

Citation Indexes and Bibliometrics

Giovanni Colavizza

The long story short

Early XXth century: quantitative library collection
management

1945: Vannevar Bush in the essay “As we may think”
proposes the memex
(a collective memory machine to deal with information
explosion)

1964: the Science Citation Index

1960s-70s: first studies in the sociology of science

1980s: proposals for using bibliometric data for evaluation



Eugene Garfield

Chemist turned linguist, information retrieval perspective.

Main goals: algorithmic historiography and tracking impact, building the memex (creator of the Science Citation Index).



Derek de Solla Price

Physicist turned historian of science,
modelling perspective.

Main goals: find laws which can predict the
accumulation of scientific knowledge and
citation cycles/patterns.



Robert Merton

Sociologist, founding father of the sociology of science, sociological perspective.

Main goals: science as a social system, e.g. cumulative advantage (Matthew effect).

The three “missions” of modern bibliometrics:

- Access (information retrieval)
- Understanding
- Evaluation

Bibliometrics: the statistical analysis of written publications

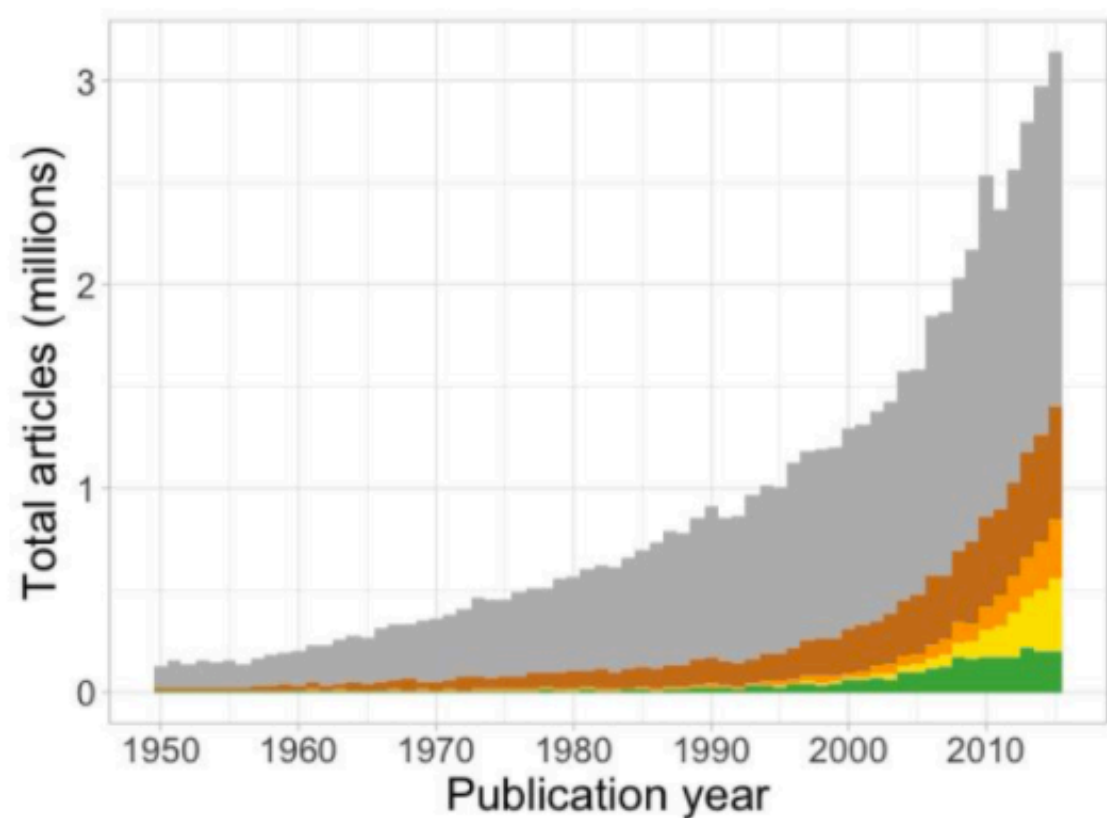
Scientometrics: the measurement and analysis of science

Altmetrics: alternative (web) metrics (within scientometrics)

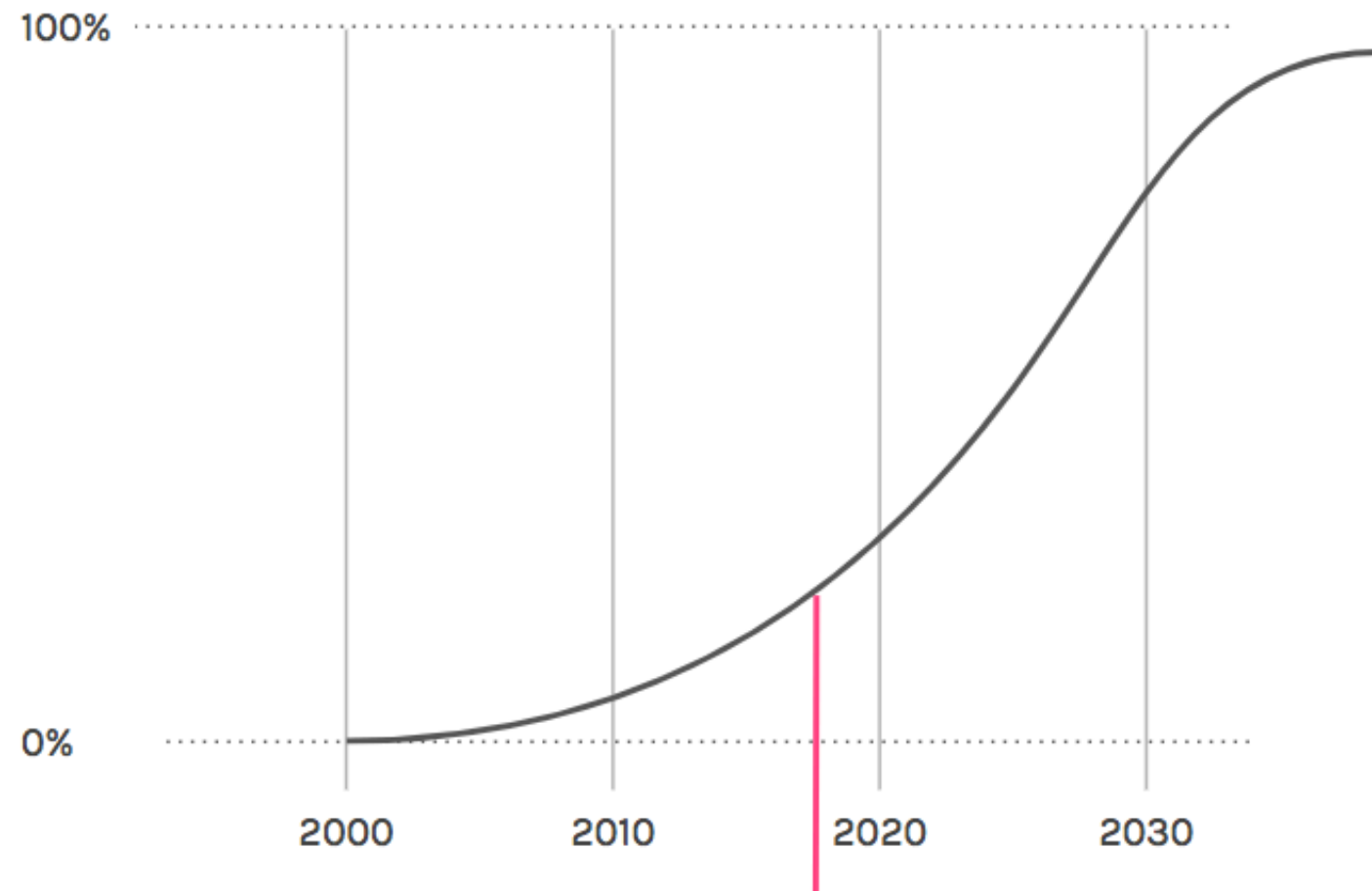
Informetrics: information measurement, more broadly

Trendy topics

Open Access



Source: Piwowar et al. "The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles", preprint 2017.



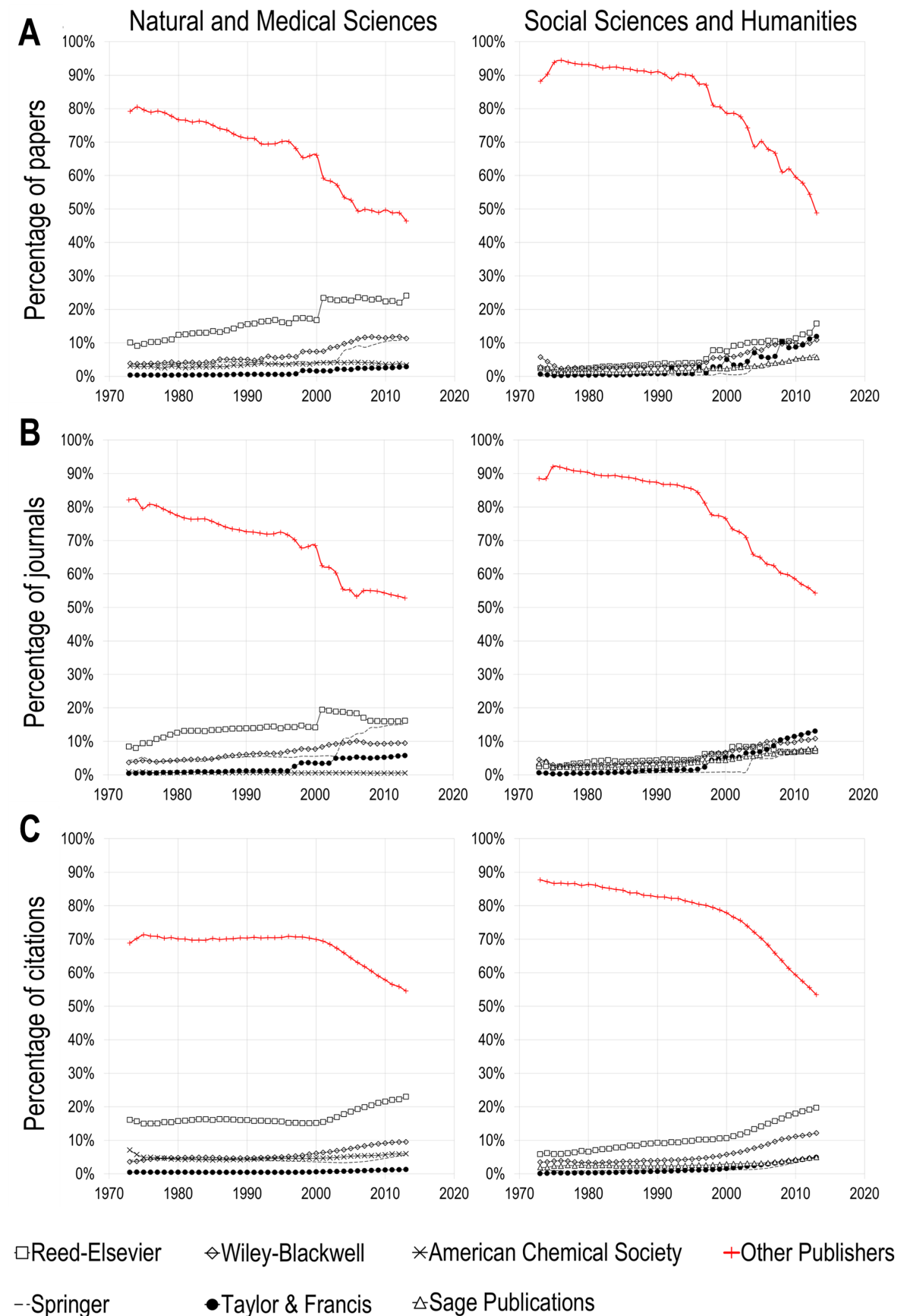
NOW
All scientific papers to be free by 2020
under EU proposals

Source: The Guardian 2016

Part of broader open science and data movement

The oligopoly of science publishing

Larivière, Vincent, Stefanie Haustein, and Philippe Mongeon. 2015. "The Oligopoly of Academic Publishers in the Digital Era." Edited by Wolfgang Glanzel. *PLOS ONE* 10 (6): e0127502.
doi:10.1371/journal.pone.0127502.





the first website in
the world to provide
mass & public access
to research papers



SCI-HUB

...to remove all barriers in the way of science

enter URL, PMID / DOI or search string



open

Oversimplification

- Publications = productivity
- Citations = impact

Uninformed use and misuse

- Impact factor
- h-index

Adverse effects

- “Salami” publishing
- Honorary authorship
- Self-citations
- Citation cartels



Altmetric

Citation Indexes

From publications to citation indexes

Citing
publication



Cited
source

Dopo la ritirata di Pipino l'imperatore di Costantinopoli inviò ad Aquisgrana, presso Carlo Magno, il legato Arsafio, il quale nell'812 stipulò un trattato col sovrano franco. In esso fu stabilito il riconoscimento a Carlo della dignità imperiale (ma non di imperatore dei "romani") da parte del *basileus*, in cam-

3. GIOVANNI DIACONO, *Istoria Veneticorum*, a cura di L.A. Berto in *Cronache*, a cura di L.A. Berto e G. Fedalto, in *Corpus Scriptorum Ecclesiae Aquileiensis*, XII/2, Roma 2003, pp. 72-74.

4. Ivi, p. 74. G. OSTROGORSKY, *Storia dell'Impero bizantino*, Torino 1968, p. 177; A. CARILE, "La presenza bizantina dell'alto Adriatico tra il VII e il IX secolo" in *AAA*, XXVII (1985), pp. 107-129. Negli *Annales*, Eginardo attribuisce, invece, la vittoria ai franchi.

REFERENCES

- Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004a: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494.
- , and Coauthors, 2004b: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.
- Bernardet, L. R., and W. R. Cotton, 1998: Multiscale evolution of a derecho-producing mesoscale convective system. *Mon. Wea. Rev.*, **126**, 2991–3015.
- Bjerknes, J., and H. Solberg, 1922: Life cycle of cyclones and the polar front theory of atmospheric circulation. *Geofys. Publ.*, **3** (1), 3–18.

Big players



☒ Articles (☒ include patents) ☐ Case law

Stand on the shoulders of giants



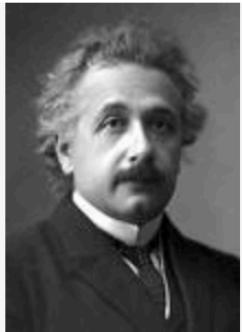
WEB OF SCIENCE™

Scopus®

Free access, higher coverage, lower quality, barebone interface, most used

High paywall, better quality data, lower coverage

Cf. How researchers keep up with the literature:
<http://www.sciencemag.org/careers/2016/11/how-keep-scientific-literature>



Albert Einstein

Institute of Advanced Studies, Princeton

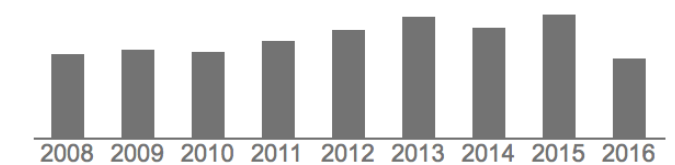
Physics

No verified email

 Follow

Google Scholar

Citation indices	All	Since 2011
Citations	97866	32160
h-index	109	65
i10-index	368	221



Title	1-20	Cited by	Year
Can quantum-mechanical description of physical reality be considered complete?	A Einstein, B Podolsky, N Rosen Physical review 47 (10), 777	14648	1935
Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt	A Einstein Ann. Phys. 17, 132-148	7927 *	1905
On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat	A Einstein Annalen der Physik 17, 549-560	6089 *	1905
Zur Elektrodynamik bewegter Körper	A Einstein	4590 *	
Investigations on the Theory of the Brownian Movement	A Einstein Dover publications	3889	1956
Eine neue bestimmung der moleküldimensionen	A Einstein Annalen der Physik 324 (2), 289-306	3471	1906

Domain specific indexes: F1000

F1000Prime F1000Research F1000Workspace Welcome Guest

F1000Prime Top experts select and rate the best articles for you


ARTICLE RECOMMENDATIONS | RANKINGS | F1000PRIME REPORTS | F1000 FACULTY | BLOG MyF1000 | SIGN IN | SUBSCRIBE

Select ☒ the **Sections** that most closely match your field(s) of interest and click the **NEXT** button on the right

dna immune bacteria


If the suggested **Sections** below do not reflect your current research interests, edit the keywords as appropriate


1 2 3 **NEXT**


☐ **Innate Immunity** **SHOW TOPICS**

Section Heads
Alan Ezekowitz | Jules Hoffmann | Dieter Kabelitz | David Raulet


☐ **Immunity to Infections** **SHOW TOPICS**

- > Acquired responses
- > Antigen processing/presentation
- > Genetic susceptibility factors in hosts
- > Immune evasion by microbes
- > Inflammation and cytokines
- > Innate recognition of microbes
- > Microbiome effects on immunity
- > Pathogenesis and resistance
- > Specific, major infectious diseases


Section Heads
Jason Cyster | Susan Swain

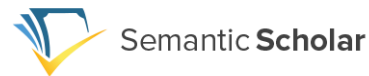
☐ **Cellular Microbiology & Pathogenesis** **SHOW TOPICS**

Section Heads
William Goldman | Ralph Isberg

☐ **Microbial Physiology & Metabolism** **SHOW TOPICS**

Section Heads
James Broach | Christophe d'Enfert | Simon Foster | Carol Gross

☐ **Medical Microbiology** **SHOW TOPICS**

Section Head
Jorg Hacker

Mainly on PubMed, manually curated

Domain specific indexes: Semantic Scholar



Deep Learning

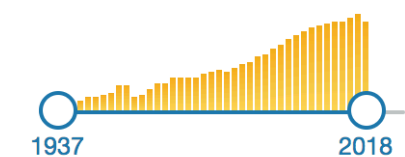
SIGN IN



Filter Results:

☐ Full text PDF available (78,494)

Publication Year



This year (10,013)
Last 5 years (53,306)
Last 10 years (76,338)

Publication Type

Author

Publication Venue

Brain Region

Cell Type

About 98,800 results

Sort by: Relevance

Vision-Based Assessment of Parkinsonism and Levodopa-Induced Dyskinesia with Deep Learning Pose Estimation

Trending

Michael H. Li, Tiago A. Mestre, Susan H. Fox, Babak Taati • ArXiv • 2017

Objective: To apply **deep learning** pose estimation algorithms for vision-based assessment of parkinsonism and levodopa-induced dyskinesia (LID). Methods: Nine participants with Parkinson's disease (PD) and LID completed a levodopa infusion protocol, where symptoms were assessed at regular intervals using the Unified Dyskinesia Rating Scale (UDysRS) and... [\(More\)](#)

Mentioned in 4 tweets



Michael Li
@MichaelHGLi

Follow

Our paper on computer vision-based assessment of [#Parkinson's](#) is now available on [@arxiv](#):
arxiv.org/abs/1707.09416 [@Babak_Taati](#) [#deeplearning](#)
pic.twitter.com/gNGLmWDbel

2:46 AM - Aug 1, 2017

3 3 6

[View on ArXiv](#) [Related Publications](#) [More...](#)

Mainly on arXiv and for computer science,
some AI components

Open access to 1,301,360 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics

Subject search and browse: 

08 Aug 2017: [A survey for users accessing arXiv programmatically](#)

20 Apr 2017: [Applied Physics](#) subject area added to arXiv

10 Mar 2017: [New members join arXiv Member Advisory Board](#)

06 Mar 2017: [arXiv Scientific Director Search](#)

10 Feb 2017: [Attention Submitters: our TeX processing system has been updated](#)

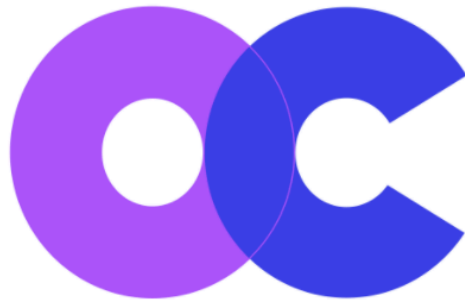
See cumulative "What's New" pages. Read [robots beware](#) before attempting any automated download

Open indexes

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.

.....



Welcome to the [OpenCitations](#) homepage!

{ } wikicite

Movement towards
free access

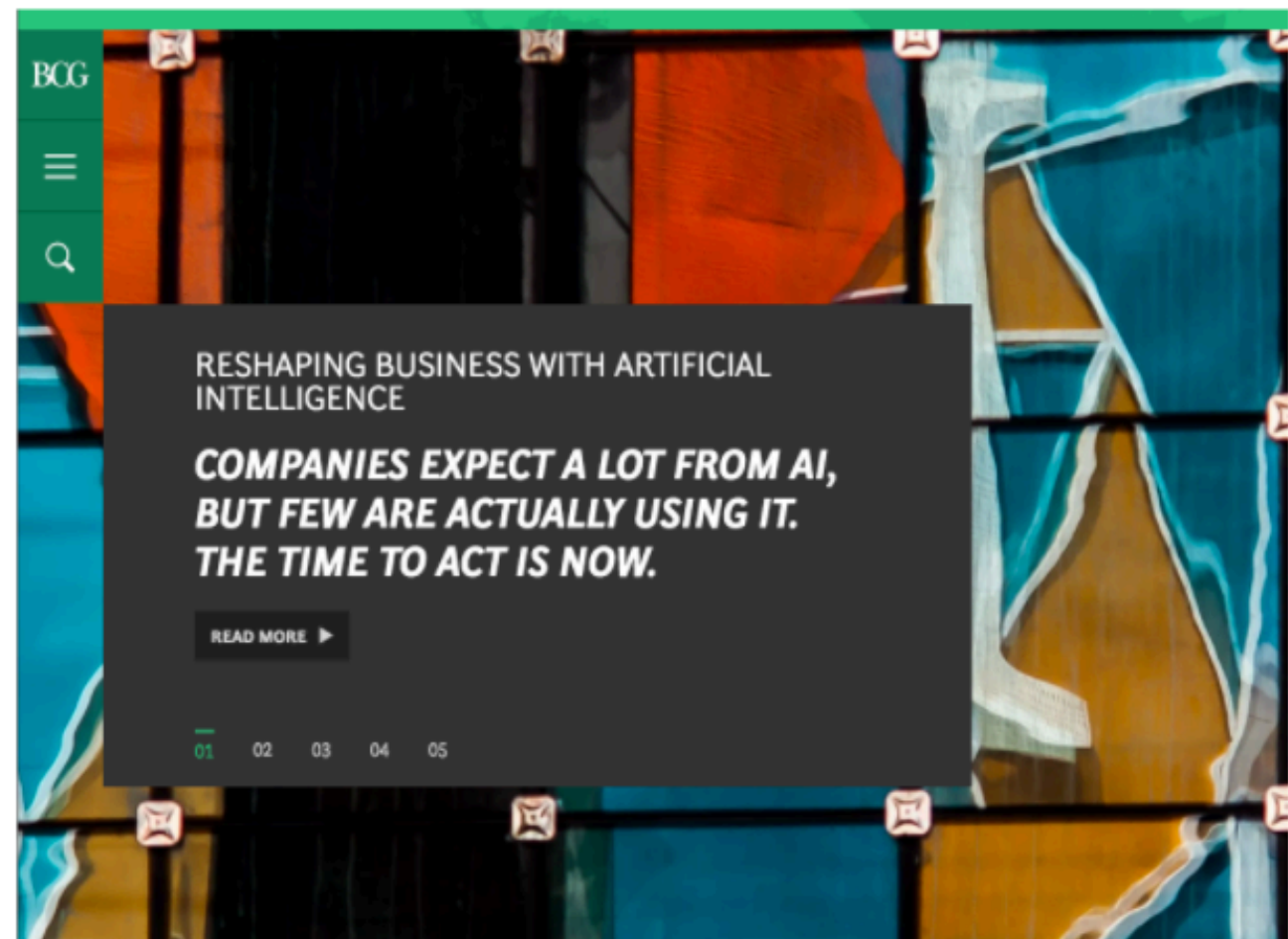
Different
alternatives for
large-scale open
repositories

Microsoft, Google and Baidu team up with Paul Allen's AI2 on Open Academic Search

Source: www.geekwire.com, June 2017



Science AAAS, 7th July 2017



Home Page Boston Consulting Group, September 2017

The humanities and social sciences

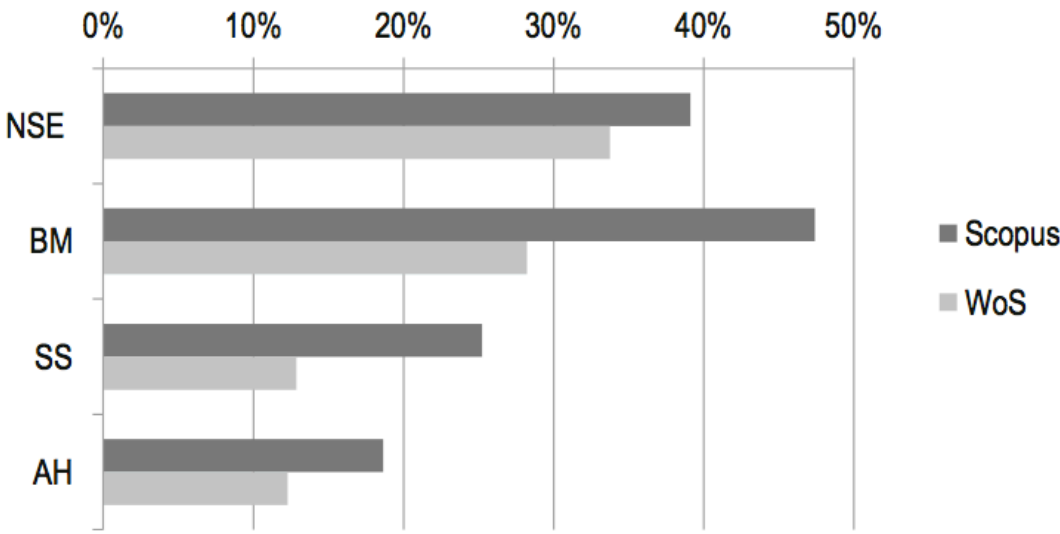


Fig. 1 Proportion of Ulrich academic journals indexed in Web of Science and Scopus

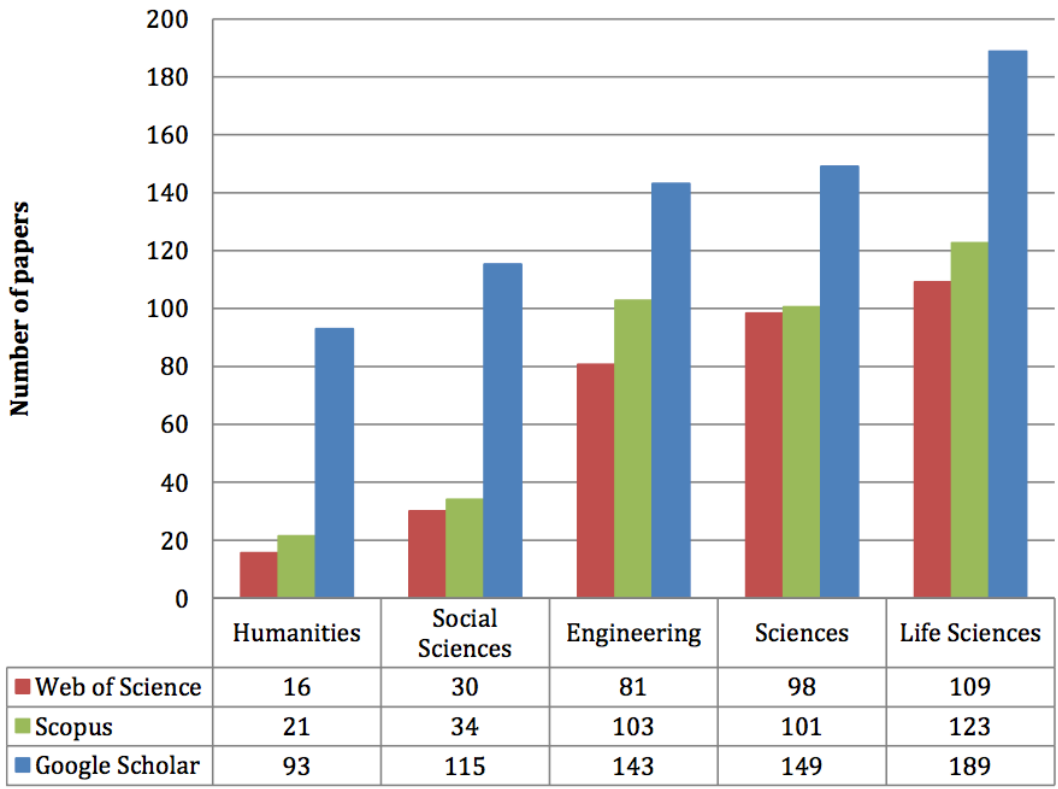


Fig. 3 Average number of papers per academic across five disciplines and three databases, July 2015

Harzing, Anne-Wil, and Satu Alakangas. 2016. "Google Scholar, Scopus and the Web of Science: A Longitudinal and Cross-Disciplinary Comparison." *Scientometrics* 106 (2): 787–804. doi:10.1007/s11192-015-1798-9.

Mongeon, Philippe, and Adèle Paul-Hus. 2016. "The Journal Coverage of Web of Science and Scopus: A Comparative Analysis." *Scientometrics* 106 (1): 213–28. doi:10.1007/s11192-015-1765-5.

More details:

1. WikiCite
2. Open Citations
3. Linked Books
(indexation in the
humanities)

<https://meta.wikimedia.org/wiki/WikiCite>

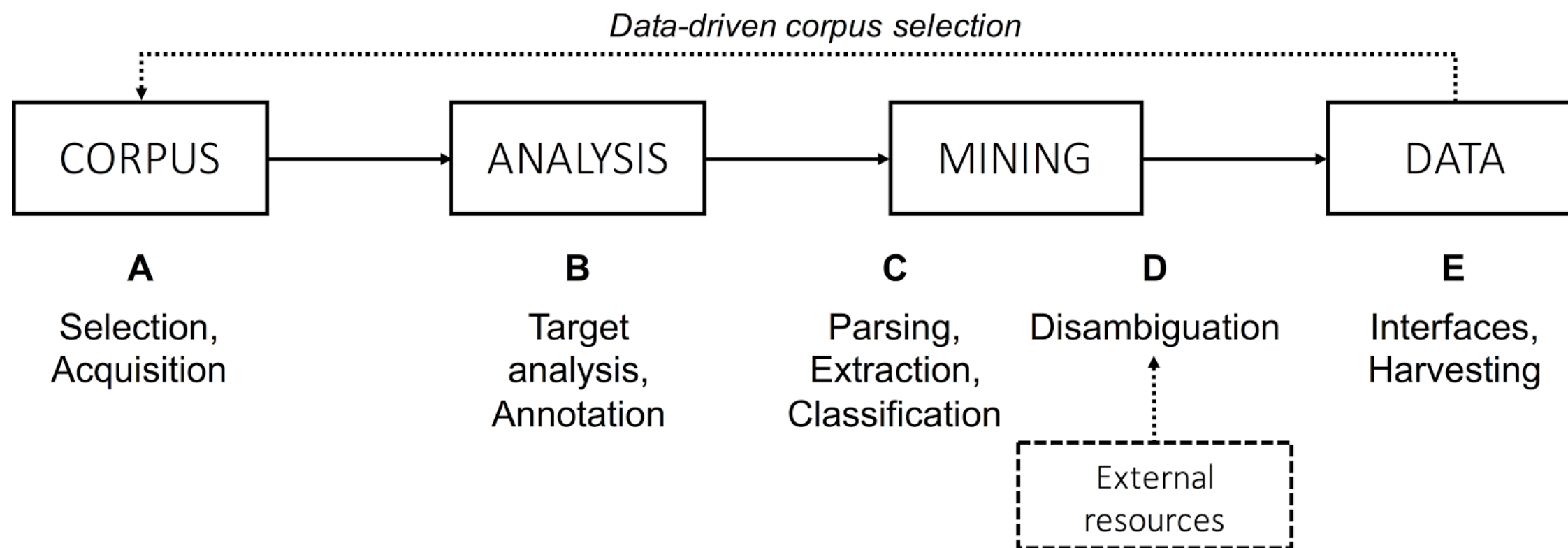
Open Citations

<http://opencitations.net/>

Swiss FNS project 2015-18

Index the literature on the history of Venice

Now ~2500 books, ~500 journal issues



Take home:

- Necessity for researches
- Many players, many unknowns
- Some clear trends with the digital turn: open access, preprints, open citations
- Openness and inclusiveness still long-term goals

Citation Indexes and Bibliometrics

Thank you!

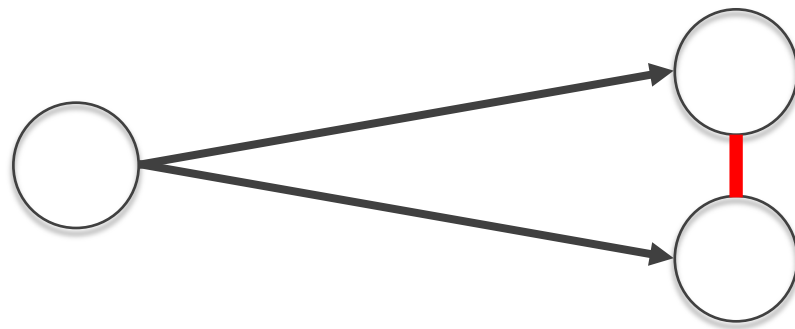
Giovanni Colavizza

Citation networks

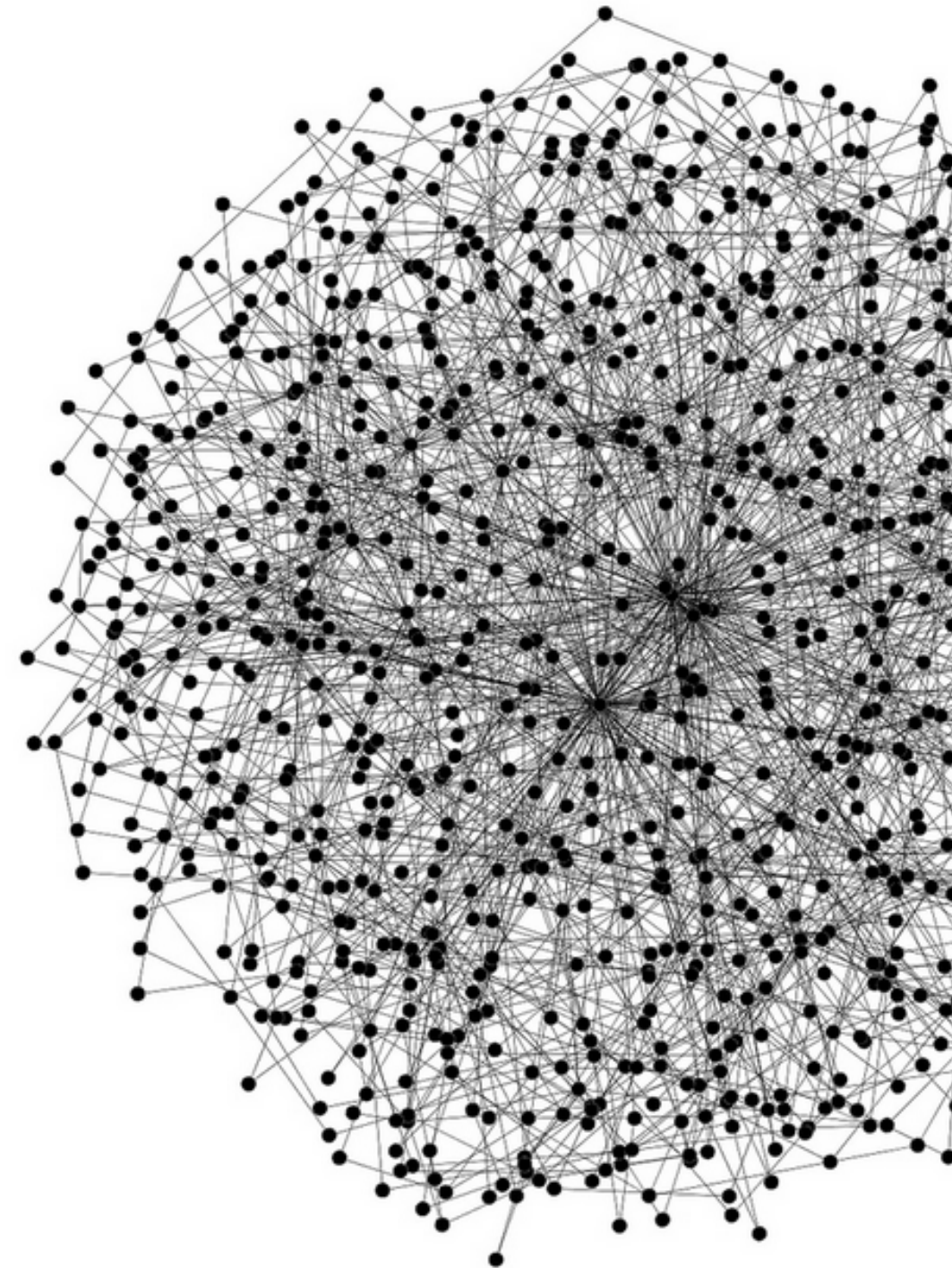
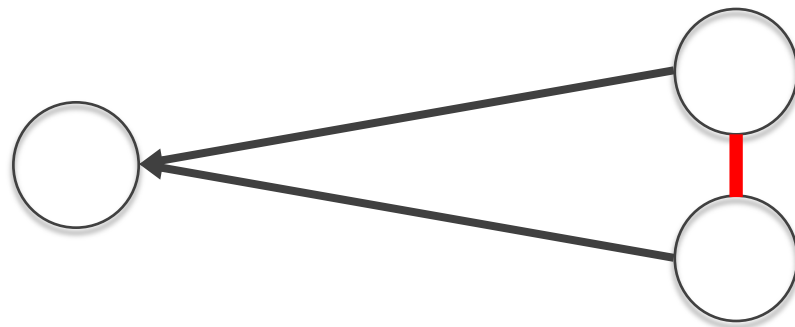
Directed network



Co-citation network



Bibliographic coupling network



Citation network analysis

Local structure:

1. Who is the most/least cited

degree centrality

1. Communities

modularity maximization (Louvain)

Global structure:

1. Who is a global broker

betweenness centrality

1. How dense is the network?

network density and diameter

Dynamic structure:

How does all change over time?

A bibliometric analysis with Gephi

https://github.com/Giovanni1085/core_literature_historians_venice

<https://gephi.org>

https://raw.githubusercontent.com/Giovanni1085/core_literature_historians_venice/master/dataset/graphs/bibc_1.graphml