

Wilsdon, James et al.

Research Report

Next-generation metrics: Responsible metrics and evaluation for open science. Report of the European Commission Expert Group on Altmetrics

Suggested Citation: Wilsdon, James et al. (2017) : Next-generation metrics: Responsible metrics and evaluation for open science. Report of the European Commission Expert Group on Altmetrics, ISBN 978-92-79-66130-3, European Commission - Directorate-General for Research and Innovation, Brussels,
<https://doi.org/10.2777/337729>

This Version is available at:
<http://hdl.handle.net/11108/321>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.

Next-generation metrics:

Responsible metrics and evaluation for open science

EUROPEAN COMMISSION

Directorate-General for Research and Innovation
Directorate — Policy Development and Coordination
Unit A.6 — *Data, Open Access and Foresight*

Contact: Rene von Schomberg

E-mail: Renevonschomberg@ec.europa.eu

RTD-PUBLICATIONS@ec.europa.eu

*European Commission
B-1049 Brussels*

Next-generation metrics: Responsible metrics and evaluation for open science

Report of the European Commission Expert Group on Altmetrics

James Wilsdon, Professor of Research Policy at University of Sheffield (UK)
Judit Bar-Ilan, Professor of Information Science at Bar-Ilan University (IL)
Robert Frodeman, Professor of Philosophy at the University of North Texas (US)
Elisabeth Lex, Assistant Professor at Graz University of Technology (AT)
Isabella Peters, Professor of Web Science at the Leibniz Information Centre for Economics and at Kiel University (DE)
Paul Wouters, Professor of Scientometrics and Director of the Centre for Science and Technology Studies at Leiden University (NL)

***EUROPE DIRECT is a service to help you find answers
to your questions about the European Union***

Freephone number (*):
00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you)

LEGAL NOTICE

This document has been prepared for the European Commission however it reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

More information on the European Union is available on the internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2017.

PDF

ISBN 978-92-79-66130-3

doi:10.2777/337729

KI-01-17-130-EN-N

© European Union, 2017.

Reproduction is authorised provided the source is acknowledged.

Contents

1	INTRODUCTION: THE OPPORTUNITY OF OPEN SCIENCE	5
1.1	The emergence of open science.....	5
1.2	Incentives and barriers to open science.....	6
1.3	The role of metrics in support of open science.....	7
1.4	Task and approach of the expert group on altmetrics	7
2	AVAILABLE METRICS: THE STATE OF THE ART	8
2.1	Bibliometrics and usage based metrics	8
2.2	Altmetrics.....	9
2.3	Research studies on altmetrics.....	10
2.4	Potentials strengths of altmetrics	11
2.5	Reservations and limitations of altmetrics.....	12
2.6	The assumptions of altmetrics.....	13
2.7	The way forward for altmetrics.....	13
2.8	The demand for next generation metrics	14
3	NEXT GENERATION METRICS FOR OPEN SCIENCE	15
3.1	Headline findings.....	15
3.2	Targeted Recommendations.....	15
3.2.1	Fostering open science	16
3.2.2	Removing barriers to open science.....	16
3.2.3	Developing research infrastructures for open science	16
3.2.4	Embed open science in society	17
4	REFERENCES	18
5	APPENDIX: LIST OF RESPONDENTS TO CALL FOR EVIDENCE.....	21

1 INTRODUCTION: THE OPPORTUNITY OF OPEN SCIENCE

"Open Science is changing every aspect of the scientific method to become more open, inclusive and interdisciplinary...Ensuring Europe is at the forefront of Open Science means promoting open access to scientific data and publications alongside the highest standards of research integrity."

Carlos Moedas, Commissioner for Research, Science and Innovation¹

1.1 The emergence of open science

Science today is in transition – from a relatively closed, disciplinary and profession-based system, toward an open and interdisciplinary structure where knowledge creation is more directly accessible to stakeholders across society. The European Commission gave its strong support to this transition in its 2016 report "Open Innovation, Open Science, Open to the World - A Vision for Europe" (EC, 2016).

Open science represents an innovation in the way research is performed: in how scientists collaborate and share knowledge with the wider world, and how scientific institutions are organized for greater societal impact. It has been driven by both digital technologies and social change. Technology provides us with new modes of communication, and generates huge volumes of data. Knowledge plays an increasing role in all walks of life. These changes have accelerated the globalization of research, while vastly increasing the public availability of scientific knowledge.

The remit of this Expert Group is to review the strengths, weaknesses, and future possibilities of next generation metrics to advance the open science agenda. For some, the ideal result might seem to be the development of a single metric through which to measure open science. We view this as impossible and undesirable. The best outcome – unrealised as yet – would be the development of a suite of metrics that offer a dynamic picture of the progress made toward the goals of open science.

Even this task represents a significant challenge, not least because the attempt to create such a suite of metrics must acknowledge the two-sided nature of the goals of open science. Metric scores typically indicate that some kind of connection has been made between a source and a recipient. Increasing such scores not only depends on 'supply side' efforts by scientific communities to provide better information to society. Just as important is the 'demand side' – the receptivity of society to scientific information and perspectives. Recent political events highlight the fact that this receptivity changes over time, as societal values change. Next generation metrics, then, must attend to both sides of this equation.

The European Commission sees the shift to an open science system as a source of competitive advantage for Europe. Open science covers the cycle of research from conceptualizing to analysing and publishing (see Figure 1). It embraces all types of scientific knowledge, from research data to journal articles to presentation slides, and all manner of stakeholders: researchers, funders, policymakers, citizens, enterprises, and publishers.

Open science is a new approach to the scientific process based on cooperative work, coupled to new tools for collaboration, and new routes for knowledge diffusion through online digital technologies. Open science entails a shift from the standard practice of publishing research results in scientific journals, towards sharing all available data and knowledge at the earliest stages of the research process. It requires a move from 'publishing as fast as possible' to 'sharing knowledge as early as possible'.

¹ Moedas, C. (2016) Introduction to *Open Innovation, Open Science, Open to the World – a Vision for Europe*. Brussels: European Commission. doi:10.2777/061652

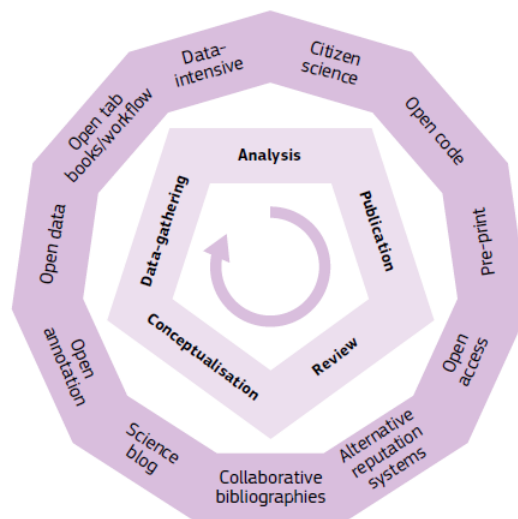


Figure 1: Open Science opens up the entire research enterprise (inner circle) by using a variety of means and digital tools (outer circle) (EC, 2016, p. 36).

1.2 Incentives and barriers to open science

Although open science now enjoys widespread support across scientific and technological communities, institutional and cultural barriers still stand in its way. The current organization of scientific research in disciplines is not sufficiently conducive to the exchange of knowledge across different fields, or between researchers and wider society. The way universities are organized into departments, and set apart from society, can hinder the accessibility of knowledge, expertise and data. Lack of investment in knowledge and data infrastructures may stymie local efforts to foster open science. Funding requirements and the use of proprietary data and instruments in research may restrict open data and open access publishing.

One of the most significant obstacles to the goals of open science lies in the incentive structures of academic research, which can often fail to recognise, value, and reward efforts to open up the scientific process (Hicks et al., 2015; Wilsdon et al., 2015; Wouters et al., 2015; Munafò et al., 2017). As a result, the career advancement of researchers may be hampered if they embrace new ways of working and publishing, rather than fitting within existing systems.

If faster and deeper change is to occur, we need robust data and empirical evidence. Metrics play an important role in any research system, including an open one. However, it is also crucial to understand the inherent limits of metrics. Metrics can be used in ways that impede, rather than accelerate, progress.

In recent years, the debate about this problem has intensified. Some see metrics as contributing to the development of an excessively managerial, audit-driven culture in universities (Collini, 2016; Martin, 2016). Concerns tend to focus on three issues. First, a narrowing of managerial or funder attention onto things that can be measured, at the expense of those that cannot. Second, a reduction in diversity, as an emphasis on particular indicators or league tables drives universities to adopt similar strategic priorities, and individual researchers to focus on lower-risk, incremental work aimed at higher-impact journals (Hicks et al., 2015). Third, a distortion of incentives, which in turn exacerbates problems of research quality, integrity and reproducibility (Benedictus, Miedema & Ferguson, 2016; Sarewitz, 2016).

Many researchers are aware of these shortcomings. As the evaluation of research has become more important, the assessment of societal impacts has grown in importance alongside that of research qualities. An open science system should make both objectives more feasible. But first, a new framework for the development and use of metrics is needed, which we call 'next generation metrics'.

1.3 The role of metrics in support of open science

- Metrics can play two roles in support of open science:
- Monitoring the development of the scientific system towards openness at all levels;
- Measuring performance in order to reward improved ways of working at group and individual level.
- These goals require the development of new indicators, as well as prompting the use of existing metrics in a more responsible fashion.
- There have been a number of high profile recent efforts to address these issues, including:
- *The San Francisco Declaration on Research Assessment (DORA)*, which called in 2012 for research to be assessed on its own merits and for ending the use of journal impact factors in funding, hiring and promotion decisions. By January 2017, DORA has over 800 organisational and 12,500 individual signatories;
- *The Leiden Manifesto*, which was published in 2015 by a group of leading scientometricians, and which sets out ten principles for the use of quantitative indicators in research evaluation (Hicks et al., 2015);
- *Science in Transition*, a movement established in 2013 by researchers in the Netherlands, with the aim of tackling systemic problems in research and university culture, which “has become a self-referential system where quality is measured mostly in bibliometric parameters and where societal relevance is undervalued” (Dijstelbloem et al., 2014);
- *The Metric Tide (2015)*: the report of an independent review of the role of metrics in research assessment and management in the UK system, which set out a framework and targeted recommendations for responsible metrics (Wilsdon et al., 2015).
- These initiatives have informed the European Commission’s Expert Group on Altmetrics, which was set up in 2016.²

1.4 Task and approach of the expert group on altmetrics

Over the past year, the Expert Group has reviewed available metrics, with special attention to altmetrics, and identified frameworks for responsible usage, in the context of the EC’s agenda for open science. This agenda is developing under five action lines: *fostering and creating incentives for open science; removing barriers for open science; mainstreaming and further promoting open access policies; developing an open science cloud; and open science as a socio-economic driver.*

A multi-stakeholder Open Science Policy Platform has been established, to advise on strategic direction and implementation.³ In May 2016, the EU Competitiveness Council issued a set of conclusions on the transition towards an open science system. It noted that the remit of the Open Science Policy Platform should include “adapting reward and evaluation systems, alternative models for open access publishing and management of research data (including archiving), altmetrics....and other aspects of open science.”⁴

This is the context in which the Expert Group on Altmetrics undertook its work, and will input findings to EC policymakers and to the Open Science Policy Platform.

The chair of the group is *James Wilsdon*, Professor of Research Policy at University of Sheffield (UK).⁵ Other members include: *Judit Bar-Ilan*, Professor of Information Science at Bar-Ilan University (IL)⁶; *Robert Frodeman*, Professor of Philosophy at the University of North Texas (US)⁷; *Elisabeth Lex*, Assistant Professor at Graz University of Technology (AT)⁸; *Isabella Peters*, Professor of Web Science at the Leibniz Information Centre for Economics and at Kiel University (DE)⁹; and *Paul Wouters*, Professor of Scientometrics and Director of the Centre for Science and Technology Studies at Leiden University (NL)¹⁰. The group’s work was facilitated by *Rene von Schomberg*, Team Leader for Open Science Policy Coordination and Development in DG Research and Innovation.

² http://ec.europa.eu/research/openscience/index.cfm?pg=altmetrics_eq

³ <http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>

⁴ <http://www.consilium.europa.eu/en/meetings/compet/2016/05/26-27/>

⁵ <https://www.sheffield.ac.uk/politics/people/academic/james-wilsdon>

⁶ <http://is.biu.ac.il/en/judit/>

⁷ <http://philosophy.unt.edu/people/faculty/robert-frodeman>

⁸ <http://www.elisabethlex.info/>

⁹ <http://www.zbw.eu/de/forschung/web-science/isabella-peters/>

¹⁰ <https://www.cwts.nl/people/paulwouters>

This report builds on the expertise of the group members, complemented by desk-research and an extensive literature review. The group also issued a call for evidence in June 2016, to gather the views of stakeholders¹¹. Respondents had one month to reply with brief submissions. They were asked to indicate whether they were making an individual or organisational response, and what role they occupied in the open science agenda. In total, twenty responses to the call for evidence were received, of which nineteen were valid answers. The list of respondents can be found in Appendix 1.

A summary of the results from the call for evidence was presented at the Science and Technology Indicators (STI) Conference in Valencia (September 15, 2016)¹² and the 3AM Conference in Bucharest (September 29, 2016)¹³. Both occasions were used to receive more feedback. The audience at the STI Conference mainly consisted of researchers in scientometrics and bibliometrics, whereas attendees at the 3AM Conference mainly came from research institutes, altmetric providers, and libraries. Feedback was mostly anonymous via plenary contributions and a paper-and-pencil-exercise during the 3AM Conference.

2 AVAILABLE METRICS: THE STATE OF THE ART

Metrics will play an important role in the successful transition to open science. The shift towards web based and networked research has created a market for novel indicators, often grouped together as “altmetrics”. What are the prospects for altmetrics? And can they replace or complement conventional indicators for research quality and impact?

To address these questions, in this chapter, we offer an overview of traditional metrics used in research assessment (see Figure 2). This is followed by an account of altmetrics that includes an explanation of the concept, a summary of the main results of altmetrics research, its strengths and potentials, its limits and problems, and an overview of the changing landscape of altmetrics. We will argue that bibliometrics and altmetrics, in tandem with peer review, offer complementary approaches to evaluation that serve a range of purposes. These findings are drawn from the literature on metrics, answers to our call for evidence, and our own deliberations.

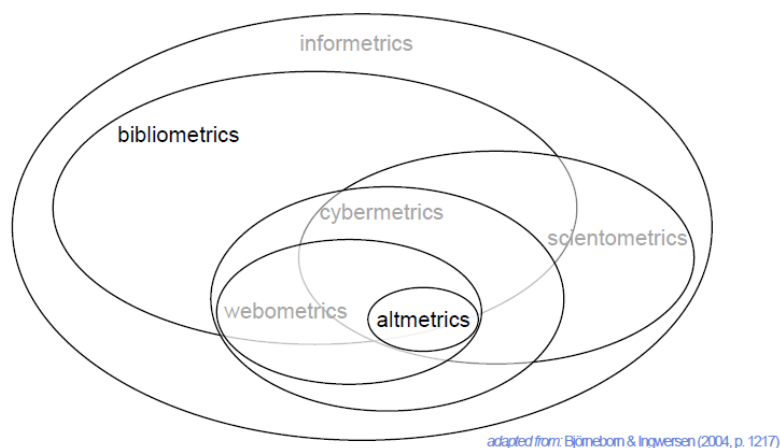


Figure 2: The basket of metrics for the evaluation of science (Haustein, 2015).

2.1 Bibliometrics and usage based metrics

Conventional metrics measure research outputs, mainly journal publications. The two basic types of metrics are the number of publications, and the number of citations the publication receives. These measures may be aggregated at different levels. An object of evaluation can be a single publication, a researcher, a research unit, an institution, or a country.

Citation and publication counts are derived from bibliometric databases (Web of Science – WOS, Scopus and to some extent from Google Scholar). More recently, bibliographic databases have also started to display citation counts of publications. The citation counts are based on the coverage of specific databases and thus vary considerably between the different data sources (Bar-Ilan, 2008).

¹¹http://ec.europa.eu/research/openscience/pdf/call_for_evidence_next_generation_altmetrics.pdf#view=fit&page=none

¹² <http://www.sti2016.org/>

¹³ <http://altmetricsconference.com/>

From the raw data (publication and citation counts) more sophisticated indicators have been created, such as the Journal Impact Factor – JIF (Garfield, 1972), the h-index (Hirsch, 2005), field normalized citation indicators (Waltman & van Eck, 2013), Eigenfactor (Bergstrom, West & Wiseman, 2008), SJR (Gonzalez-Pereira, Guerrero-Bote, & Moya-Anegón, 2010), SNIP (Moed, 2010), or the newly introduced CiteScore (Elsevier, 2016). Publication and citation cultures differ greatly across disciplines and sub-disciplines, which implies a need to normalize indicators when objects are compared from several disciplines or different aggregation levels.

The Journal Impact Factor has been subject to increasing criticism (see DORA, 2012). The main reason for this criticism is the misuse of the JIF, as an indicator of an article's "impact": the JIF is the average for the journal as a whole, and thus does not accurately capture the citation impact of individual articles. That said, it should be noted that neither the Leiden Manifesto nor the Metric Tide simply dismiss the JIF or other metrics, but rather advise that they should be used responsibly.

The h-index is also criticized – for instance, in its bias towards senior researchers, and for not reflecting the impact of highly cited publications – but still is widely used. Bibliometricians argue that it is important not to rely on numbers and indicators alone, but use these numbers together with qualitative assessment -- i.e., the narratives given by peer review -- of the object of evaluation (Moed, 2005; Hicks et al., 2015).

Besides citation and publication counts, bibliometricians also study collaborations, based on co-authorship. The Leiden Ranking¹⁴ provides information on the number of publications that are the result of inter-institutional collaboration, and the percentage of such collaboration out of the total publication of a given university. When conventional bibliometric indicators are used properly they provide valuable insights on the scientific impact of research publications. As already noted, in an evaluation context, best practice is for bibliometric indicators to be used together with qualitative assessment (peer review).

Usage metrics are considered to lie between traditional and alternative metrics. Usage is usually measured by the number of views or downloads of an item. Usage is different from citations, because there are many potential users (students, policy makers, the interested public) who read publications or use data without ever publishing. In addition, not everything a researcher reads is referenced in her publications. Usage based metrics, like the usage impact factor (Kurtz & Bollen, 2008) or lib citations (White et al., 2009) measure attention and uptake.

Usage metrics are highly relevant for open-science, not only in terms of the usage of publications, but also for tracking non-traditional publications (posts, blogs) and for the re-use of open data or open software. Open access publishers provide usage information of individual articles (e.g. PLoS), and commercial publishers are also starting to open up. Several publishers, (e.g. Springer Nature, IEEE, ACM) display the number of downloads of the specific article from their platform. Elsevier's Science Direct in cooperation with Mendeley provides information to researchers on the number of downloads of their publications from the Science Direct platform.

In sum, these indicators serve a valuable function: when they are used responsibly, they are the best quantitative measures available to assess scholarly influence, mainly of publications in journals currently available. This is not to suggest that they cannot be improved.

2.2 Altmetrics

There are additional areas of impact that are not covered by bibliometrics or usage metrics. Some of these are being explored by altmetrics. With the advent of Web 2.0 technologies, new possibilities have appeared for assessing the "impact" of scientific publications, not only of journal publications, but also books, reports, data and other non-traditional publication types. Altmetrics have become a means for measuring the broader societal impacts of scientific research.

The idea was introduced by Neylon and Wu (2009) as "article-level metrics" and by Priem and Hemminger (2010) as "Scientometrics 2.0". The concept of alternative ways to assess scholarly activities was further developed and extended in "Altmetrics: A Manifesto" (Priem, Taraborelli, Groth & Neylon, 2010), and was named "altmetrics" a shorthand for alternative metrics. Altmetrics are based mainly on social media applications, like blogs, Twitter, ResearchGate and Mendeley.

It should be noted that there are constant changes and developments in online applications: new platforms appear (e.g., Loop, WhatsApp, Kudos), while other lose their appeal (e.g., MySpace and, according to some reports, even Facebook) or even disappear (e.g., Connotea, Delicious, ReaderMeter). There are different kinds of measurable signals on social media, e.g., likes, shares,

¹⁴ <http://www.leidenranking.com>

followers, downloads, posts, mentions and comments - each indicating a different level of involvement. Thus, several categorizations of altmetrics have emerged (e.g. Lin & Fenner, 2013; Haustein, 2016).

Priem et al. (2010) emphasized the advantages of using altmetric signals – they are fast relative to citations; cover not only journal publications, but also datasets, code, experimental design, nanopublications, blog posts, comments and tweets; and are diverse, i.e. providing a diversity of signals for the same object (e.g., downloads, likes, and comments).

A theoretical framework was introduced by Haustein et al. (2016), where they define *acts* leading to online *events* that are recorded and thus can be measured. They differentiate between three levels of act categories of increasing engagement:

- Access (e.g. view metadata, accessing content, storing research object)
- Appraise (e.g. on Twitter, on a listserv, in Wikipedia, in mainstream media, in a scientific document, in a policy document)
- Apply (e.g. theories, methods, results, software code, datasets)

Currently there are three major altmetrics aggregators: Altmetric.com, PLUMx and ImpactStory each collect a slightly different set of indicators from primary sources. It should be noted that some of the data sources can be directly accessed through available APIs without the need to subscribe to the commercial aggregators. In addition, commercial aggregators often offer data for research purposes (e.g., Altmetric.com).

Article level indicators are provided by most major publishers (downloads, Mendeley readers, tweets, news mentions etc.). Most of this information is provided by the above-mentioned commercial aggregators. Springer together with Altmetric.com introduced Bookmetrix – that aggregates altmetric data from chapters within the book to the book level (Springer, 2015). Currently, this platform is only available for Springer's books, but such an indicator would be useful for open-access books as well. Author-level altmetrics are provided by Impactstory. PLUMx provides also research unit and institution level altmetrics. Datacite, Zenodo, GitHub and Figshare (and possibly other repositories) provide DOIs for uploaded data, which enables to cite data sources and to track usage, an excellent altmetric for open science.

2.3 Research studies on altmetrics

The availability of countable signals on the Web, the mentioning, discussing, reading and using of scholarly information, has raised the interest of bibliometricians. Empirical studies were and are still being conducted to understand what can be counted and how the altmetric signals relate to traditional indicators (e.g., Eysenbach, 2011; Shuai, Pepe, & Bollen, 2012; Priem, Piwowar & Hemminger, 2012; Li, Thelwall & Giustini, 2012; Mohammadi & Thelwall, 2014; Thelwall, Haustein, Larivière, & Sugimoto, 2013; Haustein et al., 2014; Zahedi, Costas, & Wouters, 2014; Costas, Zahedi, & Wouters, 2015; Bar-Ilan, 2016).

Most of the research concentrates on Mendeley reader counts and tweets. Mendeley was shown to have extensive coverage, which is important because if one seeks to use a specific altmetric for research evaluation, it is not enough that the highly visible articles are covered; one must aim to have values of the metric for all or at least most articles that are being evaluated. It has been shown that Mendeley reader counts correlate at a level of around .5 with citation counts, indicating that there is a relation between the two, but they also measure different aspects of "impact". Despite the lower coverage of scientific publications on Twitter, there is considerable interest in studying this measure. One of the reasons is that tweets spread very fast and can be viewed as early signals of impact; another reason might also be the availability of the underlying data. The main advantage of Twitter over Mendeley that tweets provide some context, although this aspect has not yet been explored.

Additional altmetrics that have been studied include blogs (Shema, Bar-Ilan & Thelwall, 2014 & 2015), F1000Prime (Mohammadi & Thelwall, 2013), data citations (Peters, Kraker, Lex, Gumpenberger, & Gorraiz, 2016) and ResearchGate views (Kraker & Lex, 2015; Thelwall & Kousha, 2016).

There are many empirical studies, but only a few that are concerned with theory and methods. Haustein et al. (2016) list possibly relevant science structure, citation and social theories. Sud and Thelwall (2014) discuss altmetric evaluation strategies and recommend that a range of methods should be used when evaluating altmetrics. Wilsdon et al. (2015) provide an extensive review of the role of metrics, including altmetrics in research assessment. For a more extensive literature review of the topic see (Wouters et al., 2015).

2.4 Potentials strengths of altmetrics

As the responses to our call for evidence have highlighted, there are several ways in which altmetrics could support the transition to open science (Figure 3). Their benefits can be grouped into three categories:

- **formats of relevance:** altmetrics can identify new formats of scholarly products to measure, which have not been considered in research assessments before, e.g., research data and software;
- **forms of impact:** these refer to the new audiences captured, who interact with or react to scholarly products and scenarios related with that, e.g., policy makers and policy documents;
- **targets and uses:** these reflect the purposes for which altmetrics can be used, e.g., budget allocation or self-assessment and career development.

Open science and altmetrics both heavily rely on (open) web-based platforms, encouraging users to contribute (via likes, shares, comments etc.). Altmetrics, then, are both drivers and outcomes of open science practices. More specifically, altmetrics can stimulate the adoption of open science principles, i.e., collaboration, sharing, networking.

Altmetrics also have potential in the assessment of interdisciplinary research and the impact of scientific results on the society as a whole, as they include the views of all stakeholders and not only other scholars (as with citations). Hence, altmetrics can do a better job at acknowledging diversity (of research products, reflections of impact etc.), providing a holistic view of users as well as providers of scientific products, and enhancing exploration of research results.

Although altmetrics are usually viewed as pure quantitative indicators, they offer the option to also analyse qualitative information about users and beneficiaries of scholarly products (e.g., via content analysis of user profiles or comments). As such, they add to the toolbox of qualitative analyses enhancing the understanding of what impact research results really have. In comparison to citations, altmetrics accumulate faster, and may reflect impact almost instantly.

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.

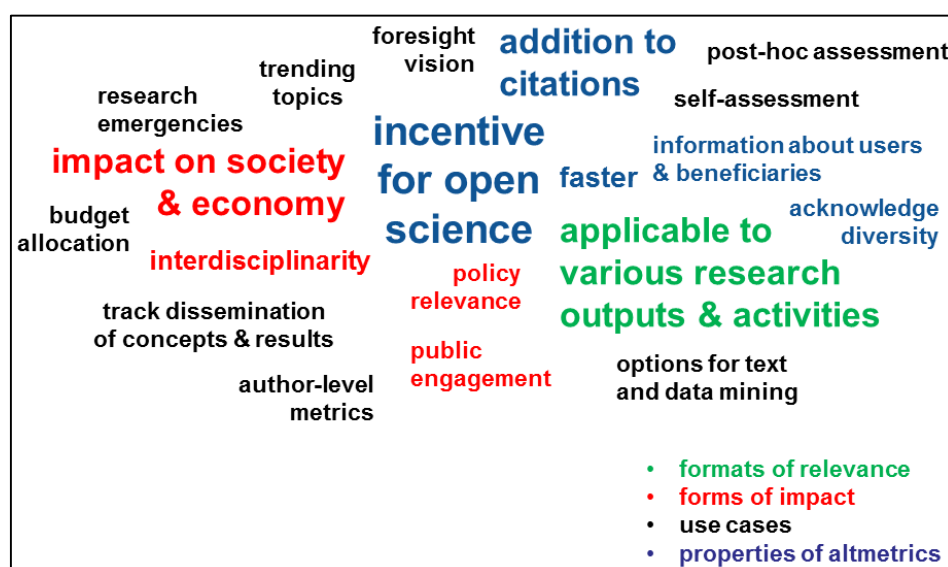


Figure 3. Word cloud compiled from the call for evidence (N=19). The terms reflect the *potential of altmetrics* as described by the respondents and which have been categorized: formats of relevance (green), forms of impact (red), targets and uses (black), and properties of altmetrics (blue).

2.5 Reservations and limitations of altmetrics

A number of challenges surround the use of altmetrics. Perhaps most prominent are the lack of robustness and the ease with which metrics-based evaluation systems can be gamed (Figure 4). Also, Goodhart's Law applies to altmetrics as to all other indicators ("when a measure becomes a target, it ceases to be a good measure"). Additional problems are created by the limited uptake of social media in several disciplines and countries.

A severe problem with altmetrics is the lack of free access to the underlying data. Some of the respondents to our call for evidence found this an insurmountable obstacle to the use of altmetrics: "it does not make sense to use an altmetric for research assessment if the data required to compute it belongs to the commercial sector and can be withdrawn and rendered difficult to access at any time" (INRIA, France; call for evidence). Since the data collection algorithms are assets of the altmetrics providers, and because standards are still in development¹⁵, scrutiny of such altmetric data is not currently possible.

On the other hand, altmetrics providers are also bound by the terms and services of the social media platforms they receive their data from (e.g., Twitter). Hence, data redistribution is often not possible in a legal way. Methods of data collection and self-collected data should be provided publicly. In the responses to the Expert Group's call for evidence, it was suggested that the reference lists should become open access (respondent STI Conference). This would then result in an open access citation database. Moreover, as a start, data from Current Research Information Systems (CRIS) could be made public (respondent STI Conference).

It is important to note that the underlying basis of altmetrics (e.g., sharing and liking behaviour, motivations for sharing, and types of users of social media platforms) is not yet well understood. This has led some to conclude that "metrics are meaningless when the score is unclear" (individual, University Southampton, UK; call for evidence). Also, the scientific community has not yet agreed on the value of altmetrics (e.g., data citations) in comparison to other metrics, such as citations in scholarly articles: "prestige is awarded by the community" (respondent 3AM Conference).



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.

The use of metrics and altmetrics also raise risks for the overall ethics of the system of science (Figure 4). Some fear that altmetrics will introduce a new form of competition not based on scientific quality. This may be advantageous to particular disciplines while others may suffer, which may be to the detriment of scientific and societal progress. As one respondent put it (German Young Academy, call for evidence): "do not mistake reach for benefit or reach for quality". The worry was expressed that researchers will need to keep pace with the development in web-based platforms, which would lead to an additional burden that can limit researchers in unleashing their creativity. A respondent from the 3AM Conference even noted that "we are so focused on measuring scholars by their scholarly output, that we ignore scholars as humans" with personal biographies that, of course, also affect their performance and that are not accounted.

¹⁵ See for example NISO Altmetrics Data Quality Code of Conduct (NISO, 2016)

2.6 The assumptions of altmetrics

Altmetrics fit within a larger ecosystem of evaluation. Traditionally, the evaluation of scientific research has been based upon the opinions of those who were in the best position to judge the merits of work because of experience and training: that is, other researchers, through the processes of peer review.

The concept of a peer has traditionally meant one or another type of expert: biologists judging biologists, and economists judging economists. Non-experts were forced to trust the judgment of experts, which raised the question of moral hazard, the danger that experts would serve their own interests rather than those of the larger community. Part of the promise of metrics is that they seem to avoid this danger, and are thus inherently more democratic: for anyone can judge one number as being greater than another.

This seems to make metrics more desirable – until we remember that every metric relies on a prior account of what can and should be measured. In other words, metrics are always dependent upon a narrative element that explains why a given metric is the one that should be relied upon. Such accounts are often buried in the footnotes; but this does not change the fact that metrics are themselves based upon a prior act of expert judgment concerning what should be measured. We can say, then, that expert judgment operates on two levels -- on the overt level known as peer review, and in a covert fashion, in deciding what and how metrics are to be devised.

This leads us to the view – as reflected by statements such as DORA, the Leiden Manifesto and The Metric Tide – that an act of judgment lies at the roots of any process of measurement. As a result, we add our voice to the conclusion that measurement and narrative, metrics and peer review should be treated as complementary tools of evaluation.

At the same time, it is worth noting the ways in which peer review itself is changing. Peer review is becoming more democratic, as society redefines who counts as a peer. This forms part of the impetus behind the open science movement. In the first instance this has been a matter of expanding the range of academic experts drawn into evaluation, in the name of interdisciplinarity. Peer review today often includes an even wider range of participants, including business people, NGOs and citizens. Such shifts are reflected in the 2010 Altmetrics Manifesto: “With altmetrics, we can crowdsource peer-review. Instead of waiting months for two opinions, an article’s impact might be assessed by thousands of conversations and bookmarks in a week” (Priem et al., 2010).

In sum, every set of metrics reflects certain embedded assumptions. Black-boxing these (ethical, epistemic, and political) assumptions doesn’t make them disappear (Briggle, 2014). It is also important to not mistake the map for the territory: for there may be any number of impacts which may be difficult or impossible to devise a metric for. This is the point of the maxim: “not all that can be counted counts; not all that counts can be counted.” A common danger with attempts at developing a metric is the built-in bias toward what *can* be measured rather than finding other ways to account for the actual effects at work. Too often, metrics start from what can be most easily measured, rather than being reverse engineered from what is most important.

2.7 The way forward for altmetrics

Altmetrics are signals that come earlier than citations. Altmetrics are usually accumulated soon after publication, and then stay more or less constant (especially tweets), while it takes more time for citations to accumulate. Early attention is not necessarily a signal of quality but may be seen as signals of visibility or communication.

In principle, altmetrics have the potential to complement more traditional impact assessment e.g., peer-review or citation-based metrics. Most metrics experts agree, however, that it should complement and not replace current procedures and measures used in research evaluation (e.g. Bornmann, 2014; Haustein, 2016; Wilsdon et al., 2015). Altmetrics can enrich current research assessment by adding new perspectives (e.g.: visibility, societal impact).

Probably the most significant challenge is to understand the meanings and uses of altmetrics. Without gaining a deeper understanding of why and how altmetric acts occur (Zahedi et al., 2014), we are just playing with numbers and do not know what we are counting. To gain a better understanding, qualitative studies should inform quantitative studies of altmetrics on why and how do people use these platforms. Correlations are not sufficient. Currently, we do not know what is the meaning of many of these metrics. Such understanding is needed in order to integrate them with traditional metrics and peer review.

2.8 The demand for next generation metrics

The altmetrics in use today are mainly concerned with identifiable scholarly articles and research data. Altmetrics are broader than bibliometrics, since they consider how other stakeholders than researchers engage with scholarly articles (which might reflect broader impact). However, they also remain too narrow (mainly considering research products with DOIs), given the opportunity to include a wide range of other research products and pieces of scientific knowledge.

Altmetrics have structural problems (e.g., representativeness), serve specific purposes (e.g., they may reflect broader impact, or relate to a wider range of research outputs), and give answers to particular questions (e.g., diversity of opinions), but they are not sufficient as metrics for open science. Rather, altmetrics need to be complemented by metrics and frameworks for use that are tailored to open science priorities.

Herb (2016) suggests that “open metrics” go further than altmetrics by meeting the following criteria:

- research products and data sources for metric development need to be logically selected, open documented, and chosen in line with the disciplinary norms;
- data that underlies metrics, indicators, and measurements need to be open and accessible (preferably via automatic processes, e.g. API);
- provision of software that was used for calculations;
- logical, scientific, and documented explanation of how data were derived and metrics calculated.
- In other words, there remains scope for improvement and a need for what we call **next generation metrics** that measure, reward and create incentives for open science.

3 NEXT GENERATION METRICS FOR OPEN SCIENCE

3.1 Headline findings

Based on our review of the literature, evidence submitted by stakeholders, and deliberations by expert group members, we offer the following **five headline findings**:

#1 An open science system should be grounded in a mix of expert judgement, quantitative and qualitative measures. Metrics cannot offer a one size fits all solution. Moreover, we need greater clarity as to which indicators are most useful for specific contexts. Many available altmetrics are not yet robust or ready to be used for purposes of research evaluation and need to be monitored carefully before they are applied in such contexts. Similarly, many altmetrics do not reflect the transition towards open science, and need to be complemented by metrics and frameworks for use that are tailored to open science priorities.

#2 Transparency and accuracy are crucial (NISO, 2016; Wass 2016). The development and application of metrics should be based on user needs, rather than on the interests of data providers. We reaffirm the conclusion of *The Metric Tide* (Wilsdon et al., 2015) and Leiden Manifesto (Hicks et al., 2015) that responsible metrics should be understood in terms of:

- **Robustness:** basing metrics on the best possible data in terms of accuracy and scope;
- **Humility:** recognising that quantitative evaluation should support – but not supplant – qualitative, expert assessment;
- **Transparency:** keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results;
- **Diversity:** accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system;
- **Reflexivity:** recognising and anticipating the systemic and potential effects of indicators, and updating them in response.

#3 Make better use of existing metrics for open science. Many available indicators can be used to measure the progress of open science. These include usage metrics (e.g., counting views or downloads; and publications saved to reference managers); collaboration (via co-authorship); and societal impact (e.g., tweets, likes, shares and followers). Most of these measures are applicable to closed science as well, which allows for a comparison between open and closed science. These indicators are not perfect; they have limitations and, sometimes, unwanted side effects. Current indicators need to be improved, and should co-exist, with additional indicators and peer review, in the policy and evaluation process.

#4 Next generation metrics should be underpinned by an open, transparent and linked data infrastructure. How underlying data are collected and processed is crucial. If we want agreed upon standardised indicators, we need to develop and promote unique, unambiguous, persistent, verified, open, global identifiers; agreed standard data formats; and agreed standard data semantics. We should also anticipate further dynamic changes in the capabilities and popularity of different social media and other web-platforms. Existing platforms may disappear while new ones may be launched.

#5 Measure what matters: the next generation of metrics should begin with those qualities and impacts that European societies most value and need indices for, rather than those which are most easily collected and measured.

3.2 Targeted Recommendations

To guide further work in this area by the EC Open Science Policy Platform, EC policymakers, research funding bodies and other stakeholders, we have developed a series of **twelve targeted recommendations**. These recommendations are organised under four of the headings of the **European Open Science Agenda**: foster open science; remove barriers to open science; develop research infrastructures for open science; and embed open science in society¹⁶. These recommendations build on and, in some cases, restate those made by earlier initiatives, such as DORA, the Leiden Manifesto and *The Metric Tide*.

¹⁶ The European Open Science Agenda is outlined here:

http://ec.europa.eu/research/openscience/pdf/draft_european_open_science_agenda.pdf

3.2.1 Fostering open science

RECOMMENDATION #1:

Ahead of the launch of its ninth research framework programme (FP9), the EC should provide clear guidelines for the responsible use of metrics in support of open science. In light of other recent initiatives, a growing number of research funders and universities across Europe and internationally are adopting formal policies on the responsible use of metrics in research evaluation, assessment and management. The position of the EC on the use of metrics across its different research funding modes should also include the European Research Council.

RECOMMENDATION #2:

The EC should encourage the development of new indicators, and assess the suitability of existing ones, to measure and support the development of open science. The starting point should be the objectives and desired outcomes of open science, from which indicators can then be developed and aligned. For example, increased collaboration in open science can be measured by comparisons between open and closed environments, or between open-access and pay-walled publications (e.g., in terms of citation advantage, usage, mentions on social platforms). Other indicators could be developed to measure the success of open science (e.g., tracking non-traditional data and information sources, interactive discussions, open pre and post-publication peer reviewing, citizen science and other signals of societal impact). To assess the efficacy of these indicators, they should be evaluated in open and closed systems.

RECOMMENDATION #3:

Before introducing new metrics into evaluation criteria, the EC needs to assess the likely benefits and consequences as part of a programme of 'meta-research'. Most research funding is still discipline, topic or challenge-specific, and not well directed towards solving problems or developing solutions that apply across the entire research enterprise. The EC should invest more resources in the emerging field of 'meta-research' (Ioannidis et al., 2015) and on interdisciplinary research into research systems and practices through the final stages of Horizon 2020 and into FP9. This should include scenario modelling the likely effects of different indicators, and further research into evaluation methods and practices.

3.2.2 Removing barriers to open science

RECOMMENDATION #4:

The adoption and implementation of open science principles and practices should be recognised and rewarded through the European research system, via the currency of citations (by linking every piece of scientific knowledge to a unique and persistent identifier), and in career advancement and funding decisions. The importance of this agenda has been emphasised by the EC High Level Expert Group on the European Open Science Cloud in its recent report (EC High Level Expert Group on EOSC, 2016) and is the on-going focus of work by the EC Expert Group on Reward Systems in Open Science, which will report later in 2017.

RECOMMENDATION #5:

The EC should highlight how the inappropriate use of indicators (whether conventional or altmetrics or next generation metrics) can impede progress towards open science. There is legitimate concern that some of the quantitative indicators being used to support decisions around research quality can be gamed, leading to perverse or unintended consequences. Similarly, there is a strong possibility of emergent alternative indicators will introduce unhelpful biases into the system and ignore other, harder to quantify goods. These consequences need to be identified, acknowledged and addressed in the EC's overall policy position on responsible metrics (see RECOMMENDATION #1).

RECOMMENDATION #6:

In EU research policymaking, funding and evaluation, metrics derived from private platforms should always be accompanied by open metrics to enable proper validation. The EC and other funders should encourage researchers to use private platforms (e.g., Twitter, Facebook, ResearchGate) as an additional – rather than primary – outlet for scholarly communication and collaboration.

3.2.3 Developing research infrastructures for open science

RECOMMENDATION #7:

Realising the vision for the European Open Science Cloud (EOSC) will rely on linked meta-data that can become the basis for open, publicly available data infrastructure. We endorse the October 2016 proposals of the High Level Expert Group on the European Open Science Cloud, and agree with its observations that: "the majority of the challenges to reach a functional

EOSC are social rather than technical” (EC High Level Expert Group on EOSC, 2016). How underlying data are collected and processed – and the extent to which they are open to interrogation – is crucial to addressing these challenges. Without the right identifiers, standards and semantics, we risk developing metrics that are not contextually robust or properly understood. The systems used across EU funding bodies, higher education institutions (HEIs) and other actors in the research system need to interoperate better, and definitions of research-related concepts need to be harmonised. The EOSC should also provide infrastructure that enables scientists to more easily cite datasets, for example by coupling to Datacite.

RECOMMENDATION #8:

The European research system and Open Science Cloud should adopt ORCID as its preferred system of unique identifiers, and an ORCID iD should be mandatory for all applicants and participants in FP9. Unique identifiers for individuals and research works will gradually improve the robustness of metrics and reduce administrative burden. ORCID provides researchers with a unique ID and associates this ID with a regularly updated list of publications. It is already backed by a growing number of funders across Europe (<http://about.orcid.org/>). The EC and ERC should utilise ORCID IDs for grant applications, management and reporting platforms, and the benefits of ORCID need to be better communicated to researchers and other stakeholders (Galsworthy & McKee, 2013).

RECOMMENDATION #9:

The EC should encourage scholarly publishers across Europe to reduce emphasis on journal impact factors as a promotional tool, and only use them in the context of a variety of metrics that provide a richer view of performance. This broader indicator set could include 5-year impact factor, EigenFactor, SCImago, editorial and publication times. Publishers, with the aid of Committee on Publication Ethics (COPE), should encourage responsible authorship practices and the provision of more detailed information about the specific contributions of each author. Publishers should also make available a range of article-level metrics to encourage a shift toward broader assessment based on the academic quality of an article.

3.2.4 Embed open science in society

RECOMMENDATION #10:

The EC should identify mechanisms for promoting best practices, frameworks and standards for responsible use of metrics in support of open science. Building on recent initiatives and existing networks, such as the Research Data Alliance, the EC should use its influence to coax stakeholders in the research system towards more responsible approaches to the use of metrics. Relevant efforts in this context include: DORA, Leiden Manifesto, The Metric Tide, FAIR principles for data sharing¹⁷, Science Europe Position Statement on Research Information Systems (Science Europe, 2016), and NISO Altmetrics Data Code of Conduct (NISO, 2016). Further progress could be achieved by encouraging EU university and research leaders to develop a clear statement of principles on institutional uses of quantitative indicators; encouraging funding, recruitment and promotion panels to be more open about the criteria used for decisions; and encouraging Individual researchers to be more mindful of the limitations of particular indicators in the way they present their own CVs and evaluate the work of colleagues.

RECOMMENDATION #11:

The agenda of this Expert Group should be taken forward by a European Forum for Next Generation Metrics. This forum should note the twin roles of metrics for open science – monitoring the transformation of the scientific system, and restructuring scholarly reward and assessment processes – and ensure close liaison between metrics experts and key stakeholders across the open science system, and with related work underway through the EC High Level Expert Group on the European Open Science Cloud and the EC Expert Group on Reward Systems in Open Science.

RECOMMENDATION #12:

Over the next 18-24 months, the European Forum for Next Generation Metrics should focus on FP9 and the design of a next generation research data infrastructure, which can ensure greater efficiency and interoperability of data collection, and its intelligent and responsible use to inform research strategy, assessment, funding prioritisation and evaluation in support of open science.

¹⁷ The FAIR Principles for Data Sharing are Findable, Accessible, Interoperable and Re-usable.

4 REFERENCES

- Bar-Ilan, J. (2008). Which h-index? — A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257-271.
- Bar-Ilan, J. (2016). Ebooks and ebook chapters on Mendeley: Work in progress. Presented at the Altmetrics16 workshop. Retrieved from http://altmetrics.org/wp-content/uploads/2016/09/altmetrics16_paper_9.pdf
- Benedictus, R., Miedema, F., & Ferguson, M. W. (2016). Fewer numbers, better science. *Nature*, 538(7626), 453-455.
- Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor™ metrics. *The Journal of Neuroscience*, 28(45), 11433-11434.
- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895-903
- Briggle, A. (2014). Opening the black box: The social outcomes of scientific research. *Social Epistemology*, 28(2), 153-166.
- Collini, S. (2016). Who are the spongers now? *London Review of Books*, 38(2).
- Costas, R., Zahedi, Z., & Wouters, P. (2015). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66(10), 2003-2019.
- Dijstelbloem, H., Huisman, F., Miedema, F., & Mijndhardt, W. (2014). *Why science does not work as it should? And what to do about it*. Retrieved from <http://www.scienceintransition.nl/wp-content/uploads/2013/10/Science-in-Transition-Position-Paper-final.pdf>
- DORA (2012). *San Francisco Declaration on Research Assessment*. Retrieved from <http://www.ascb.org/files/SFDeclarationFINAL.pdf?x30490>
- EC (2016). *Open Innovation, Open Science, Open to the World - A Vision for Europe*. Brussels: DG Research and Innovation. doi:10.2777/061652.
- EC High Level Expert Group on the European Open Science Cloud (2016). *Realising the European Open Science Cloud*. Brussels: DG Research and Innovation. doi:10.2777/940154
- Elsevier (2016). *CiteScore metrics FAQ*. Retrieved from https://journalmetrics.scopus.com/downloads/CiteScoreMetrics_FAQ_Scopus_Dec2016_LO.pdf
- Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, 13, e123. doi: 10.2196/jmir.2012
- Galsworthy, M., & McKee, M. (2013). Europe’s “Horizon 2020” science funding programme: how is it shaping up? *Journal of Health Services Research & Policy*, 18(3), 182-185. <http://doi.org/10.1177/1355819613476017>
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178, 471-479.
- Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, 7(1), 113-122.
- Glänzel, W., & Gorraiz, J. (2015). Usage metrics versus altmetrics: Confusing terminology? *Scientometrics*, 102(3), 2161-2164. doi:[10.1007/s11192-014-1472-7](https://doi.org/10.1007/s11192-014-1472-7).
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals’ scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379-391.
- Haustein, S. (2015). Scholarly communication and evaluation: From bibliometrics to altmetrics. Presentation at *COAR-SPARC Conference 2015*, Porto, Portugal (16 April 2015). Retrieved from <http://www.slideshare.net/StefanieHaustein/haustein-coar-sparcapril2015>
- Haustein, S. (2016). Grand challenges in altmetrics. *Scientometrics*, 108, 413-423. doi:10.1007/s11192-016-1910-9
- Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting “altmetrics”: viewing acts on social media through the lens of citation and social theories. In C. R. Sugimoto (Ed.), *Theories of informetrics and scholarly communication. A Festschrift in Honor of Blaise Cronin* (pp. 372-405). Berlin: De Gruyter. Retrieved from <http://arxiv.org/abs/1502.05701>
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1145-1163.
- Herb, U. (2016). Impactmessung, Transparenz & Open Science [Impact Measurement, Transparency & Open Science]. *Young Information Scientist*, 1, 59-72.

- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 16569-16572.
- Ioannidis, J. P. A., Fanelli, D., Dunne, D. D., Goodman, S. N. (2015) Meta-research: Evaluation and Improvement of Research Methods and Practices. *PLOS Biology* 13(10): e1002264
- Kraker, P., & Lex, E. (2015). A critical look at the ResearchGate score as a measure of scientific reputation. In *Proceedings of the Quantifying and Analysing Scholarly Communication on the Web workshop (ASCW'15)*, Web Science Conference.
- Kurtz, M. J., & Bollen, J. (2010). Usage bibliometrics. *Annual Review of Information Science and Technology*, 44(1), 1-64.
- Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*, 91(2), 461-471.
- Lin, J., & Fenner, M. (2013). Altmetrics in evolution: defining and redefining the ontology of article-level metrics. *Information Standards Quarterly*, 25(2), 20-26.
- Martin, B. (2016), What is happening to our universities? *SPRU Working Paper Series*, 2016-03. Brighton: SPRU.
- Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht, Netherlands: Springer
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265-277.
- Mohammadi, E., & Thelwall, M. (2013). Assessing non-standard article impact using F1000 labels. *Scientometrics*, 97(2), 383-395.
- Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, 65(8), 1627-1638.
- Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., du Sert, N. P., ... & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021. doi:10.1038/s41562-016-0021.
- Neylon, C., & Wu, S. (2009). Article-level metrics and the evolution of scientific impact. *PLoS Biology*, 7(11): e1000242. doi: 10.1371/journal.pbio.1000242
- NISO (2016). Altmetrics data quality code of conduct – Draft for public comment. NISO RP-25-201X-3. Retrieved from http://www.niso.org/apps/group_public/document.php?document_id=16121&wg_abbrev=altmetrics-quality
- Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. (2016). Research data explored: An extended analysis of citations and altmetrics. *Scientometrics*, 107(2), 723-744.
- Priem, J. & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday*, 15(7). Retrieved from <http://firstmonday.org/article/view/2874/2570>
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv preprint arXiv:1203.4745*.
- Priem, J., D. Taraborelli, D. Groth, P., & Neylon, C. (2010), Altmetrics: A manifesto, 26 October 2010. <http://altmetrics.org/manifesto>
- Sarewitz, D. (2016). The pressure to publish pushes down quality. *Nature*, 533(7602), 147-147.
- Science Europe (2016). Position Statement on Research Information Systems. Science Europe, November 2016.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2014). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. *Journal of the Association for Information Science and Technology*, 65(5), 1018-1027.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2015). How is research blogged? A content analysis approach. *Journal of the Association for Information Science and Technology*, 66(6), 1136-1149.
- Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations. *PLoS One*, 7(11), e47523.
- Springer. (2015). *Springer and Altmetric to launch new platform for book impact at the London Book Fair: Bookmetrix*. Retrieved from <http://www.springer.com/gp/about-springer/media/press->

[releases/corporate/springer-and-altmetric-to-launch-new-platform-for-book-impact-at-the-london-book-fair--bookmetrix/56006](https://www.springer.com/press-releases/corporate/springer-and-altmetric-to-launch-new-platform-for-book-impact-at-the-london-book-fair--bookmetrix/56006)

Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, 98(2), 1131-1143.

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS One*, 8(5), e64841.

Thelwall, M., & Kousha, K. (2016). ResearchGate articles: Age, discipline, audience size, and impact. *Journal of the Association for Information Science and Technology*, doi:0.1002/asi.23675

Waltman, L., & van Eck, N. J. (2013). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics*, 96(3), 699-716.

Wass, J. (2016). Crossref Event Data. Presented at the altmetrics16 workshop. Retrieved from http://altmetrics.org/wp-content/uploads/2016/09/altmetrics16_paper_4.pdf

White, H. D., Boell, S. K., Yu, H., Davis, M., Wilson, C. S. and Cole, F. T. H. (2009), Libcitations: A measure for comparative assessment of book publications in the humanities and social sciences. *Journal of the American Society for Information Science and Technology*, 60(6), 1083-1096. doi:10.1002/asi.21045

Wilsdon, J. R., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., ..., & Johnson, B. (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

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., de Rijke, S., Rushforht, A., & Franssen, T. (2015). *The Metric Tide: Literature Review* (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). HEFCE. doi: 10.13140/RG.2.1.5066.3520

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, 101(2), 1491-1513.

5 APPENDIX: LIST OF RESPONDENTS TO CALL FOR EVIDENCE

institution that submitted the letter	country	contact person	type of institution (library, non-university center, publisher, etc.)	research university, university, publisher, etc.)
Vetenskap & Allmänhet (VA)	Sweden	Cissi Billgren Askwall (Secretary General), Martin Bergman (Project Manager), Maria Lindholm (Director of Research)	non-profit organisation	membership
University of Southampton	UK	Michael Head	individual	
Mercator Research Institute on Global Commons and Climate Change (MCC) GmbH	Germany	Prof. Dr. Ottmar Edenhofer (Director)	non-university center	research
CNRS-DIST (Department of Scientific and Technical Information)	France	Herbert Gruttemeier	non-university center	research
Frontiers	Switzerland	Frederick Fenter (Executive Editor)	open-access publisher	
Institute for Research in Computer Science & Automation (INRIA)	France	François Sillion (Deputy CEO for Science)	national research institution	
Elsevier	The Netherlands	Dr. Lisa Colledge (Director of Research Metrics)	publisher	
European Physical Society (EPS)	France	David Lee	learned society/membership organisation	
Faculty of 1000 (F1000)	UK	Liz Allen (Director of Strategic Initiatives)	online publisher	
Helmholtz Open Science Coordination Office (Helmholtz-Gemeinschaft Deutscher Forschungszentren e. V)	Germany	Christoph Bruch	non-university center	research
Global Young Academy	Germany	Martin Dominik (Member of GYA Executive Committee)	society of young scientists (world-wide)	
XPANSA Group	Poland	Roman Gurinovich (Director, XPANSA Group)	data analysis company	
(University of Derby)	UK	Prof. Fionn Murtagh	Individual	
University of Craiova	Romania	Adriana Schiopoiu Burlea	Individual	
Joint Information Systems Committee (JISC)	UK	David Kernohan (senior co-design manager)	non-profit-organisation	
League of European Research Universities (LERU)	Belgium	Dr. Paul Ayris	association of universities	
Digital Science	UK	Mike Taylor (Head of Metrics Development)	publisher	

Curtin University		UK		Cameron Neylon	individual
Research (RCUK)	Councils	UK	UK	Dr. Ian Viney	non-departmental public body

How to obtain EU publications

Free publications:

- one copy:
via EU Bookshop (<http://bookshop.europa.eu>);
- more than one copy or posters/maps:
from the European Union's representations (http://ec.europa.eu/represent_en.htm);
from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or
calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (<http://bookshop.europa.eu>).

The Expert Group on Altimetrics outlines in this report how to advance a next-generation metrics in the context of Open Science and delivers an advice corresponding to the following policy lines of the Open Science Agenda: Fostering Open Science, Removing barriers to Open Science, Developing research infrastructures and Embed Open Science in society

Studies and reports

