Clarivate

Bibliometrics Workshop

Using scholarly citation metrics in evaluating research performance

Pavel Kasyanov | 25 June 2025

What is bibliometrics?

The econon	nic lives of the poor		Citation Network	
Ву	Banejee, AV (Banejee, Abhijit V.) ; Duflo, E (Duflo	Banejee, AV (Banejee, Abhijit V.) ; Duflo, E (Duflo, Esther)		
Source	JOURNAL OF ECONOMIC PERSPECTIVES Volume: 21 Issue: 1 Page: 141-167 DOI: 10.1257/iep.21.1.141		Create citation alert	
Published	^w Abbiiit Ra	anorioo	750 Times Cited in All	
Indexed				
Document Type	Facts			
Abstract				
	w or da a b b b b b b b b b b b b b b b b b	Abhijit Banerjee The Sveriges Riksbank Prize in Ed Memory of Alfred Nobel 2019 Born: 21 February 1961, Mumbai, Affiliation at the time of the awar of Technology (MIT), Cambridge, Prize motivation: "for their exper alleviating global poverty" Prize share: 1/3	conomic Sciences in India d: Massachusetts Institute , MA, USA imental approach to	
Clarivat	te [™] Mahmoud			

Bibliometrics is the use of statistical methods to analyze scholarly books, articles and other publications

Evaluating research

With bibliometric methods

- Idea first introduced by Dr. Eugene Garfield in 1955
- In 1963, Eugene Garfield established the Institute of Scientific Information (ISI) in Philadelphia
- In 1996, Web of Science was launched (prior to that it existed as separate databases, like Science Citation Index Expanded, Social Sciences Citation Index, etc.)
- In 2008, InCites was launched



Data used in bibliometric analysis

Web of Science

Accelerate your institution's research with an integrated suite of trusted resources



© Clarivate 2023 5



• 28 selection criteria in total

•24

quality criteria to select for editorial rigor and publishing best practice

• 4

impact criteria to select for the most impactful journals

InCites Benchmarking & Analytics

Reports and Visualizations instead of searching



InCites is an analytical environment for all documents indexed in Web of Science Core Collection since 1980.

Web of Science document record

And the data (or metadata) used

The economic lives of the poor

Ву	Banejee, AV (Banejee, Abhijit V.) ; Duflo, E (Duflo, Esther)				
Source	JOURNAL OF ECONOMIC PERSPECTIVES				
	Volume: 21 Issue: 1 Page: 141-167 DOI: 10.1257/jep.21.1.141				
Published	WIN 2007				
Indexed	2007-12-01				
Document Type	Article				
Abstract	The 1990 World Development Report from the World Bank defined the "extremely poor" people of the world as those who are currently living on no more than \$1 per day per person. But how actually does one live on less than \$1 per day? This essay is about the economic lives of the extremely poor: the choices they face, the constraints they grapple with, and the challenges they meet. A number of recent data sets and a body of new research allow us to start building an image of the way the extremely poor				



Clarivate[™]

Citations vs Cited References

What's the difference?

Cited References

- As a rule, lead to **earlier works**
- Their number is fixed in a document
- Can lead either to the documents indexed in Web of Science Core Collection, thus **increasing their times cited**, or lead to the documents outside of the Web of Science Core Collection

Citations

- As a rule, come from the documents published after the cited document
- Their number can change over time
- If they come from the documents that are **not covered** by Web of Science Core Collection, they don't count

What is a scholarly citation?

Is not equal to a quote on a news item or mentioning on social media



Clarivate[®]

Objects of bibliometric studies

What can we measure using bibliometric methods?

- Individual research outputs
- Authors
- Journals
- Organizations
- Countries
- Trending research topics
- Research funders
- Research collaborations
- etc.

"Is this research paper **an influential one**?" "Should we promote **researcher A or researcher B**?" "Which scientific journal is **the most important** in my field?" University rankings, national research assessment programs

"Is genome sequencing **still a hot research topic**?" "Is our foundation **supporting the best** applications?" "Who should we **partner with** to maximize synergies?"



Document types

What can we measure using bibliometric methods?

- Article
- Review article
- Proceedings Paper
- Book chapter
- etc.

- Letters
- Meetings Abstracts
- Editorial Materials
- Poetry
- Dance reveiw
- *etc*.

As Web of Science **Core Collection** indexes its sources cover-to-cover, there might be different document types in the dataset, some of them having greater research value than the others

Clarivate[™]

Subject area schemas

There are dozens, but these are the most frequently used

Web of Science

- Journal-level
- 254 subject areas
- Each Web of Science Core Collection source (and every document published in that source) belongs to at least one and up to four of the areas
- Multidisciplinary: any source that belongs to more than 4 subject areas

Essential Science Indicators

- Journal-level
- 22 subject areas
- Only Articles and Reviews
- Every document only falls into one subject area
- This schema is used for calculating Highly Cited Papers (and, as a result, Highly Cited Researchers)

Citation Topics

- Document-level
- 3 levels: macro, meso, micro
- Thousands of topics on the lower ('micro') level
- Created by clustering Web of Science Core Collection documents using citation links between the documents

Clarivate[®]

Open Access vs Subscription journals

Key differences

Subscription journals

- To publish in them is free of charge for the author
- However, certain subscription journals might offer an author to make their article open access for an additional fee. This offer can be rejected without any influence on the publishing chances
- The "paywall" problem

Open Access journals

- Research results published there are available free of charge
- But the author needs to pay the Article Processing Charge
- Problem #1: finding money to pay the APC
- Problem #2: predatory open access

Basic bibliometric indicators

And what can be measured with them

Basic Bibliometric Indicators





First five indicators

Available by clicking the "Create Citation Report" button on Web of Science Core Collection search results page



Clarivate[™]

Journal Impact Factor

- Available on click at the journal title on Web of Science Core Collection search results page
- Only if the organization is subscribed to Journal Citation Reports

ELECTRICAL ENGINEERING

Publisher name: SPRINGER

Journal Impact Factor ™ 1.8 2022	1.8 Five Year	
JCR Category	Category Rank	Category Quartile
ENGINEERING, ELECTRICAL & ELECTRONIC in SCIE edition	200/275	Q3

Source: Journal Citation Reports 2022. Learn more 🔀

Number of Web of Science documents

Or documents count



Helps answer the following questions:

• How much of a research has been conducted that ended up in publishing its results in one of the top journals in the world?

• What can be measured: author, group of authors, institution, country, etc.

- Important consideration: different research areas have different average timeframes to publish results
- Can be measured over years
- Volume-dependent indicator

Sum of the times cited Or total citations



Helps answer the following questions:

• How did it influence the context of future research?

- What can be measured: author, group of authors, institution, country, etc.
- Important consideration: different research areas have different average bibliography sizes and, as a result, different average citation rates
- Citations accumulate with years
- Volume-dependent indicator

Clarivate[™]

Average citations per item Or citation impact

Times Cited



Helps answer the following questions:

• How does one average document from the dataset influence the context of subsequent research?

• What can be measured: author, group of authors, institution, country, etc.

- Important consideration: different research areas have different average citation rates
- Citations accumulate with years
- It doesn't make a lot of sense to track this indicator over years
- Volume-independent indicator

Self-citations

- It's not bad to occasionally reference your own previous works
- 20-30% of self-citations is generally acceptable
- Web of Science doesn't calculate this percentage automatically

68,203

64,077 Without self-citations

- Excessive self-citation is considered a way of citation manipulation
- Can damage academic reputation
- In modern bibliometric analysis, rather than counting self-citations, it is advised to focus on monitoring excessive concentration of citations coming from the same author or group of authors, organization, source, etc.

Clarivate[™]

H-index

One value to measure both documents and citations

Equals N if in there are N documents in the dataset which were cited N or more times

Documents	Times Cited
1	128
2	64
3	32
4	16
5	8
6	4
7	2
8	1
9	0
10	0



Suggested by Jorge Hirsch in 2003

> 51 H-Index

• Just like times cited, is an indicator of research influence

H-index

- Show the number of really important on the scale of the dataset being analyzed documents
- Influenced by the same inconsistencies across subject areas, years, and document types

One value to measure both documents and citations

- Physicists: came to an informal agreement that for every year spent in science it should increase by 1
- But: there's no such consensus in other subject areas. Even physics can be different.
- Almost ignored in modern bibliometric analysis

Suggested by J	orge Hirsch

51 H-Index

Clarivate[®]

Journal Impact Factor

The Journal Impact Factor (JIF) reflects the average number of citations per article and is used to assess a journal's influence within its field.



Basic Bibliometric Indicators

Wrap-up

Publications Times Cited 348.29 2,617 911,474 Average per item Total Total Number of Documents Sum of the Times Cited **Average Citations per Item** Research influence / Volume of research / Research impact research importance research intensity

905,925 Without self-citations

Self-Citation

Clarivate[™]

447 H-Index

H-Index

Research influence / research importance

Journal Impact Factor ™ 2020 1.479

Journal Impact Factor

Journal influence / Journal importance

What unites this indicators

0 0 Times Cited 0 Publications **Citing Articles** 40 H-Index 9,361 8.51 1.1007,563 Analyze Average per item 7.115 Analyze 8.328 From 1900 v to 2023 v Without self-citations Without self-citations

Unfortunately, neither of these metrics allows us to benchmark versus a generally accepted reference value in order to conclude, is the current value high or low

Almost neither of them (with an exception of number of documents and the Journal Impact Factor) can be analyzed by years

Neither of them is good for comparing research made in different subject areas

These metrics are good for scholarly search and discovery. But what should we use for professional research evaluation?

Let's see in the next section.

Modern bibliometric indicators

And what can be measured with them

What's the benchmark?

Is 735 a lot? How much is it versus the benchmark?

The economic lives of the poor

Ву	Banejee, AV (Banejee, Abhijit V.) ; Duflo, E (Duflo, Esther)	
Source	JOURNAL OF ECONOMIC PERSPECTIVES Volume: 21 Issue: 1 Page: 141-167 DOI: 10.1257/jep.21.1.141	
Published	WIN 2007	
Indexed	2007-12-01	
Document Type	Article	
Abstract	The 1990 World Development Report from the World Bank defined the "extremely poor" people of the world as those who are currently living on no more than \$1 per day per person. But how actually does one live on less than \$1 per day? This essay is about the economic lives of the extremely poor: the choices they face, the constraints they grapple with, and the challenges they meet. A number of recent data sets and a body of new research allow us to start building an image of the way the extremely poor	

Citation Network In Web of Science Core Collection 735 Citations **Create citation alert** 750 Times Cited in All Databases + See more times cited View citing preprints 48 Cited References View Related Records →

Citation normalization

CNCI: Category Normalized Citation Impact

Times cited

CNCI_{document}=

^t Average citations of all the documents of the *same type*, published in the *same year* and in the *same subject area*

CNCl_{document} > 1: research is cited above world average CNCl_{document} < 1: research is cited below world average

$$CNCI_{group of documents} = \frac{NCI_1 + NCI_2 + ... + NCI_N}{N}$$

Category Normalized Citation Impact is calculated automatically in InCites Benchmarking & Analytics

What's the benchmark?

Is 735 a lot? How much is it versus the benchmark?

Use case Analyzing UiTM Category Normalized Citation Impact over years

💮 Global Baseline 🏳 Baseline (MALAYSIA) 💶 – Universiti Teknologi MARA

Clarivate[™]

Use case Top 9 subject areas at the UiTM

Research Area	Rank …	Web of Science ··· Documents	Category Normalized Citation Impact	% Documents in Q1 Journals
7 Engineering & Materials Science	1	765	1	48.4%
2 Chemistry	2	1,455	0.95	37.47%
3 Agriculture, Environment & Ecology	3	986	0.96	34.59%
8 Earth Sciences	4	233	0.96	33.61%
1 Clinical & Life Sciences	5	2,075	0.64	30.92%
5 Physics	6	253	0.68	29.81%
 4 Electrical Engineering, Electronics & Computer Science 	7	1,263	0.47	26.39%
9 Mathematics	8	76	0.57	23.33%
6 Social Sciences	9	1,943	0.67	23.08%

- Subject areas schema: Citation Topics – Macro
- Publication Years: 2020-2024
- Sorted by % documents in Q1 journals

Highly Cited Documents

Papers that were published in the last 10 years and that perform in the top 1% based on the number of citations received when compared to other papers published in the same field in the same year

Their number is normally much smaller than the total number of Web of Science records, which makes it easier to analyze them in more detail

Show the volumes of world-class research

Academic Ranking of World Universities: 20% of a university's position in this ranking is the number of highly cited authors

Highly Cited Documents are also a source of data for Research Fronts

Highly Cited Documents with UiTM affiliation by years

--- Universiti Teknologi MARA

Clarivate[®]

Top UiTM researchers

By the number of Highly Cited Documents

Person Name	Rank …	Web of Science Documents	Web of Science ResearcherID	Highly Cited Papers ÷ …	Category Normalized … Citation Impact	ORCID	Primary Affiliation (and ••• others)
🗌 Jawad, Ali H. 🤗	1	232	J-5364-2019	13	2.46	0000-0002-4827- 9093	Universiti Teknologi MARA +12 affiliation(s)
🗌 Yusoff, Khalid 🖉	2	106	I-7029-2019	12	5.08	n/a	Universiti Teknologi MARA +13 affiliation(s)
🗌 Ghani, Rohana Abdul	3	13	GBC-7791-2022	6	35.82	0000-0002-3520- 8472	Universiti Teknologi MARA +1 affiliation(s)
🗌 Nawawi, Hapizah 🤗	4	221	N-7446-2019	3	1.1	0000-0002-1372- 453X	Universiti Teknologi MARA +18 affiliation(s)
🔲 Thu, Hnin Ei 왿	4	50	AFR-8954-2022	3	1.43	0000-0002-9665- 7650	Universiti Teknologi MARA +12 affiliation(s)

Journal Impact Factor quartiles

Why switch to them?

- The absolute value of a Journal Impact Factor is important
- But it can mean absolutely different level of journal depending on the subject area
- Because different subject areas suppose different average citation rates, it is not recommended to compare journal impact factors across subject areas
- We break down the journals into each subject category, and rank them within each category from highest to lowest
- Then, within each of the subject areas, we break down these ordered lists into 4 equal parts
- Each part represents a quartile, the first quartile being the most impactful and prestigious

Journal impact

UiTM's research output

By Journal Impact Factor quartile

Clarivate[®]

One more important consideration

Regarding the use of journal metrics in evaluating research output

- Journal metrics exist for comparing journals, not individual research outputs published there
- We don't encourage using journal metrics to evaluate individual research papers
- However, it is also important to understand that classifying a university's – or a country's – research outputs retrospectively by Journal Impact Factor quartiles can provide valuable insights on the university's – or a country's – publication strategy

What we've discussed today

Let's recap

- Bibliometrics allows us to analyze scholarly output with statistical methods
- We can use a variety of citation metrics, but it is important to remember what we are measuring, what their limitations are, and how we should interpret our results
- Modern indicators allow us to compare research made in different subject areas

For further reading

- Profiles, not Metrics: <u>https://clarivate.com/webofsciencegroup/campaigns/profiles-not-metrics/</u>
- Research Fronts 2023: <u>https://clarivate.com/blog/tenth-annual-research-fronts-report-highlights-hot-and-emerging-fields/</u>
- InCites indicators handbook: <u>https://incites.help.clarivate.com/Content/Indicators-</u> <u>Handbook/ih-about.htm</u>
- Clarivate Developer Portal: <u>https://developer.clarivate.com/</u>

Thank you

Pavel Kasyanov pavel.kasyanov@clarivate.com

About Clarivate

Clarivate is the leading global information services provider. We connect people and organizations to intelligence they can trust to transform their perspective, their work and our world. Our subscription and technology-based solutions are coupled with deep domain expertise and cover the areas of Academia & Government, Life Sciences & Healthcare and Intellectual Property. For more information, please visit <u>clarivate.com</u>

© 2023 Clarivate

Clarivate and its logo, as well as all other trademarks used herein are trademarks of their respective owners and used under license.