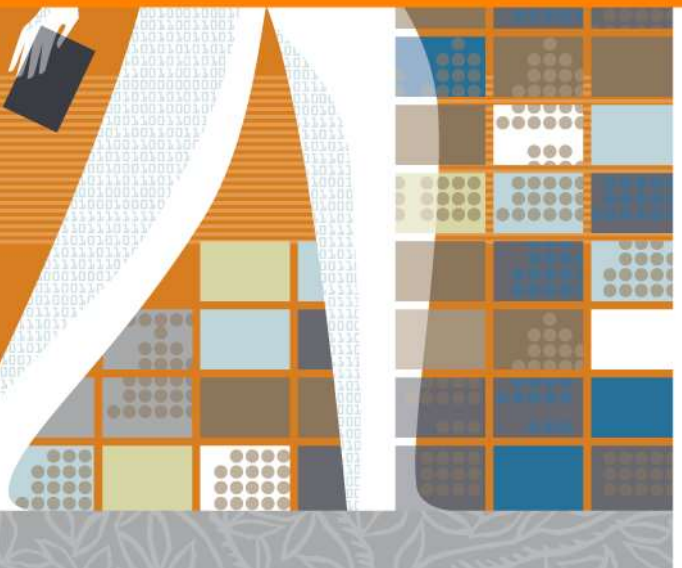
ELSEVIER

Elsevier Publishing Campus | Publishing Connect

Bibliometrics

big data and the world of scientific publishing



Prof. dr. Henrik Rudolph

Editor-in-Chief Applied Surface Science

Workshop Brazil Oct. 2018

Defining bibliometrics (in scientific publishing)

Bibliometrics [ˌbɪblɪə(ʊ)ˈmɛtrɪks]

The application of statistical analysis of books, articles, or other publications.

This means “reducing” the published scientific work into key figures, that (hopefully) say something about the quality and/or quantity of the work

Warning: By projecting a multi dimensional space onto a single parameter you loose (a lot of) information

Generating information (big data)

We generate information about ourselves constantly !!

- Information has a monetary value. Most social networks are “free”



- You pay with information about yourself and your desires
- You even allow the social networks to sell information about you
- Search engines thrive on your behaviour (page ranking)
- Your previous searches might even determine, what you pay for your next online purchase (e.g. airline ticket)

Google Bing Expedia

Publishing Connect YouTube YouTube

The “old school” social networks: scientific publishing

1439
Gutenberg and Moveable Type

Henry Oldenburg (1618- 1677)
Founding editor and commercial publisher of the first scientific journal

1580
Founding of the House of Elzevir

6th March 1665
“Philosophical Transactions of the Royal Society”

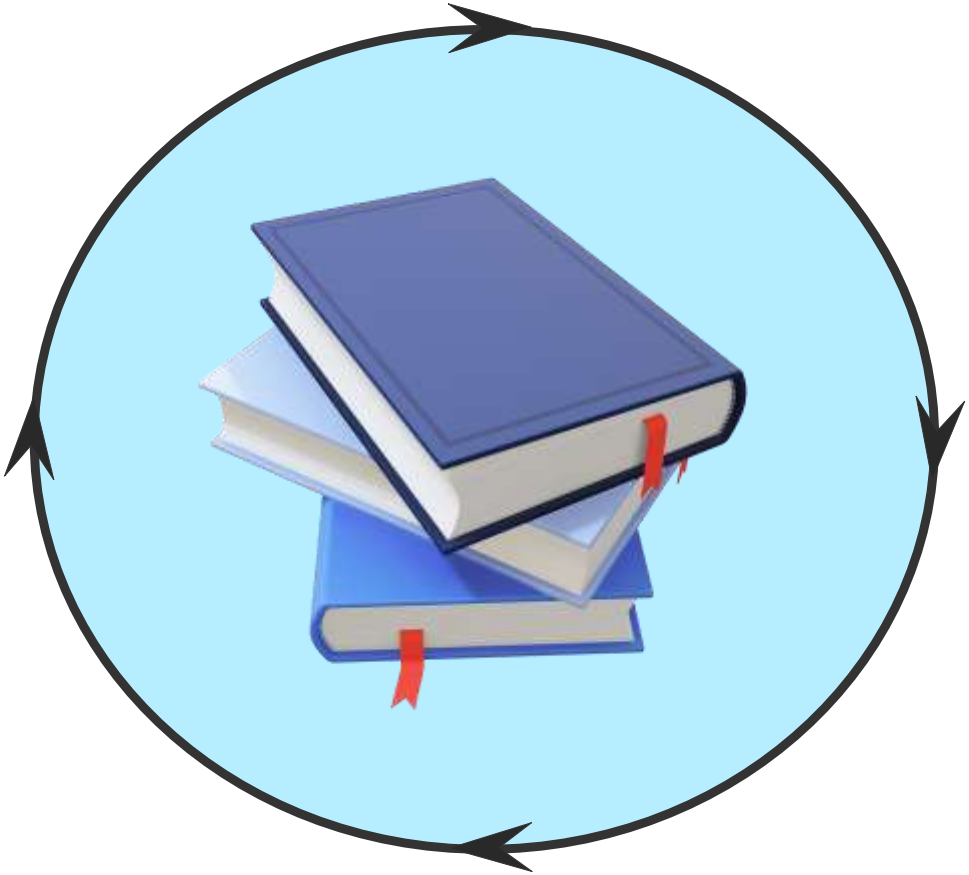
- **First true scholarly journal**

Publishing Connect

Stakeholders in the publishing circle

They all have a responsibility in the process

1) Funding “agency”
2) Academic institution
3) Author and co-authors
4) Publisher/journal
5) Editor
6) Reviewers
7) Libraries
8) Readers



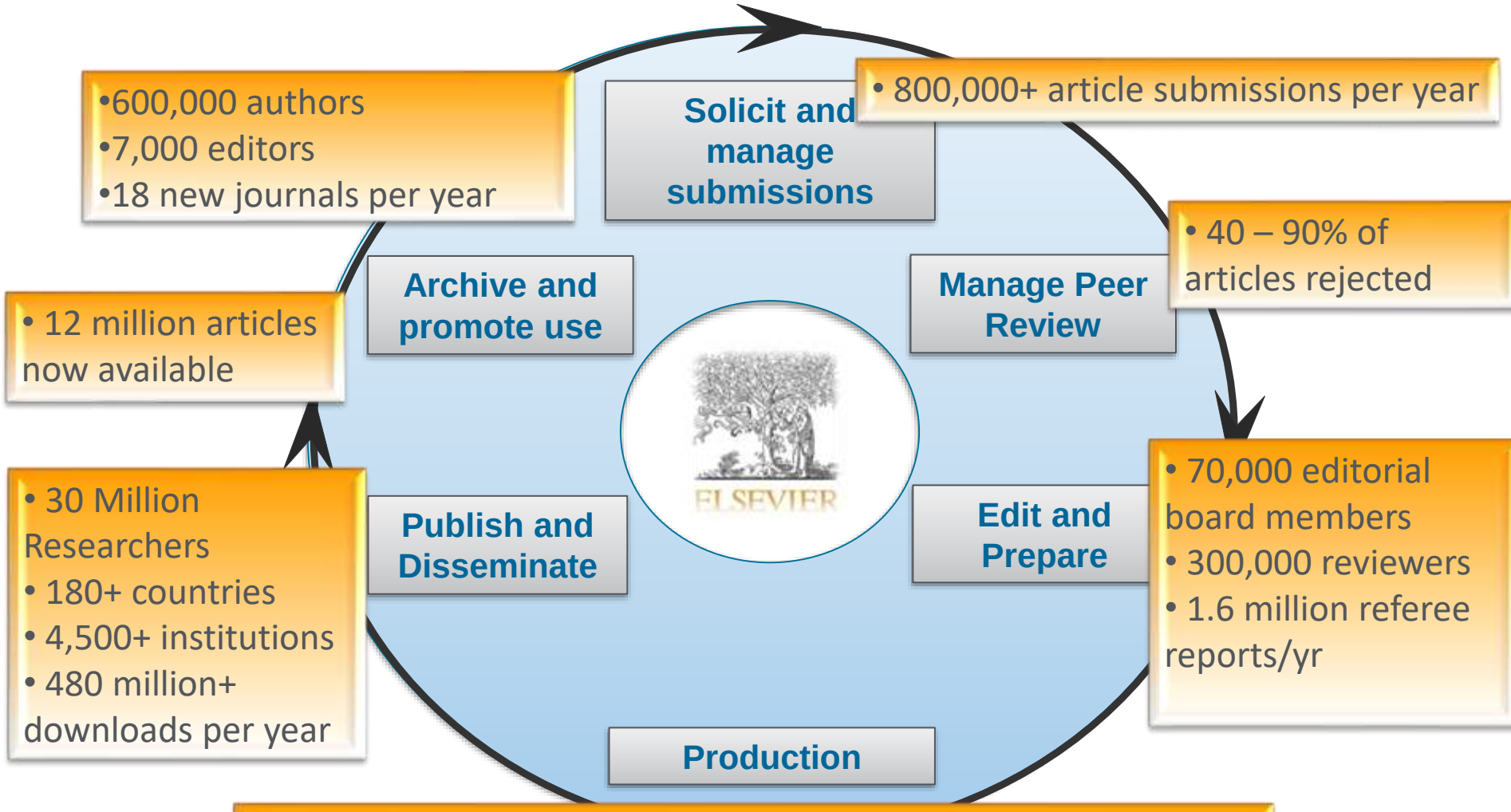
Scientific publishing

What data is being used and what data could be used: bibliometrics

All ***stakeholders*** in the publishing process make decisions based on data

- Authors decide what journal to publish in
- Readers decide what authors or journals to keep track of
- Reviewers decide what to review for
- Libraries decide what journals to subscribe to
- Funding agencies decide what journals are important
- Universities often decide about the career of a scientist based on numbers

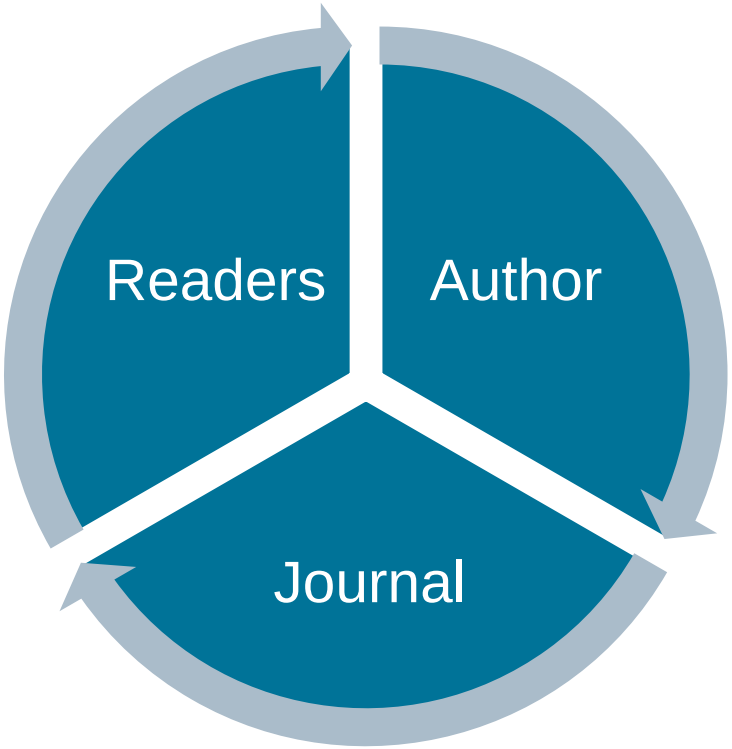
Publishing in the 21st century – the publishing circle



The three central actors of (scientific) publishing

Putting numbers on these actors

Average time spent
downloads
Areas of interest
Personal subscription

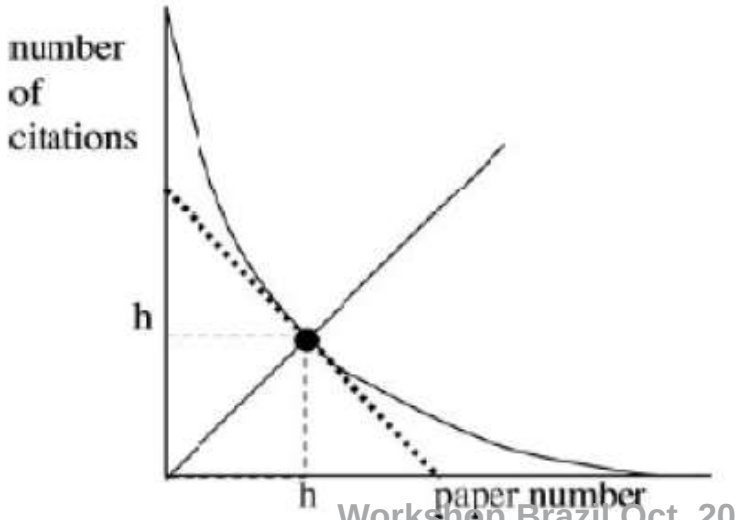


published papers
of citations
h-index
high impact papers
papers in high impact journals

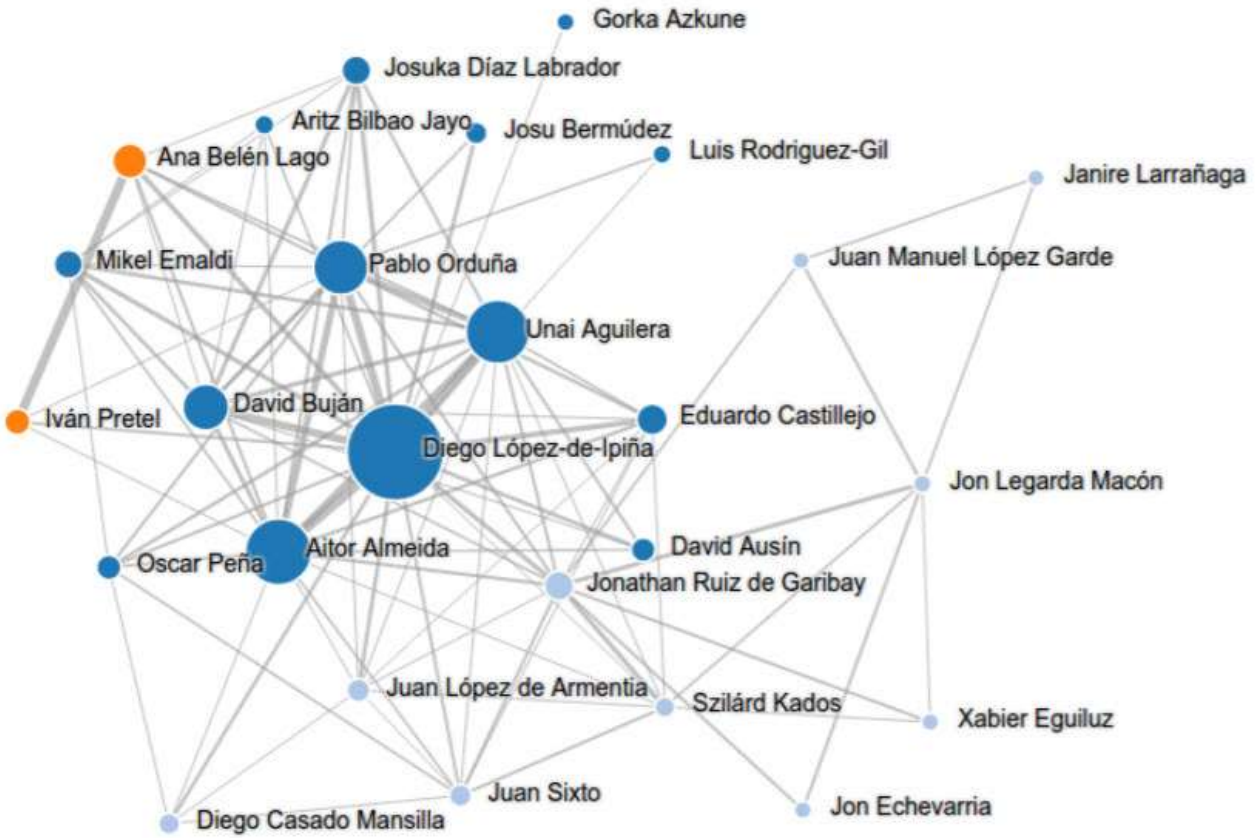
published papers (size)
Acceptance rate
Impact factor
Eigenfactor

h-index

- **Proposed by physicist Jorge Hirsch in 2005**
- **Rates individual scientist based on career publications**
- **Incorporates both quantity (no. publications) and quality (no. citations)**
- **A scientist has index *h if h* papers have at least *h citations* each**



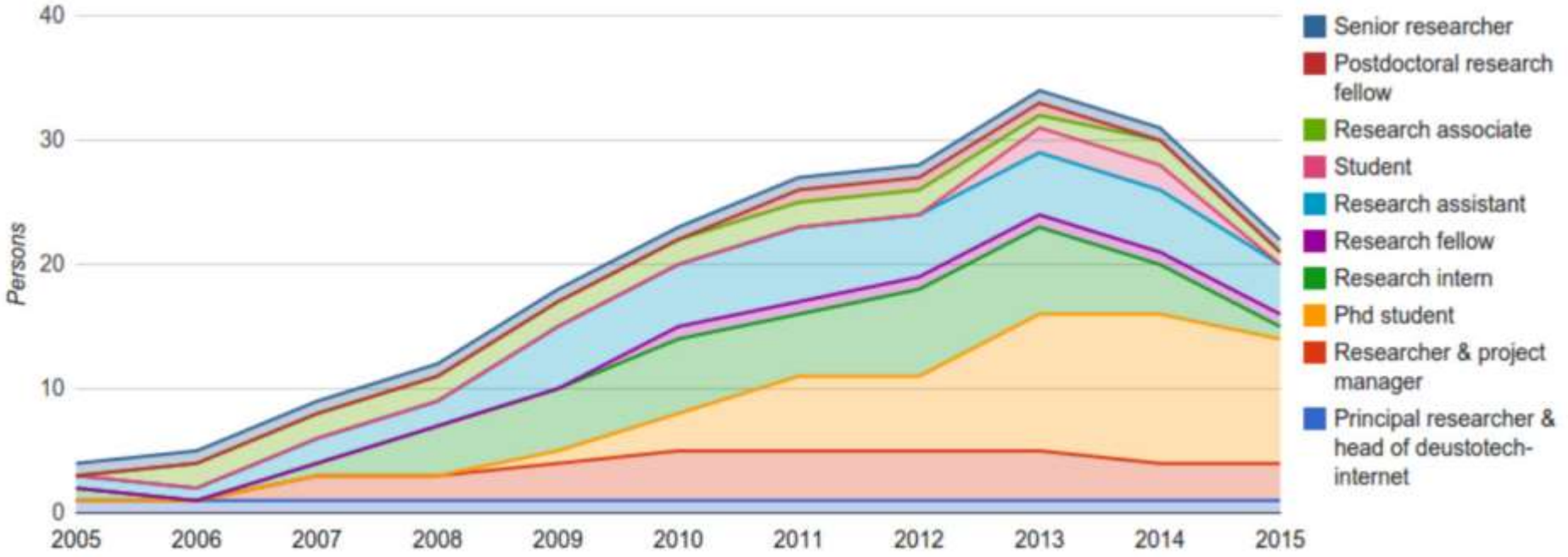
Mapping collaborations and relations



Pena *et al.*
DOI 10.1007/978-3-319-24129-6_25
ISSN 1865-0929

(Conference paper)

Looking at the group

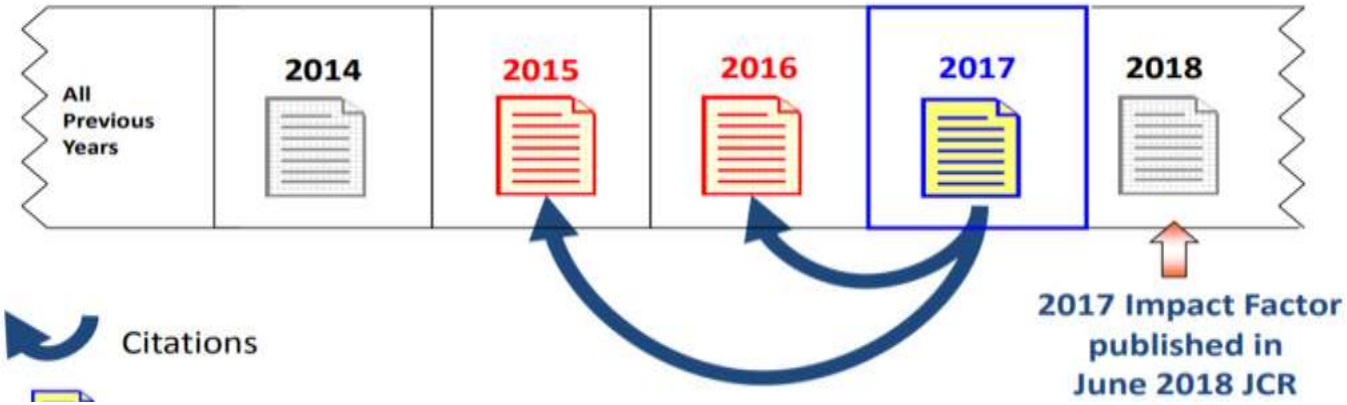


Pena *et al.*
DOI 10.1007/978-3-319-24129-6_25
ISSN 1865-0929

(Journal) Impact factor

The impact factor of a journal for year x, is the ratio between the sum of references in year x to papers in the journal in year x-1 and x-2 divided by the number of publications in year x-1 and x-2 (all based on Web of Science data – ca. 11000 journals) - “Clarivate Analytics Journal Citation Reports”

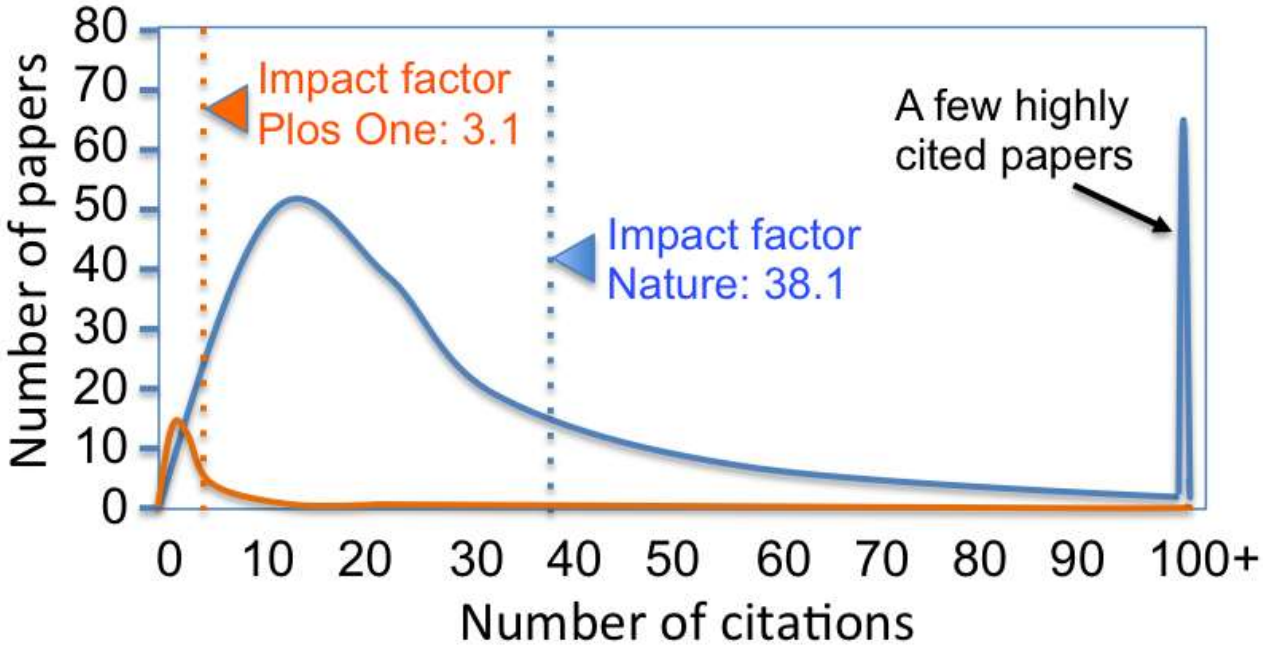
$$IF_x = \frac{Citations_{x-2} + Citations_{x-1}}{Publications_{x-2} + Publications_{x-1}}$$



(Journal) Impact factor …. *Please note:*

- JIF can only be calculated after completing a minimum of 3 years of publication. New journals have NO IF
- JIF is a quotient factor only and is not a quality factor.
- JIF is not related to quality of content and quality of peer review.
- Review Journals which publishes more review articles will get highest impact factors.
- JIF is very discipline dependent.
- JIF is dependent on the database use (WOS)
- JIF may be independent of the real 'impact' of the work among investigators and scientific communities.
- Self-citations and purposeful addition of the same journal article with the favorable editorial policies may cause the journals or publishers to be not considered for the evaluation of the impact factor.

Few articles with many citations determine IF



Journal impact factors are influenced heavily by a small number of highly cited papers. In general, most papers published in 2013–14 received many fewer citations than indicated by the impact factor. Two journals (Nature [blue], PLOS One [orange]) are shown to represent a highly cited and less cited journal, respectively. Note that the high citation impact of Nature is derived from relatively few highly cited papers. Modified after Callaway 2016

Other bibliometric measures

Cited half-life: the median age of the articles that were cited in Journal Citation Reports each year. E.g., if Cited half-life time for a journal in 2017 is 5, it means that half of the articles cited in 2017 were from before 2012 and the other half from 2012-17

CiteScore based on citation counts in a given year (e.g. 2015) to documents published in **three** previous calendar years (e.g. 2012 – 14), divided by the number of documents in these **three** previous years (e.g. 2012 – 14) using Scopus (ca. 36000 journal of which 23000 are still active

Accept rate ratio between number of accepted articles and received articles

Other bibliometric measures

More elaborate bibliometrics

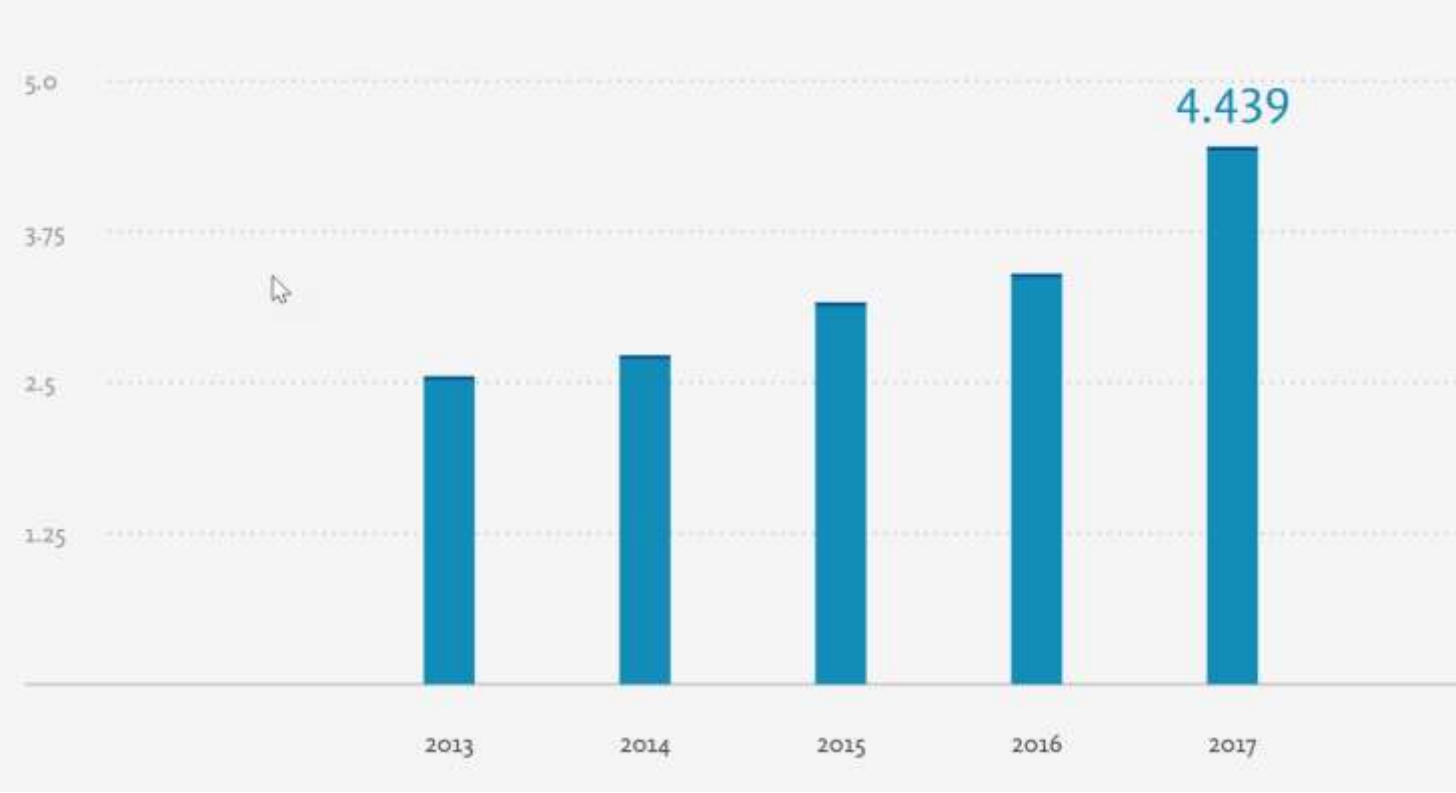
Eigenfactor® score is a rating of the total importance of a specific journal. Journals are rated according to the number of incoming citations, with citations from highly ranked journals weighted to make a larger contribution to the ***Eigenfactor***® than those from poorly ranked journals. As a measure of importance, the ***Eigenfactor***® score scales with the total impact of a journal. Journals generating higher impact to the field tend to have larger ***Eigenfactor***® scores.

Article Influence is calculated by dividing the ***Eigenfactor***® score by the percentage of all articles recorded in the Journal Citation Reports that were published in a specific journal. The ***Article Influence*** score measures the average influence per article of the papers in a journal.

http://www.eigenfactor.org

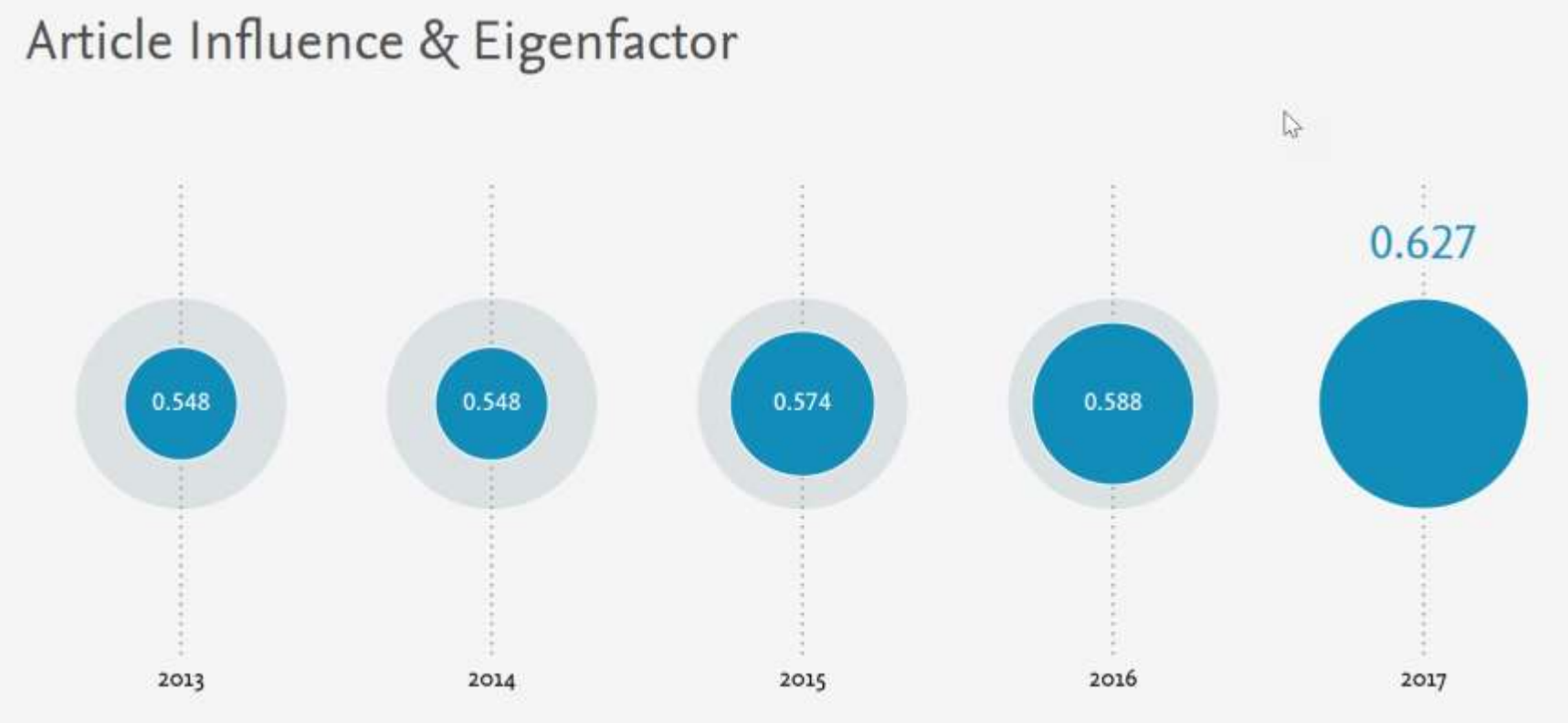
Bibliometric data

Development of a journal



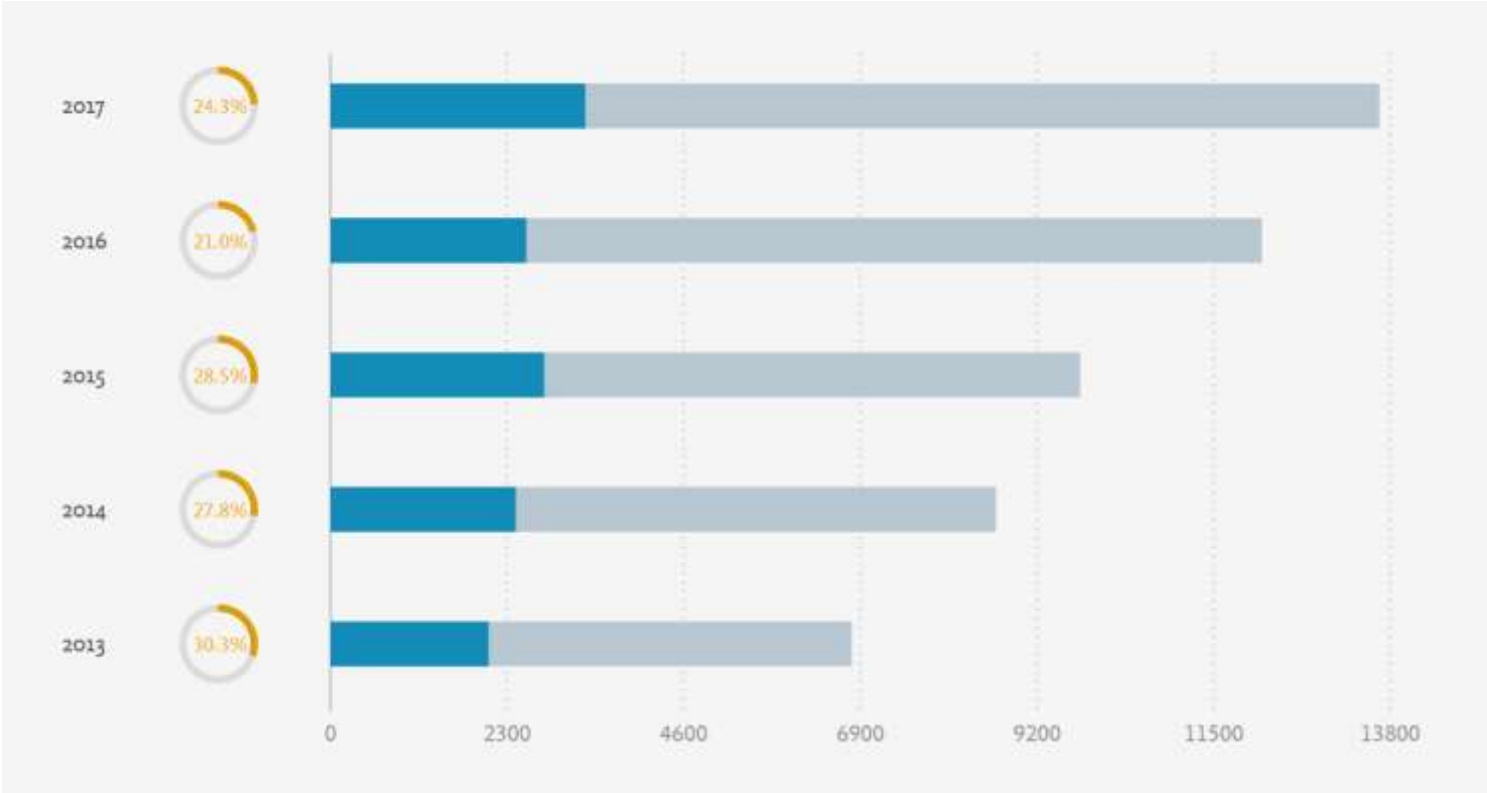
Bibliometric data

Development of a journal



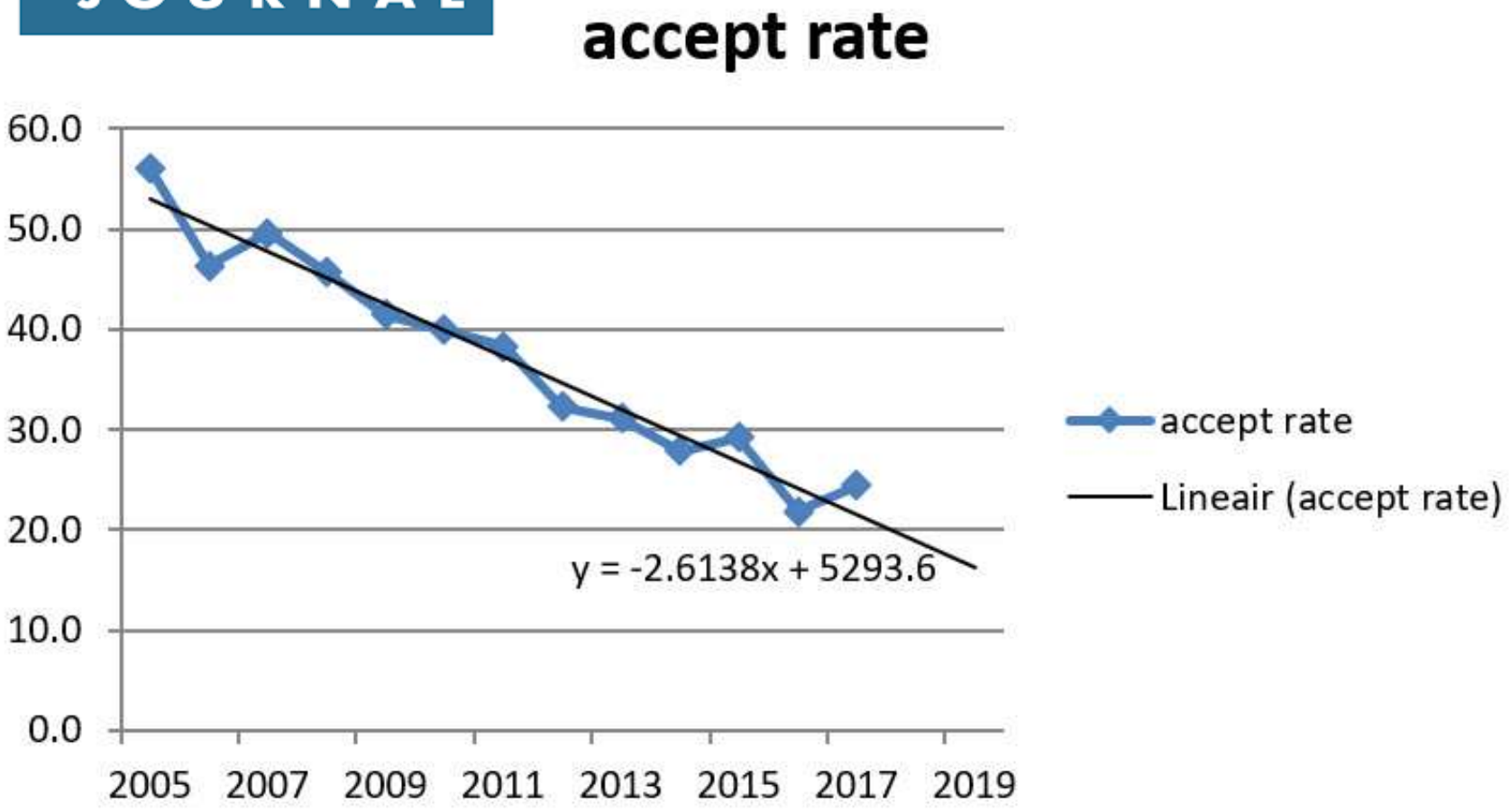
Bibliometric data

Development of a journal



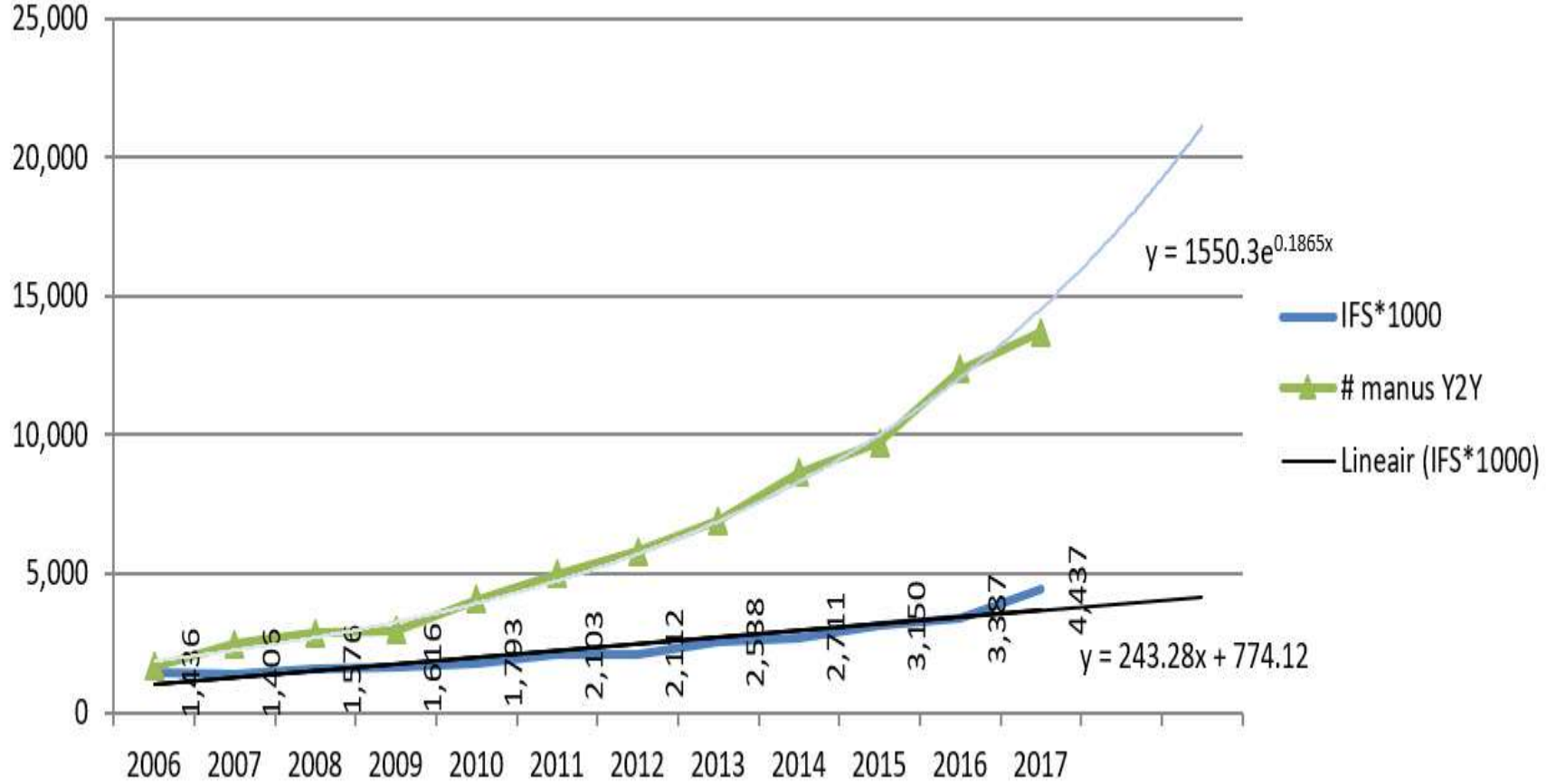
THE JOURNAL

Is Applied Surface Science !!



Publishing Connect

Applied Surface Science



THE JOURNAL

Applied Surface Science

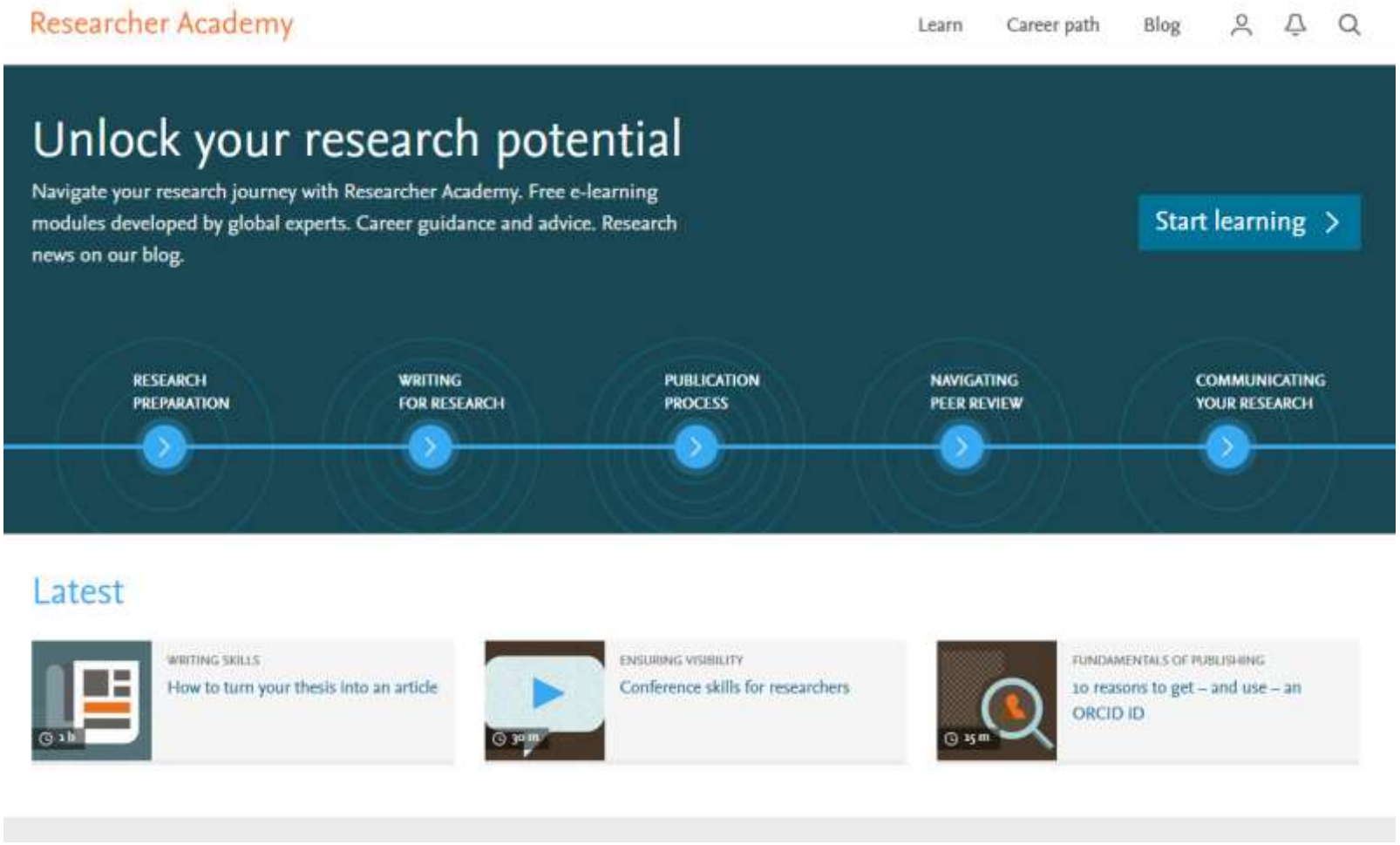
Country	Total # submissions	Accept rate
Argentina	39	43.6
Brazil	189	24.9
Chile	17	23.5
Colombia	20	30.0
Mexico	75	24.3
China	7193	24.2
India	1256	19.7
Overall	13705	24.6

The future for bibliometric measures

- More complex multiparameter measures
- Bibliometric measure “vectors”
- Discipline normalized bibliometric parameters
- Institution subscriptions for insights of own and other scientists

Resources for the (young) scientist

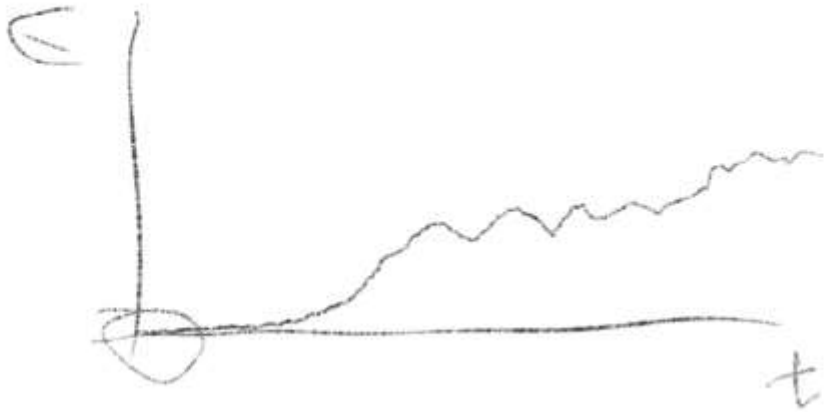
https://researcheracademy.elsevier.com/



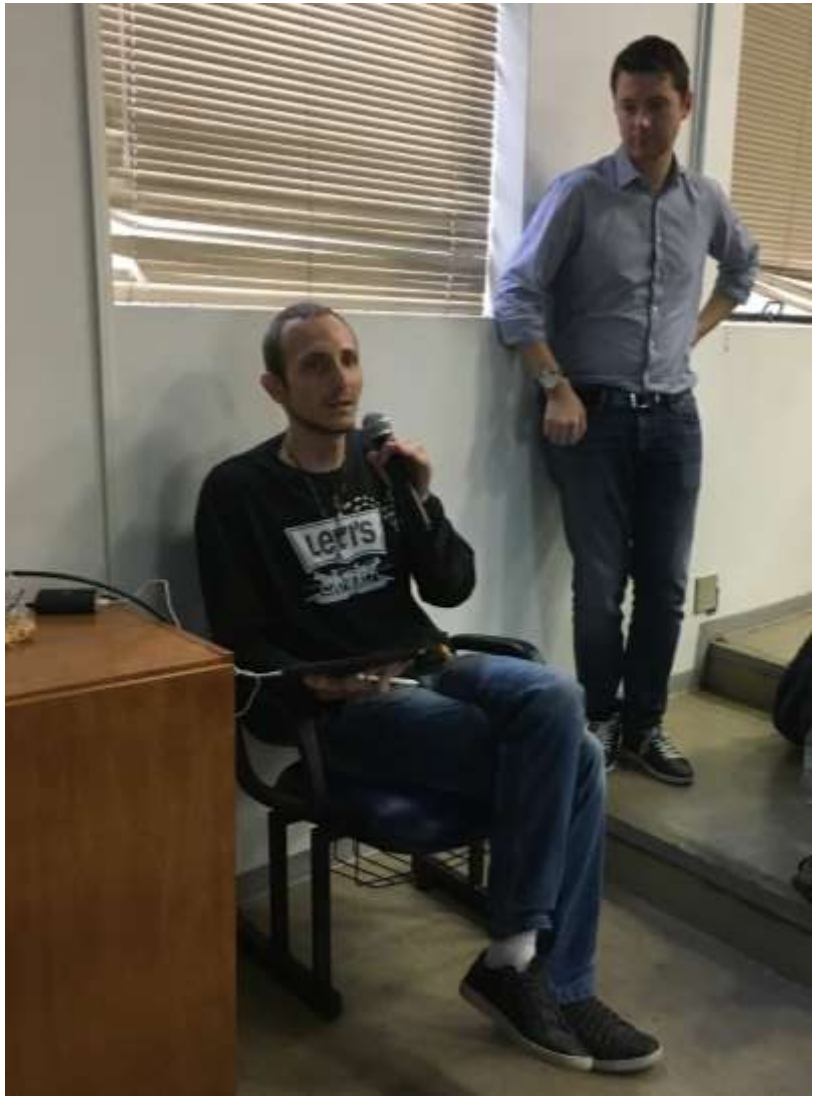
New bibliometric measure from Sao Paulo

$$E(R_{ESS}) = \sum W_j R_{ESS}$$

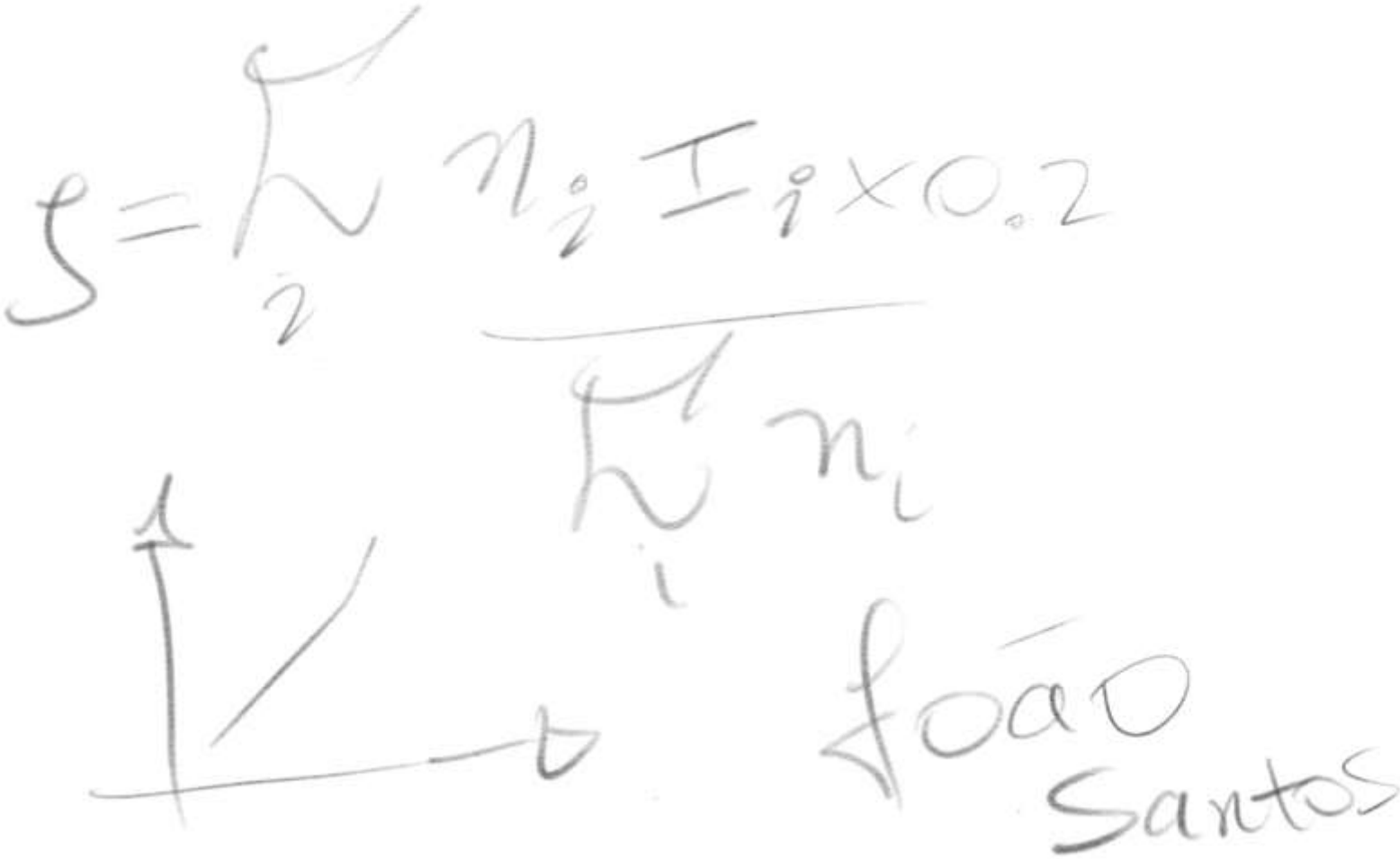
GUILHERME D. ARAUJO



It took a lot of concentration



New bibliometric measure from Natal



Also in Natal it required concentration

Advertisement…..

Thank you