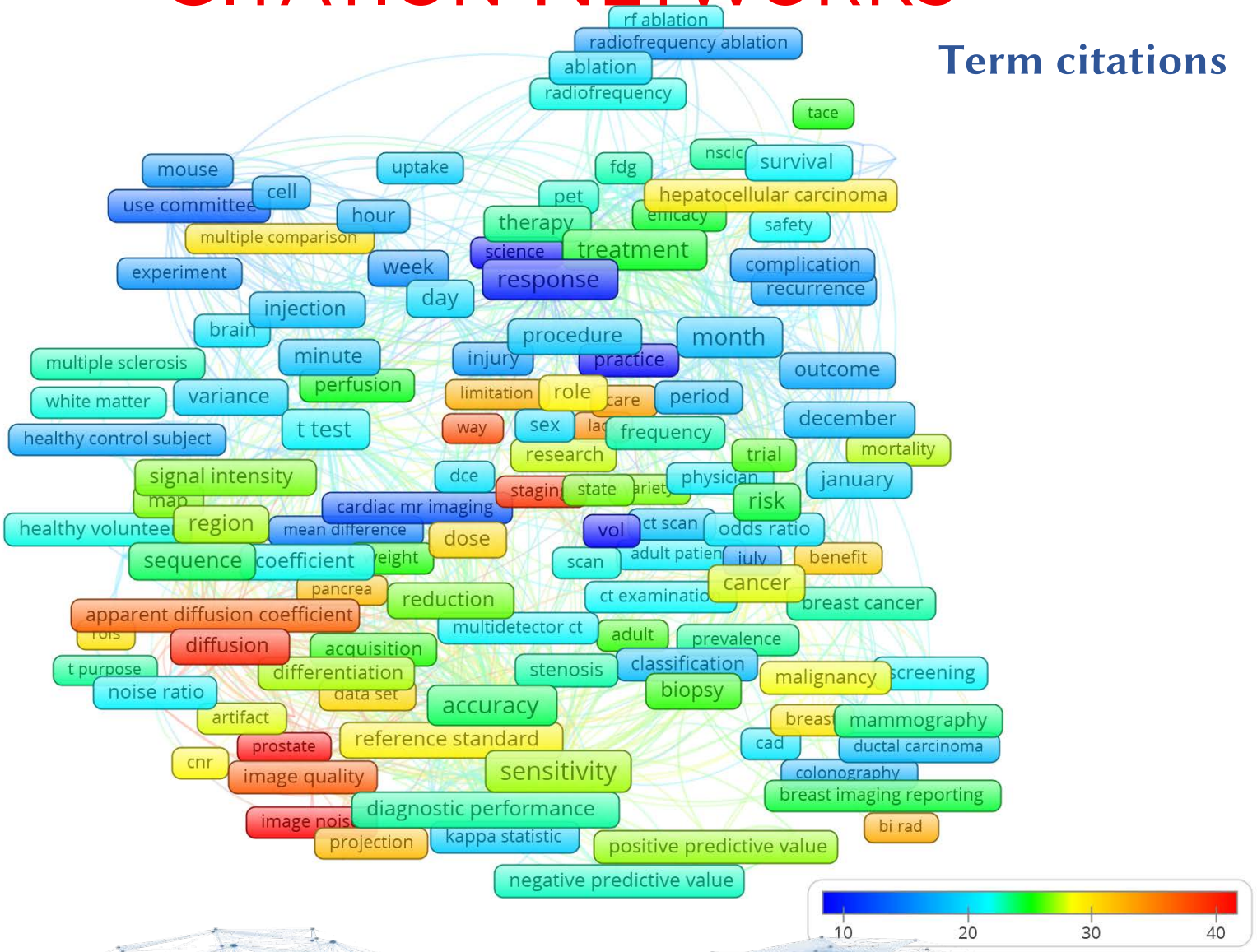
Citation Networks

Term density



Citation Networks

Term citations



Using Bibliometrics Well

Metrics must be normalized.

- Citation and publication behavior varies widely according to discipline. If a journal's scope bridges fields, its actual peers may not be who you think they are. For this, you need citation dynamics.
- Likewise, if using metrics at the author-level, be sure to normalize by career age.

Do not conflate journal-level metrics with performance of individual articles.

- Ready-made metrics are meaningless on their own – they can be skewed by a small number of articles. Go back to the source!

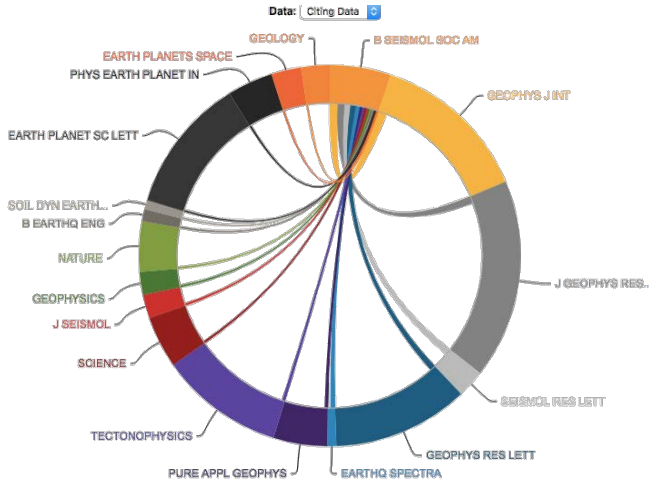
There is no such thing as a one-size-fits-all metric

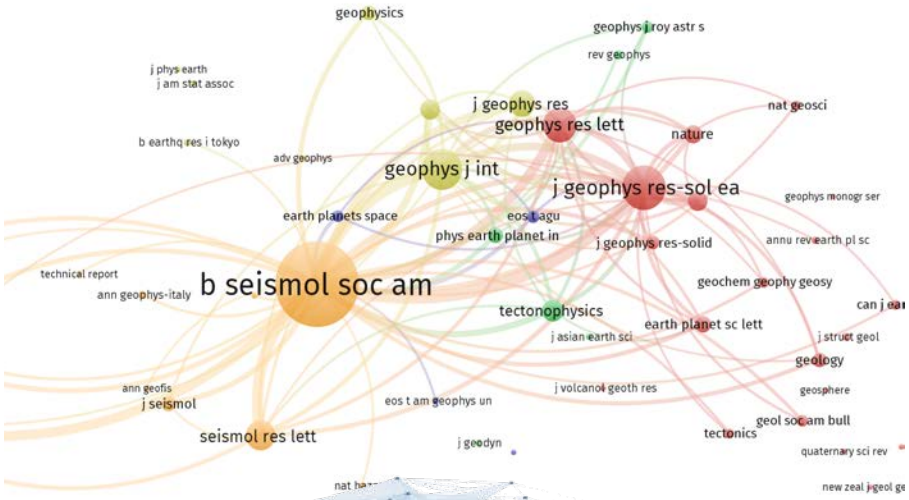
- Use a set of metrics to gauge all levels of journal health, especially when undertaking new initiatives. Track early and often.

Editorial Applications

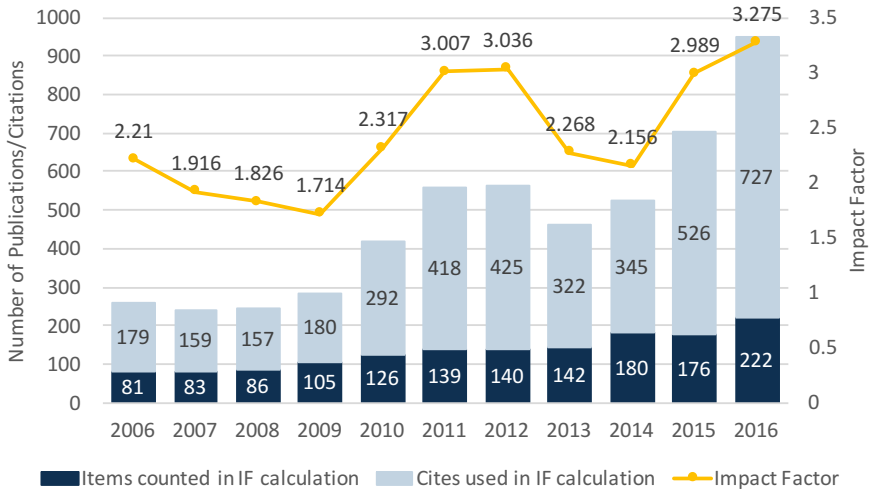
How does your journal stack up against its peers?

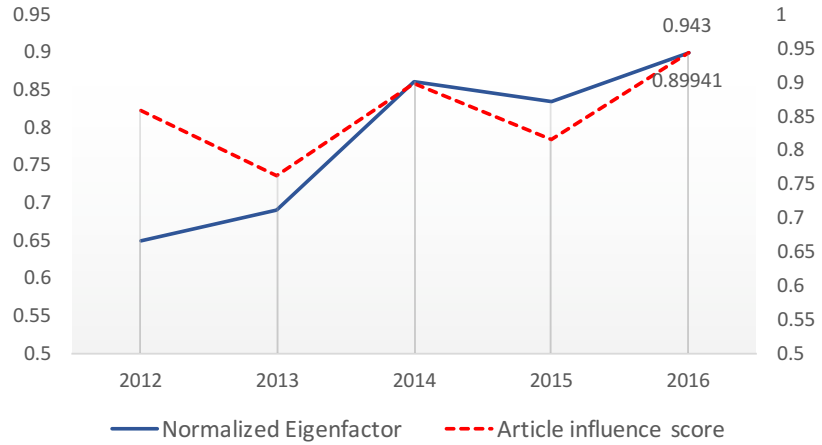
- Find its cohort. Which journals do your authors cite? Which journals cite your journal's content?
- Either look at the citing/cited journals in JCR/Scopus or consider visualizing its co-citation universe.
- Can run a topic search to see if highly cited content in your journal's scope is being published/cited elsewhere.





Editorial Applications





Why did our impact factor (or other metric) rise/fall?

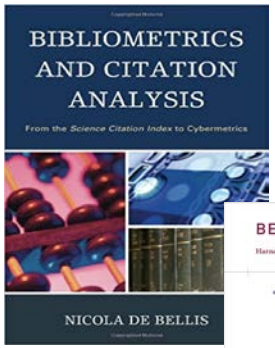
- Go back to the source – examine lists of citation data feeding into the metric.
- Impact factor will be affected by **volume** of citable content and **timing** of print publication.
- Eigenfactor and SJR will register increased citations from top-tier journals (overall or closely related)
- Think about meaning of metric changes before making any decisions regarding content or process.

Cheat Sheet

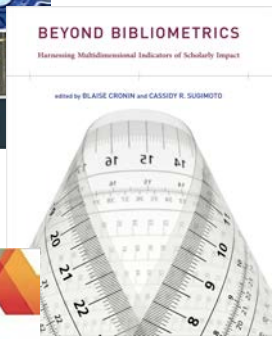
Editorial interest	What to look at
Classic papers	• Citation count • Citation peak map
Competitor title identification	• Cited by/citing data • Network co-citation map
Competitor analysis	• Article type breakdowns & citation counts • Lists of highly cited articles • Never-cited rates and h-indices
Editorial board candidates	• H-index corrected for career maturity • Altmetrics and community engagement • Centrality in journal network space
Digital efficacy	• Altmetrics, backlinks, and conversion to views • Immediacy index
Highly cited topics	• Browse through highly-cited lists (but know your peak!) • Topical citation maps

Further Reading

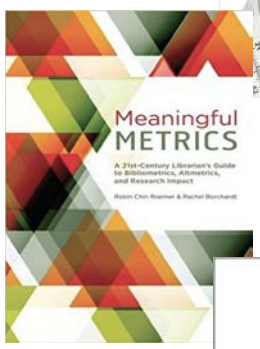
- *Bibliometrics and Citation Analysis* by Nicola De Bellis
- *Beyond Bibliometrics* by Blaise Cronin and Cassidy R. Sugimoto
- *Meaningful Metrics by* by Robin Chin Roemer and Rachel Borchardt
- *An Introduction to Bibliometrics: New Development and Trends* by Rafael Ball (upcoming release! Sept 28, 2017)



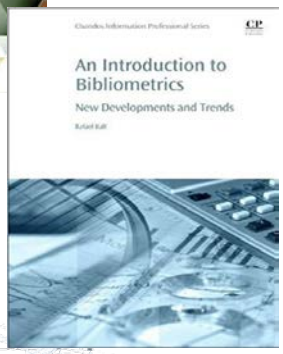
2009



2014



2015



2017

Thank You!

jmavzer@gmail.com
consult@coronisgroup.com