Conventional and Innovative Methods of research Assessment

Najko Jahn, Dr Birgit Schmidt
Göttingen State and University Library

Riga, 15 November 2017

Agenda

14:05 - 14:20 Welcome & Round of introductions

14:20 - 14:40 A review of research assessments principles (Birgit Schmidt)

14:40 - 15:00 Discussion

15:00 - 15:20 Data sources, metrics and methods (Najko Jahn)

15:20 - 15:40 Demo

15:40 - 16:00 Discussion

16:00 - 16:30 Coffee break

16:30 - 17:00 Exercises (and report of outcomes)

17:00 - 17:15 Summary & Close

Dr Birgit Schmidt

Project Coordinator
Scholarly Communication Researcher

Göttingen State and University Library / Electronic Publishing

Projects, e.g. OpenAIRE, OpenUP, FOSTER, EOSCpilot, ERC Open Science Study

Committees, e.g. LIBER, Knowledge Exchange, Research Data Alliance

Member of EC Expert Group on the Future of Scholarly Publishing and Scholarly Communiction

Education: PhD Mathematics, MA LIS

ORCID: orcid.org/0000-0001-8036-5859

Other: @bschmid1

Our back-grounds

- Not research assessment as such but management of research information

- Open Science policies and their implementation, incl. OA monitoring, at institutional, national, European levels

- Data analytics, e.g. in support of institutional information management, proposal writing, benchmarking

RESEARCH ASSESSMENT REVISITED
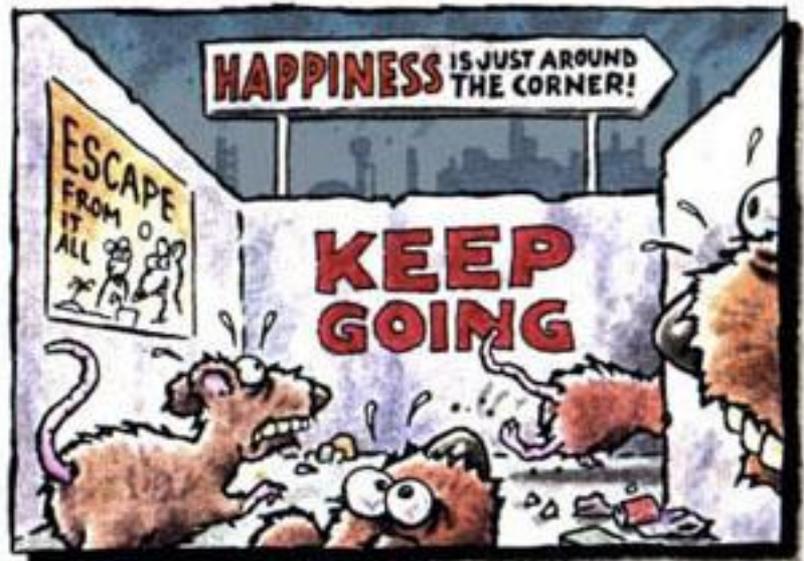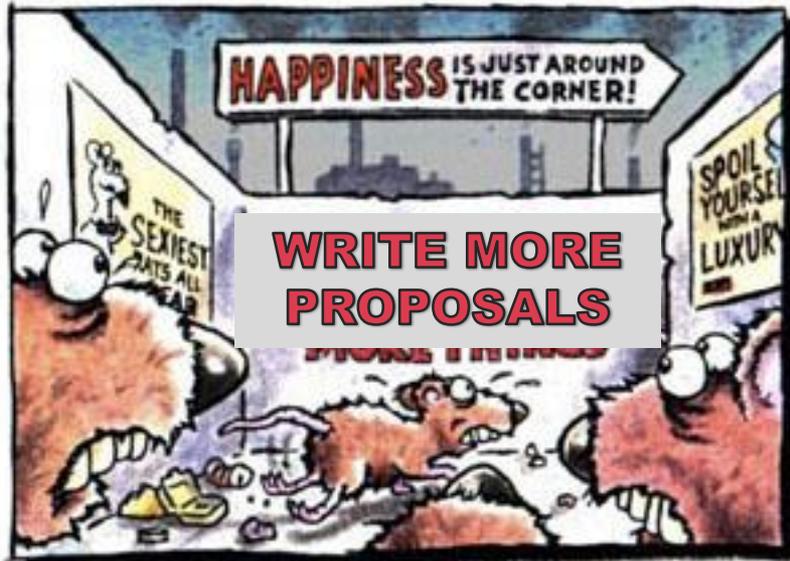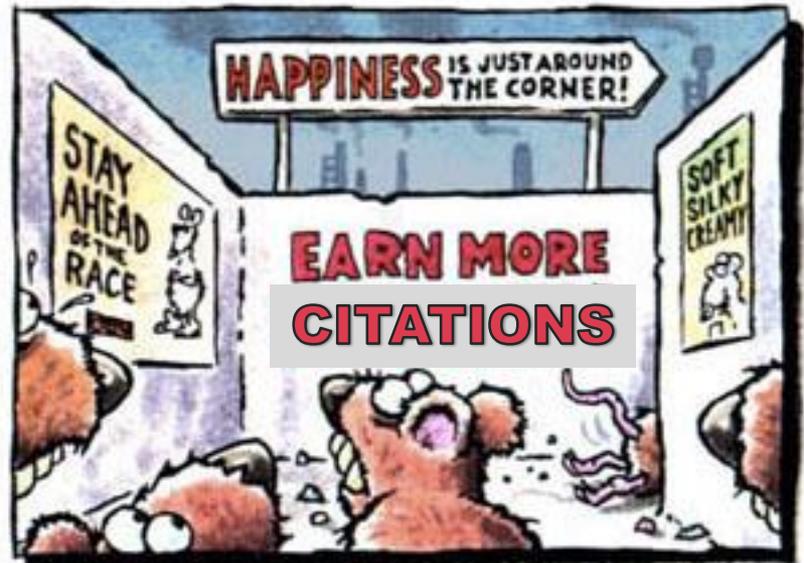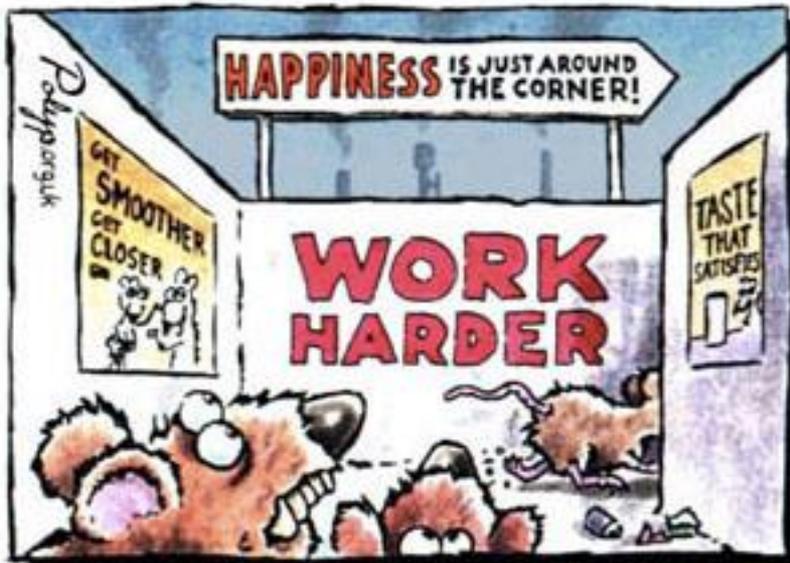
Source: Polyp cartoons, http://www.polyp.org.uk/cartoons/consumerism/polyp_cartoon_Rat_Race.jpg

Research Assessment 101: An introduction

In upcoming issues of Research Trends we dedicate attention to research assessment. Here we explain why we have chosen this subject, how it is defined, its historical background, how the article series is built up, and which topics will be addressed. We also highlight a few fundamental principles that underlie the subsequent articles in the series.

Measuring returns on investment

Research assessment is a broad endeavour. At root it is an attempt to measure the return on investment in scientific-scholarly research. Research assessment includes the evaluation of research quality and measurements of research inputs, outputs and impacts, and embraces both qualitative and quantitative methodologies, including the application of bibliometric indicators and mapping, and peer review.

Research performance is increasingly regarded as a key factor in economic performance and societal welfare. As such, research assessment has become a major issue for a wide range of stakeholders, and there is consequently an increasing focus on research quality and excellence, transparency, accountability, comparability and competition.

https://www.researchtrends.com/issue23-may-2011/research-assessment-101-an-introduction/

Targets & methods

„evaluation of research quality"
>> typically through peer review

„measurement of research inputs and outputs"
>> quantitative measures, comparison with benchmarks

„measurement of impacts"
>> what type of impact is of interest – scholarly, societal/cultural, technological, economic

Aspects of the assessment process

- What is the **unit of the assessment?**
  A country, an institution, a research group, an individual, or a research field or an international network? In which discipline(s) is it active?

- Which **dimension of the research process** must be assessed? Scientific-scholarly impact? Social benefit? Multi-disciplinarity? Participation in international networks?

- What are the **purpose and the objectives** of the assessment? Allocate funding? Improve performance? Increase regional engagement? Which "meta assumptions" can be made on the state of the units of assessment?

Source: Moed & Halevi, 2015

Examples of research assessment

- **Allocation of funding** based on research performance during a previous period
- **Benchmarking of institutions** against as group of (competing) institutions, by region/country/discipline etc.
- Information system to **inform students** about quality of research at institutions
- **Self-assessment** to identify areas which would benefit from a revised strategy
- … and many more

RA Approaches I: University rankings

UL as the only university from Latvia ranks among 1.5 % prestigious institutions in EECA region

UL Communication and innovation department
18.10.2017

The prestigious QS University Rankings show that the University of Latvia (UL) showed a performance, which placed it among the top 1.5 % higher education institutions in EECA region, ranking 44th rank and becoming the only university in Latvia to reach TOP 50.

"The QS University Rankings is one of the leading internationally recognised rankings. By attaining the rank among the top 1.5 % universities in EECA region, UL has shown truly outstanding performance. The results also show the aspects that require particular attention, a greater emphasis should be placed upon creation of the globally cited publications and engagement of international lecturers," asserts the Rector of the UL Indriķis Muižnieks.

The global ranking is based on different indicators, the key ones being the Academic Reputation and Employer Reputation, and 2938 universities have been evaluated.

The University improved on its 2016/17 performance, rising by 3 places. The result is a successful outcome of the UL reforms, resulting in higher number of foreign lecturers, academic staff with doctorate degree, and better academic and employer reputation.

"We are glad to see the UL's continuous development and improving performance which are rewarded, ifs I may say, with a recognition affirmed by entry into these rankings. . The UL has received not only

national but also international evaluation. Most importantly, the University must carry on with this accomplishment and continue to improve the study process and quality, invest in its students, and develop scientific research. We believe in our capacity to continue our progress and attain even higher position in the rank, and we, the students, are ready to work in order to reach this goal," says the Chairperson of the UL Student Council Signe Skutele.

Universities from 18 regional countries are among the TOP 100 in the ranking. The 55th place in the ranking goes to the Rīga Technical University, whereas Rīga Stradiņš University has attained the 108th place, while Latvia University of Agriculture ranks as 139th.

Overall, the universities of Russia prevail in this ranking. The 1st place for the third consecutive year goes to the Mikhail Lomonosov Moscow State University, the 2nd – to the Novosibirsk State University, but the 3rd is taken by the University of Tartu (Estonia), rising by two places since the last year.

More information about the QS University Rankings – EECA rating results can be found here.

The assessment within the current ranking was given to universities of Armenia, Azerbaijan, Belarus, Bosnia and Hercegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Turkey, and Ukraine.

QS WORLD UNIVERSITY RANKINGS

EECA UNIVERSITY RANKINGS 2018
TOP 200

FACT FILE

UNIVERSITY OF LATVIA

Some observations about university rankings

- **Cannot be ignored** (both at national/ regional/international levels) – as it might have effects on funding, rate of international staff and students, etc.

- Regularly leads to **good and bad news** (up in one, down in another, etc.)

- The **group of winners** (top10+) is relatively fixed across all major rankings

- **Strong differences in design**, weighting of indicators and underlying data (a lot is a black box)

RA approaches II: Scholarly impact through the lense of bibliometric indicators

Performance measures based on publication output and citations, derived from **bibliographic databases & related products,** e.g. InCites/Clarivate

Pros
- High-quality data and analytical functionalities

Cons
- Access fees
- Dependency on what the product offers, analysis not fully reproducible
- Many „unknowns": exact scope of data collection, black box indicators/metrics, …
- Biases baked into the product (language, disciplines, regional, gender, etc.)

Why not use the Journal Impact Factor for rankings of publications? (e.g. Nature Chemistry, JIF 2014)
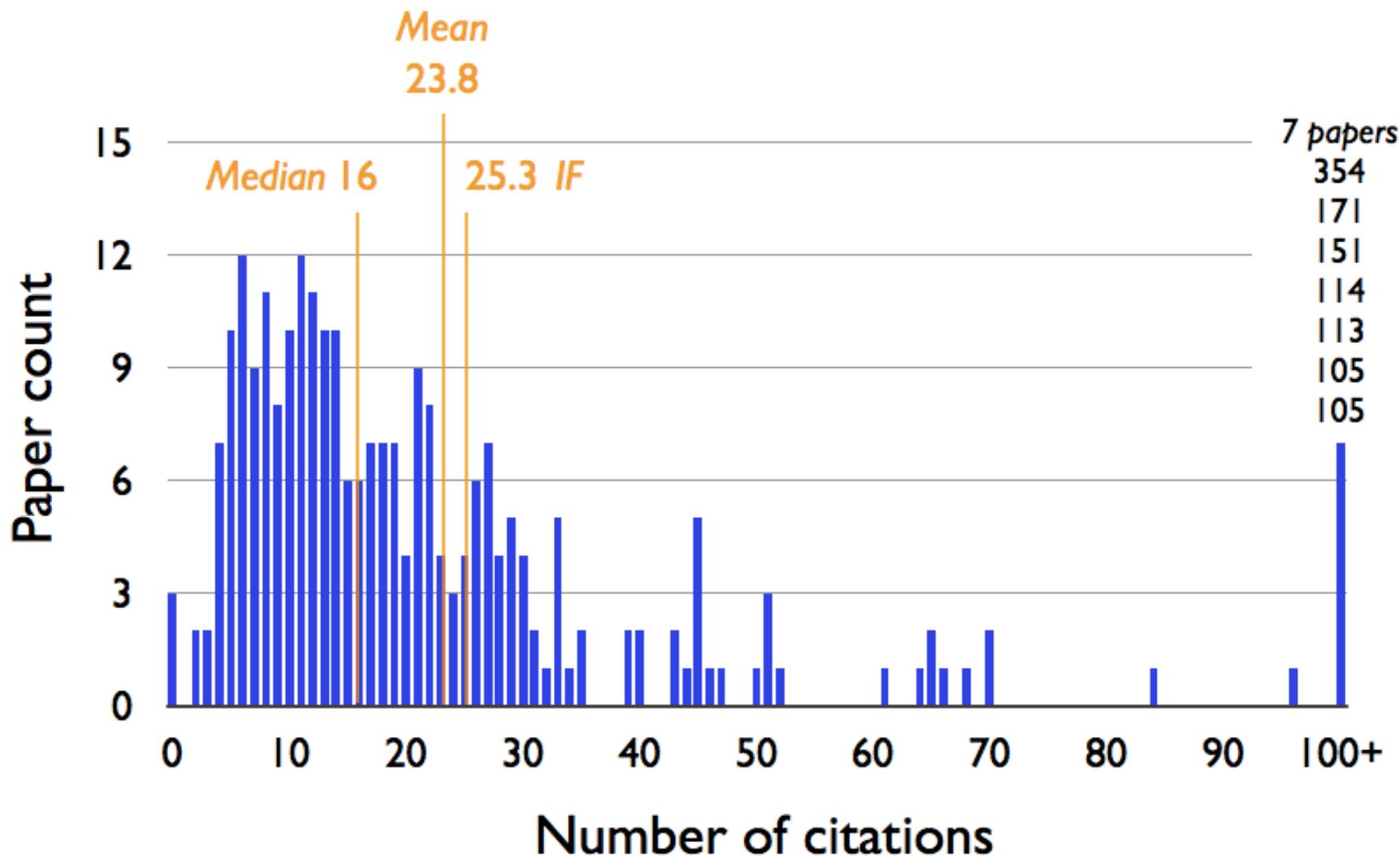
Table 2: Percentage of papers published in 2013-2014 with number of citations below the value of the 2015 JIF.

Journal	JIF	% citable items below JIF
eLife	8.3	71.2%
EMBO J.	9.6	66.9%
J. Informetrics	2.4	68.4%
Nature	38.1	74.8%
Nature Comm.	11.3	74.1%
PLOS Biol.	8.7	66.8%
PLOS Genet.	6.7	65.3%
PLOS ONE	3.1	72.2%
Proc. R. Soc. B	4.8	65.7%
Science	34.7	75.5%
Sci. Rep.	5.2	73.2%

A simple proposal for the publication of journal citation distributions
Vincent Lariviere, Veronique Kiermer, Catriona J MacCallum, Marcia McNutt, Mark Patterson, Bernd Pulverer, Sowmya Swaminathan, Stuart Taylor, Stephen Curry
bioRxiv 062109; doi: https://doi.org/10.1101/062109

San Francisco Declaration on Research Assessment (2012)

San Francisco

D⚙RA

Declaration on Research Assessment

DORA
About DORA
Sign The Declaration
Media Inquiries
Inspiration and Good Practices
A Letter to Thomson Reuters

The San Francisco Declaration on Research Assessment (DORA), initiated by the American Society for Cell Biology (ASCB) together with a group of editors and publishers of scholarly journals, recognizes the need to improve the ways in which the outputs of scientific research are evaluated. The group met in December 2012 during the ASCB Annual Meeting in San Francisco and subsequently circulated a draft declaration among various stakeholders. DORA as it now stands has benefited from input by many of the original signers listed below. It is a worldwide initiative covering all scholarly disciplines. We encourage individuals and organizations who are concerned about the appropriate assessment of scientific research to sign DORA.

The Declaration

There is a pressing need to improve the ways in which the output of scientific research is evaluated by funding agencies, academic institutions, and other parties. To address this issue, a group of editors and publishers of scholarly journals met during the Annual Meeting of The American Society for Cell Biology (ASCB) in San Francisco, CA, on December 16, 2012. The group developed a set of recommendations, referred to as the San Francisco Declaration on Research Assessment. We invite interested parties across all scientific disciplines to indicate their support by adding their names to this Declaration. The outputs from scientific research are many and varied, including: research articles reporting new knowledge, data, reagents, and software; intellectual property; and highly trained young scientists. Funding agencies, institutions that employ scientists, and scientists themselves, all have a desire, and need, to assess the quality and

News About DORA

MRC Preprint Policy Supports Evaluating Research by Content, Not by Where It Is Published.

DORA News Archives

Learn More about DORA

Download the Declaration (PDF)   **Download the DORA Logo (PDF)**   **Download the DORA Poster (PDF)**

For Organizations That Supply Metrics

- Be transparent
- Provide access to data
- Discourage data manipulation
- Provide different metrics for primary literature and reviews

For Publishers

- Cease to promote journals by Impact Factor; provide an array of metrics
- Focus on article-level metrics
- Identify different author contributions
- Open the bibliographic citation data
- Encourage primary literature citations

For Research Institutions

- When hiring and promoting, state that scientific content of a paper, not the JIF of the journal where it was published, is what matters
- Consider value from all outputs and outcomes generated by research

For Funding Agencies

- State that scientific content of a paper, not the JIF of the journal where it was published, is what matters
- Consider value from all outputs and outcomes generated by research

For Researchers

- Focus on content
- Cite primary literature
- Use a range of metrics to show the impact of your work
- Change the culture!

San Francisco

D🌼RA

Declaration on Research Assessment

See the full text of DORA at **www.ascb.org/SFdeclaration.html.** Sign the Declaration!
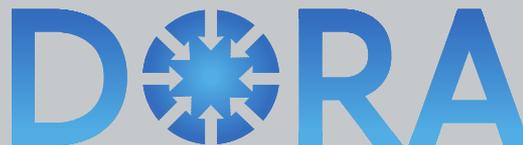
Leiden Manifesto (2015) – Ten Principles

1) **Quantitative evaluation should support qualitative, expert assessment.**

2) Measure performance against the **research missions** of the institution, group or researcher.

3) Protect **excellence in locally relevant research**.

4) **Keep data collection and analytical processes open, transparent and simple**.

5**) Allow those evaluated to verify data and analysis**.

6) Account for **variation by field** in publication and citation practices.

7) Base assessment of individual researchers on a **qualitative judgement of their portfolio.**

8) **Avoid misplaced concreteness and false precision**.

9) Recognize the **systemic effects** of assessment and indicators.

10) **Scrutinize indicators regularly and update them.**

The Leiden Manifesto for Research Metrics

https://vimeo.com/133683418

#1 Quantitative evaluation should support qualitative, expert assessment.

„The sole reliance on citation data provides at best an incomplete and often shallow understanding of research—an understanding that is valid only when reinforced by other judgements.

Numbers are not inherently superior to sound judgments.“

Source: Citation Statistics, A Report of the International Mathematical Union in cooperation with the International Council of Industrial and Applied Mathematics and the Institute of Mathematical statistics, R. Adler, J. Ewing and P. Taylor, Statistical Science (2009), 1-14. Available at

http://arxiv.org/pdf/0910.3529.pdf?origin=publication_detail

#6 Account for variation by field in publication and citation practices.

Journal rankings: very many high-qual international journals, loyalty to particular editors or geographical areas, response to earlier published paper

Citation indices: small community of researchers, very long half-life of papers, long gap between submission and publication, citation data better covered in subject-specific databases

Source: London Mathematical Society, Input for call for evidence, June 2014https://www.lms.ac.uk/sites/lms.ac.uk/files/HEFCEresponseMetrics.pdf

The Metrics Tide (2015)

Report on the role of metrics in research assessment and management

Some conclusions

- Proposal of the notion of **responsible metrics**: robustness, humility, transparency, diversity, reflexivity
- **Recommendations on the application of quantitative indicators**, e.g.
  - awareness of limitations,
  - explicit criteria,
  - greater transparency and interoperability of data providers, data infrastructures and measurement systems,
  - reduced emphasis on the JIF as a promotional tool on publisher websites,
  - specific recommendations related to the REF (Research Excellence Framework, UK), etc.
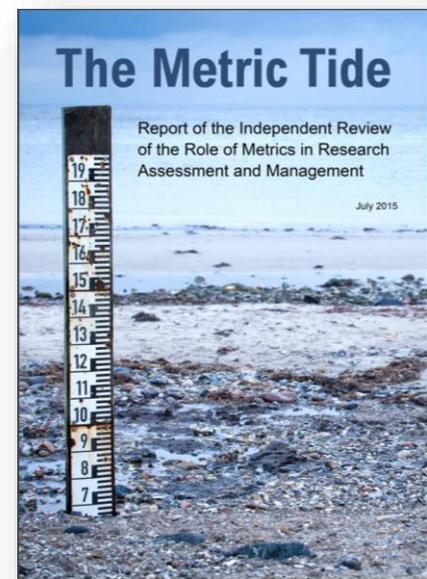
New generation metrics: strength & weaknesses

Review of metrics and their potentials in the context of the EU Open Science policies

Altmetrics: online events which can be recorded and thus be measured (e.g. downloads, likes, tweets, comments)

The strengths of altmetrics can be summarized as follows:

- Broadness - altmetrics can measure not only scholarly influence, but impacts on other audiences as well;

- Diversity – they have the ability to measure different types of research objects (e.g. data, software tools and applications);

- Multi-faceted - the same object can be measured by multiple signals (e.g. comments, tweets, likes, views, downloads);

- Speed - altmetric signals appear faster than conventional metrics.
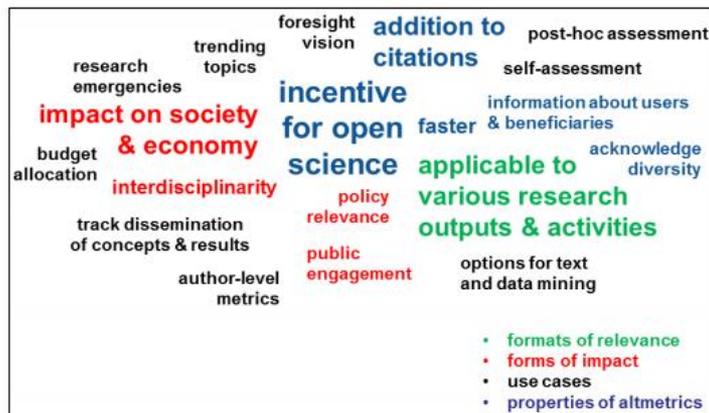
Figure 3. Word cloud compiled from the call for evidence (N=19). The terms reflect the *potential of altmetrics* as described by the respondents and which have been categorized: formats of relevance (green), forms of impact (red), targets and uses (black), and properties of altmetrics (blue).
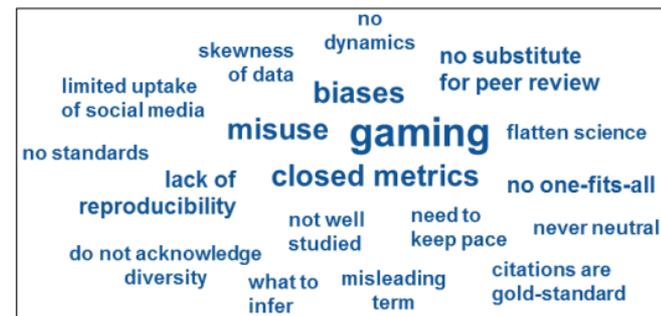
Figure 4. Word cloud compiled from the call for evidence (N=19). The terms reflect the *reason for not using metrics and altmetrics* as described by the respondents.

Recommendations from the report

#1 EC should provide **clear guidance for the responsible use of metrics in support of Open Science**

#2, #3 Encourage the **development of new indicators & assess suitability of existing ones**, to measure and support the development of OS

#4 The **adoption of OS should be recognized and rewarded** through the European research system

#5 Highlight how **inappropriate use of metrics** can impete progress towards OS

#6 **Complement metrics based on private platforms by open metricsadequate research infrastructures for OS**

#7-9 Provide

#10-12 **Embed OS in society**, suggestions for further work

RA Approaches III: Research Excellence Framework UK

1) It costs too much
2) It is not peer review anyway
3) It undermines collegiality
4) It discourages innovation
5) It is redundant
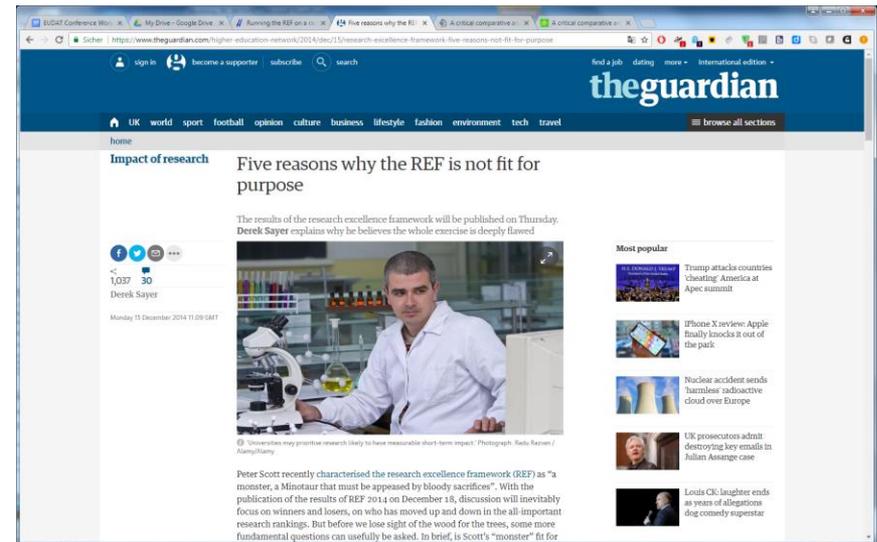
5) It is redundant

11 British universities made the top 100 in the 2013-14 Times Higher Education World University Rankings, which don't use RAE/REF data and rely heavily on the citation metrics that Hefce rejected for the REF. Eight of these universities were in the top 10 in RAE and the other three in the top 20. Other rankings give similar correlations. Almost 85% of Hefce's quality-related funding in 2013 went to Russell Group and former 1994 Group universities and nobody expects the outcomes of REF to significantly change this.

In short, not only is the REF an expensive, cumbersome and divisive procedure that is much more likely to inhibit innovative research than foster research excellence, but it mostly tells us what we already know. It is time it was replaced by something whose costs are more proportionate to its benefits both for the universities and the taxpayer.

Do metrics match peer review? (Harzing, 2017)
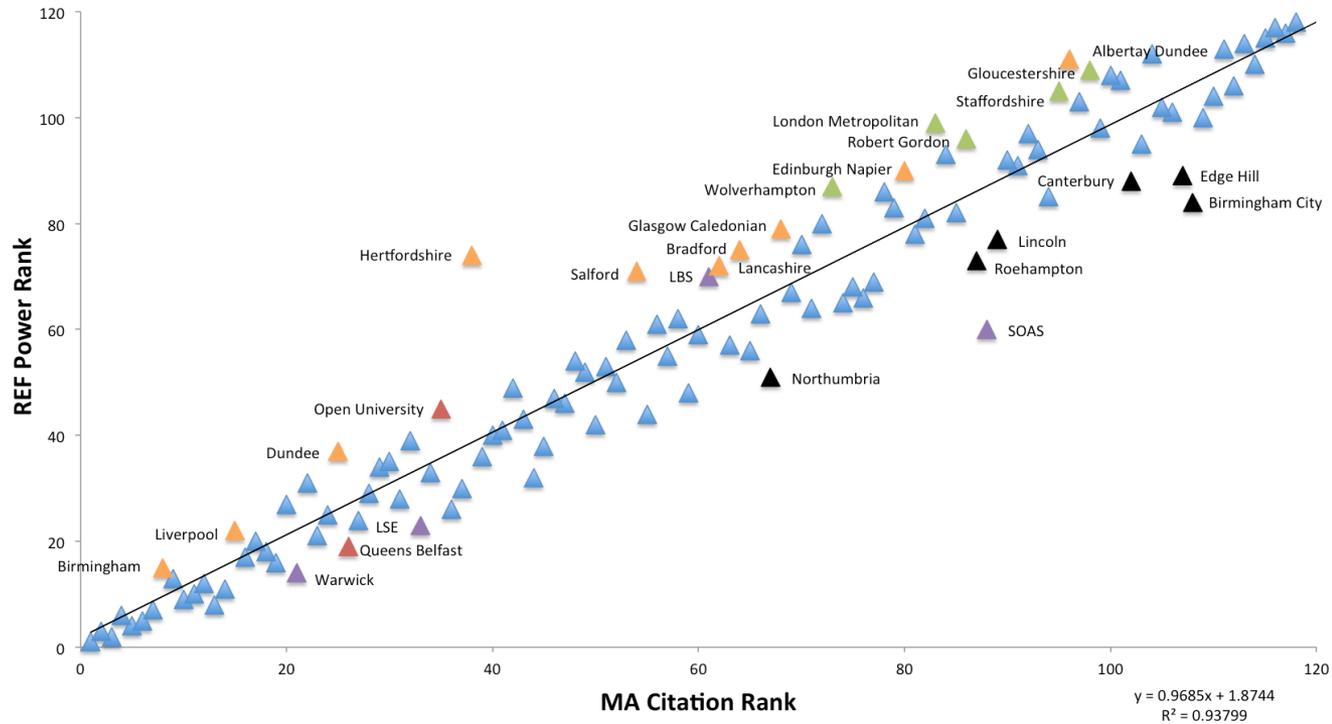
Data from REF2014 vs. citations data from Microsoft Academic (retrieved in July 2017) for articles published 2008-2013

Some issues
- **identifying affiliations** (e.g. too few matched with Queens Belfast, too many with Open University)
- **citation practices differ by discipline**:
  - huge consortia with many citations, e.g. in particle physics and gene technology
  - few authors and lower citation rates in the Social Sciences and Humanities (with some exceptions, e.g. London Business School, London School of Economics)

orange: heavy concentration in the Life Sciences; purple: strong presence in SSH; red: issues with affiliation matching

How about peer review?

The main formal quality assurance mechanism.

However, traditional peer review is also subject to various criticisms (Ross-Hellauer, 2017):

- **Unreliability and inconsistency:** weak levels of agreement, inability to prevent error and fraud, etc.
- **Delay and expense:** delay from submission to publication, resubmission often means new review process, etc.
- **Lack of accountability and risks of subversion:** the „black box" nature of traditional peer review gives a lot of power with the reviewers
- **Social and publication bias**
- **Lack of incentives**
- **Wastefulness**

>> **Open peer review** can address many of these weaknesses

Conclusions: How to make research assessments more reliable

- **Clear definition of procedures** (aims, who, what, when)
- **Well-defined workflows** for the **collection and analysis of data**
- **Careful choice of indicators** (what to measure, what strength, wheat weaknesses)
- Involvement of all **units under assessment** for fact-checking
- Use of **open data wherever possible**
- **Reproducible data analysis**
- **Transparent communication** of observations and recommendations (allow feedback and corrections)

Rewards and/or marketing gimmick

CHINADAILY.COM.CN 中国日报网

US EUROPE AFRICA ASIA
中文 双语 Français

Sign in   Subscribe

HOME   CHINA   WORLD   BUSINESS   LIFESTYLE   CULTURE   TRAVEL   WATCHTHIS   SPORTS   OPINION   REGIONAL   FORUM   NEWSPAPER ⌄   MOBILE

China   News   Society   Innovation   HK/Taiwan/Macao   Cover Story   Photos   Environment   Health   Military

• Home / China / Society

Lancet restaurant gives medical professionals food for thought

By Li Hongyang | China Daily | Updated: 2017-11-02 07:14

f ✔ in +

Xi's Moments

Decoding Xi's report: Chinese wisdom

Xi's report: Unlocking the words

Under the rules of the restaurant, scientists, medical professionals and social scientists are eligible for a discount if they have recently published papers in journals that are included on internet databases such as the Science Citation Index and the Social Sciences Citation Index.

The paper's impact factor is multiplied by 10 to determine the discount, which can account for as much as 30 percent of the bill.

Controls on solid waste tightened

Fishery survey ship nearly ready to sail

State Council News

China promotes Internet Plus manufacturing to boost innovation.

Customers eat lunch at the Lancet Barbecue, which gives discounts to people who have had papers published in academic journals. Wang Zhuangfei / China Daily

28

Discussion

- Your questions & comments?

- What are your good/bad experiences with research assessments? For what purposes?

- What methods have you applied? What are your recommendations?

Measuring impact

References

- AUBR (2010). Assessing Europe's university-based research. Expert Group on the Assessment of University-Based Research. European Commission – DG Research. https://publications.europa.eu/s/dZ9H

- Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., … de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLOS ONE, 12(10), e0185809. doi:10.1371/journal.pone.0185809

- Harzing, A.-W. (2017). Running the REF on a rainy Sunday afternoon: Do metrics match peer review? Available at: https://harzing.com/publications/white-papers/running-the-ref-on-a-rainy-sunday-afternoon-do-metrics-match-peer-review

- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. Nature, 520(7548), 429–431. doi:10.1038/520429a

- Moed, H., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. JASIST 66(10). doi: 10.1002/asi.23314

- Moed, H. F. (2016). A critical comparative analysis of five world university rankings. Scientometrics, 110(2), 967–990. doi:10.1007/s11192-016-2212-y. Open access postprint available at: https://arxiv.org/ftp/arxiv/papers/1611/1611.06547.pdf

- Ross-Hellauer, T. (2017). What is open peer review? A systematic review. F1000Research, 6, 588. doi:10.12688/f1000research.11369.2

- San Francisco Declaration on Research Assessment (DORA), http://www.ascb.org/dora/

- Wilsdon, J., et al. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. DOI: 10.13140/RG.2.1.4929.1363

- Wilsdon, J. et al. (2017). Next-generation metrics: Responsible metrics and evaluation for open science. Report of the European Commission Expert Group on Altmetrics. Available at: https://ec.europa.eu/research/openscience/pdf/report.pdf