NEW INDICATORS FOR OPEN SCIENCE

POSSIBLE WAYS OF MEASURING THE UPTAKE AND IMPACT OF OPEN SCIENCE

DIETMAR LAMPERT, MARTINA LINDORFER, ERICH PREM, JÖRG IRRAN AND FERMÍN SERRANO SANZ

KEYWORDS: open science, uptake and impact indicators, altmetrics, new open science indicators, open science evaluation

ABSTRACT

Open science (OS) opens up new ways of creating and sharing knowledge and of disseminating various kinds of results, such as traditional articles, research data, computational and mathematical codes, 3D models, interactive visualisations, or micro-insights. Moreover, OS offers the chance to introduce new ways of evaluating science in a more nuanced, fair, and precise way. As the recent ‘altmetrics’ push has shown, there is wide agreement that conventional approaches to science evaluation are inadequate. With the open movement becoming stronger – especially in science –, it is a good time to reflect on potential new indicators to gauge the uptake and impact of OS. This conceptual work aims to offer a vantage point for more substantial discussions among the key stakeholders.

INTRODUCTION

The traditional way of evaluating science comprises, among other things, an approximation of impact – typically the number of received citations (cf. Mingers & Leydesdorff 2015; Garfield, Eugene, and Alfred Welljams-Dorof 1992, Weingart 2005). Although the advent of the internet made it significantly easier to calculate such indicators, warning signs of their misuse appeared early (cf. e.g. Kostoff 1998, Gläser et al. 2002, Butler 2003, or Weingart 2005). The critique – not just of the misuse and unintended bad consequences of indicators but also policies and practices that had adverse systemic effects on scientific research – culminated in a series of declarations and manifestos that went hand in hand with the widening of the open movement.

This article does not intend to examine OS (Open Science) from a history point of view (for that, cf. David 1998, 2003; Willinsky 2005; Bartling & Friesike 2014a) nor does it examine the various concepts of OS (for that, cf. Bartling & Friesike 2014b; Fecher & Friesike 2014; Buschmann et al. 2015; Delfanti & Pirelli 2015). It suffices to recognise that Open science (OS) does not only open up new ways of creating and sharing knowledge, of disseminating results of individual components along the scientific research process but of evaluating science more nuanced, precise, and fair. This means that much is expected of OS, which is further underlined by a recent report by the OECD (2015a) that states that the positive factors associated with OS are, for instance, increasing transparency and quality in the research validation process, improving efficiency in science, or increasing the knowledge spillovers to the economy.

This is not only a matter of technological developments but also of change in cultural practices. It is yet unclear how the uptake and impact of OS practice ought to be monitored and measured, on research in general but on society in particular. This article is based on the results of a study on Open Science conducted for the European Commission and represents conceptual work as, so far, no substantial work has been done before in this regard.

In the literature review and in the interviews with OS experts that we conducted, there is a general consent that possible new indicators for the monitoring and assessment of scientific production and its impact need to be agreed on by all stakeholder groups, in light of a major redesign of the scientific process provoked by OS.

WHAT IS ALREADY BEING MEASURED

The most prominent attempt to move beyond the traditional impact indicators and towards more open, extensive ones is altmetrics (cf. Priem, Piwowar, Hemminger 2012; Galloway & Pease 2013; Bornmann 2014). Although it employs indicators that are enabled by new technology and extend their reach to capture impact on society, the concept is still in its infancy. Moreover, it is yet unclear how altmetrics can actually signify (cf. Mingers & Leydesdorff 2015).

It has become evident (cf. EC 2015), that new evaluation systems are needed – evaluation of research that is not solely based on bibliometric indicators and that does take into account the whole array of contributions to and resulting from the research process (data, methods, code, insights, ideas, trainings, participations in all kinds of activities, etc.). One can see that altmetrics do not go far enough – the open concepts involved in OS exceed that scope by far.

---

1 see e.g. the Budapest Open Access Initiative (2002), the Berlin declaration on Open Access (2003), the Declaration on Access to Research Data from Public Funding (2004), the Open Science Rome Declaration (2012), the San Francisco Declaration on Research Assessment (2013), the Liber Statement on Open Science (2014), the Leiden Manifesto for research metrics (2015), or even the Amsterdam Call for Action on Open Science (2016).

2 tender SMART 2014/0007 “Open Digital Science”
WHAT TO ACTUALLY MEASURE

One of the main objectives of the study underlying this article was to propose a framework for an OS observatory which monitors the progress of OS in Europe on a continuous basis. The indicators suggested in the article shall therefore be useful to monitor the uptake and impact of OS. Also, indicators shall measure if OS practices make science more accessible for a wider audience, whereby Fecher and Friesike (2014:19) see accessibility in the double sense: (a) accessibility of the research process and (b) comprehensibility of the research result. This understanding suggests that the relationship between science and society must be reflected in the indicators in any case.

Unbeknown to the project, RAND Europe had been tasked by the European Commission to develop the Open Science Monitor that was to accommodate a whole range of indicators to monitor and measure OS trends in the EU. They conducted their work in parallel to our project. Before the writing of this article, we had a chance to scrutinise their results (Smith et al. 2016), which yielded similarities but also differences compared to our results, which we will mention below in the indicators sections.

METHODOLOGY

To come up with reasonably sound results, our project employed a mix of methods that started with a thorough desk research on the status quo of OS concepts, metrics, good practices, policies, programmes, and stakeholders, predominantly in the EU. To better understand the technology characteristics inherited by OS and to predict its potential evolution in the near future from a technological point of view, a trend analysis was conducted.

The next phase consisted of a series of consultations with roughly 60 EU experts from research, industry, policy, and RTD management that was kicked off with interviews on the OS vision, metrics of OS uptake and impact, and the involved main players and surfaced good practices. Based on this work, six distinct future OS scenarios were created to provide the necessary level of concreteness for the development of a first set of OS uptake and impact indicators that were scrutinised through a wider online consultation. Finally, a focus group served to validate the results and explore concrete policy options.

AN INITIAL SET OF NEW INDICATORS

The application of the above-mentioned methodology yielded, among others, a first set of new possible indicators for measuring the uptake and impact of OS. We could observe mainly two major dimensions – one pertains to the scientific process itself, i.e. the way science is conducted; the other pertains to the system level and thus the framework conditions. Each of these two major dimensions has several sub-dimensions:

1. the scientific process:
   - conceptualisation and data gathering/creation
   - analysis
   - diffusion of results
   - review and evaluation

2. the system level:
   - reputation system, recognition of contributions, trust
   - open science skills and awareness
   - science with society

These sub-dimensions are not exhaustive; they merely pose a categorisation that aligns well with the identified, new potential OS indicators. It goes without saying that this categorisation will need to be revised and refined the further the indicators are being developed.

Each of the above-mentioned dimensions entails a cluster of indicators. Those will be presented below in terms of their nature, their relevance, and the stakeholder group responsible for adopting and further developing an indicator. This article will not cover the entirety of indicators elaborated by the project team but only a subset of those sub-dimensions that are most relevant for the theme of the Open Evaluation conference. That said, the other sub-dimensions will at least provide a rough description to provide context and make it easier to understand the scope of the cluster.

Comparing these results with the ones generated by RAND Europe (Smith et al. 2016), there are similarities in terms of indicators that pertain to the scientific process, like open access publications (e.g. percentage of publications from each year that are open access, rate of green open access publications compared to journal publications, number of preprints, or journal policies on open access), open research data (e.g. number of data repositories, or funder policies on data sharing), and open scholarly communication (e.g. percentage of peer reviews that are published, journal policies on open peer review, use of altmetrics platforms/number of mentions of publications in media and social media, or articles published before peer review). Their work offers little with regard to the system level, though, which is the biggest difference compared to our work. Our consultation has shown that the necessary framework conditions need to be in place to foster an open culture.

Figure 1: Stakeholder groups - abbreviations and colour

| R  | researchers |
| RO | research (conducting) organisations |
| RFO | Research-funding organisations |
| PM | policy-makers |
| PU | publishers |

The presented indicators contain the stakeholder group that is – not solely but – mainly responsible for further developing and adapting an indicator. In some instances, more than one stakeholder group is responsible, i.e. when an indicator is fairly complex to design, maintain, or yield data. In any case, these stakeholder groups are defined as follows:

Each presented indicator will also have a mean rating that pertains to the consulted experts’ view on the relevance of said indicators – a 10 means the highest relevance, 0 no relevance at all; we have eliminated all of the roughly 60 indicators that did not achieve an above-average rating of at least 7.5.
Scientific work must no longer be restricted to measuring final products (such as articles), but should measure the development of the individual steps of the scientific workflow. Furthermore, results will differ according to disciplines, fields, or data types. Indicators in this dimension cover e.g. research funding organisations requiring the open provision of data/code, the accessibility of data/code, or the availability of metadata.

### INDICATORS CLUSTER I: CONCEPTUALISATION & DATA GATHERING/CREATION

Important questions in this dimension are whether the quality of data and information is adequate, e.g. whether the data were properly cleaned, whether they are curated, whether metadata are provided, etc. Recent policy trends involve mandatory rules and requirements (most commonly, funding agencies mandate public access to funded research), and the development of infrastructure to enable OS. Fewer initiatives relate to non-monetary incentive mechanisms like the definition of new reward/promotion systems.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Mean Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of research funders that mandate the provision of the data / software code produced in the context of the funded activity AND who mandate the conformity to data (exchange) standards</td>
<td>7.9</td>
</tr>
<tr>
<td>Accessibility</td>
<td>9.1</td>
</tr>
<tr>
<td>Accessibility of open data / code as % of all data / code produced by publicly (co-)funded projects</td>
<td></td>
</tr>
<tr>
<td>Machine-readable</td>
<td>7.5</td>
</tr>
<tr>
<td>% of machine-readable data / metadata</td>
<td></td>
</tr>
<tr>
<td>Availability of metadata</td>
<td>7.5</td>
</tr>
<tr>
<td>Availability of explanatory metadata as % of all available data (resulting from publicly (co-)funded research)</td>
<td></td>
</tr>
<tr>
<td>Quality of metadata</td>
<td>8.2</td>
</tr>
<tr>
<td>Quality of metadata (versioning, volume, data format, description of fields, etc.)</td>
<td></td>
</tr>
<tr>
<td>Simulation results</td>
<td>7.5</td>
</tr>
<tr>
<td>Usability of simulation results (models, data, and code)</td>
<td></td>
</tr>
<tr>
<td>Data services</td>
<td>8</td>
</tr>
<tr>
<td>(types of) open data services offered</td>
<td></td>
</tr>
<tr>
<td>Data compilation/publication costs incorporated</td>
<td>7.6</td>
</tr>
<tr>
<td>% of funded projects incorporating costs for data compilation / publication and maintenance (of the repository/data sets)</td>
<td></td>
</tr>
<tr>
<td>Long-term availability</td>
<td>8.2</td>
</tr>
<tr>
<td>Is the (long-term) availability of the data guaranteed (availability of a sustainability plan (yes/no))</td>
<td></td>
</tr>
<tr>
<td>Sharing policies</td>
<td>7.6</td>
</tr>
<tr>
<td># of sharing policies in research organisations (sharing of data, organisms, etc.)</td>
<td></td>
</tr>
</tbody>
</table>
limited verification of scientific results (cf. OECD 2015). Open peer review is often mentioned as an alternative, but not without the same amount of criticism. In the Open Science community, however, there is certain agreement that transparency measures need to be taken in the review and evaluation process. A multitude of suggestions have been put forward, some of which are considered as “incremental”, meaning that they would not do much harm to the current review procedure, while others are regarded “radical” or transformative. Adding transparency to the review process can happen at various stages. One option would be to make grant proposals publicly accessible at various points of time, e.g. after the project has ended, along with the final project reports, at the beginning of a project, at the point of announcing funding decisions, upon submission to the funder and during the drafting phase (cf. Mietchen 2014). Another would be to make the peer review public. This can again happen in an incremental form, meaning that some knowledge within the peer review process is made openly accessible, or in a radical form, meaning that transparency of knowledge becomes a separate pillar of legitimacy itself (cf. Gurwitz, Milanesi, and Koenig 2014). Open peer review is currently a highly contested field and so is the choice of respective indicators. This can also be said for the question how societal relevance of research should be treated and assessed in evaluation. A rather easy measure could be to make the “impact statement” of a proposal publicly accessible. A labelling system for expected impact (oriented on e.g. the Sustainable Development Goals) could be an option to create clearer evaluation references. Again, there are several options to develop new indicators but only a few concrete ones passed the threshold or were further suggested.

INDICATORS CLUSTER II: ANALYSIS

Respondents in this cluster argue that open methods contribute to improving the reliability of research results but that the impact of the open methods were still marginal because their use is not spread widely yet in the research community. Indicators in this cluster that are easier to design and monitor are data citations and code/software citations, a possible new one might be content citations.

INDICATORS CLUSTER III: DIFFUSION

We deliberately chose the term “diffusion” (of results) instead of the term “publication” which is most commonly used in academia. We want to stress that diffusion can and – some would argue – should start well before the results are out. In our online assessment, several comments underpinned the need to get away from the traditional paper publishing models and find indicators that gauge the growth of dissemination channels other than journals. Participants stated that journals are becoming irrelevant in many fields already. Impact of OS can more easily be captured in those cases where open communication and responsive attitude to feedback have actually changed the trajectory of research, e.g. a side-line turned into the main thing, a bug/design issue was detected, or the project just responded (or even emerged in response) to what is happening in society.

INDICATORS CLUSTER IV: REVIEW AND EVALUATION

Currently, peer review is the standard practice to assure quality of scientific output. Traditional peer review has well known shortcomings, though, such as little credit given to reviewers, lack of transparency and limited verification of scientific results (cf. OECD 2015). Open peer review is often mentioned as an alternative, but not without the same amount of criticism. In the Open Science community, however, there is certain agreement that transparency measures need to be taken in the review and evaluation process. A multitude of suggestions have been put forward, some of which are considered as “incremental”, meaning that they would not do much harm to the current review procedure, while others are regarded “radical” or transformative. Adding transparency to the review process can happen at various stages. One option would be to make grant proposals publicly accessible at various points of time, e.g. after the project has ended, along with the final project reports, at the beginning of a project, at the point of announcing funding decisions, upon submission to the funder and during the drafting phase (cf. Mietchen 2014). Another would be to make the peer review public. This can again happen in an incremental form, meaning that some knowledge within the peer review process is made openly accessible, or in a radical form, meaning that transparency of knowledge becomes a separate pillar of legitimacy itself (cf. Gurwitz, Milanesi, and Koenig 2014). Open peer review is currently a highly contested field and so is the choice of respective indicators. This can also be said for the question how societal relevance of research should be treated and assessed in evaluation. A rather easy measure could be to make the “impact statement” of a proposal publicly accessible. A labelling system for expected impact (oriented on e.g. the Sustainable Development Goals) could be an option to create clearer evaluation references. Again, there are several options to develop new indicators but only a few concrete ones passed the threshold or were further suggested.

INDICATORS CLUSTER V: REPUTATION SYSTEM, RECOGNITION OF CONTRIBUTIONS, TRUST

The uptake of OS practice in the research process is unlikely to flourish if researchers fear it is not properly acknowledged and officially recognised. This is underpinned in the initially mentioned surveys on researchers’ attitudes towards OS, which reveal low factual progress in putting OS into practice. Reward mechanisms for data sharing are currently especially weak and researchers might choose rather not to spend a serious amount of time in cleaning and curating their data for the reuse of others. Some organisations (datacite, ORCID, Figshare, Dryad Digital Repository, ResearcherID) have propositions for data citation tools which would credit authors for data and metadata sharing, but “in most countries the existing framework does not promote sharing efforts, especially with respect to results, data sets or other research material at the pre-publishing phase” (OECD 2015a, p. 89). Formal recognition of a variety of contributions along the scientific process (e.g. to the selection of research topics, formulation of hypotheses, project participations, review activities, etc.) has yet to be adopted. To understand the importance of the recognition of contributions, it serves to recall the various roles that are involved in the scientific process (see figure below).

Platforms that may provide data on data citation: DataCite, ORCID, Figshare, The Dryad Digital Repository, ResearcherID.
INDICATORS CLUSTER VII: SCIENCE WITH SOCIETY

This cluster is about finding indicators that assess effects of OS on the promotion of the engagement of citizens in science and research. As Mietchen, Mounce, and Penev (2015) observed, most of the research process is hidden from public view through multiple layers of obfuscation that stems from inherited conventions and habits from the paper era. This has begun to change, though, not least because digital technologies enable engagement and popularisation. Populatisation activities are understood as targeting a wide audience and a non-specialised public. Consequently, relevant new indicators gauge, among others, citizens’ engagement in (open) science, research communication (beyond academia), or the accessibility of data that are of public interest.

Although the importance of this sub-dimension has been recognised, only one of the suggested indicators was rated high enough to reach the predefined threshold.

Further options to explore are for example the % of publications in Open Access Journals (with or without impact factor) or availability of means to easily publish negative results.

INDICATORS CLUSTER VI: OS SKILLS & AWARENESS

OS-related skill development across disciplines will be a crucial factor for the maturation of OS in Europe. Researcher’s skills in OS (e.g. curating and maintaining large data sets) differ across disciplines due to different traditions or training opportunities in digital tools and data handling. There is a substantial need for further training of researchers and scientists in handling big, multi-layered and complex data sets. Accordingly, indicators in this cluster cover e.g. the monitoring of skilled personnel, research personnel active in OS, or the awareness and use of open standards.
CONCLUSION AND OUTLOOK

Designing indicators to measure the uptake and impact of OS (Open Science) is a challenge, not least because the concept itself is still evolving. OS is necessarily broad because it is composed of many dimensions (e.g. along the scientific research process) and embedded in a larger system that involves e.g. new skills, a new reputation scheme, or the wider public.

Most indicators proposed in this article are new and not gathered/surveyed/evaluated automatically (yet). Consequently, a first vital step is to put the necessary mechanisms in place. To achieve this, we propose stakeholder groups that are primarily involved in/responsible for designing, measuring, interpreting, and/or adapting an indicator.

It should be of prime concern to avoid the early mistakes of bibliometrics that had severe unintended negative consequences on the research system. An essential precondition to circumnavigate Campbell’s law and to make indicators work as intended is that all concerned stakeholder groups are involved in their design and evolution. They all need to agree on what an indicator should measure (and what it should not) and how it should be used (and what it must not be used for). Furthermore, indicators need to be flexible enough to accommodate differences, e.g. in research fields, and allow the emergence of new developments. The differences in research fields can be considerable, as is the pace at which OS is being adopted in those fields. Those differences will need to be elaborated and reflected in the respective indicators.

Furthermore, all stakeholders need to make sure that the OS indicators are and remain a means to an end and never become an end in themselves; otherwise, Campbell’s law would apply again.

Finally, new indicators need to be tested—not just discussed—before being adopted on a larger scale. This can be done in small experiments by using individual, selected indicators.

BIBLIOGRAPHIC REFERENCES

Liz Allen et al. (2014): Credit where credit is due.


Bartling, Sönke, and Sascha Friesike (2014b): Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing. Springer Open.


Mietchen, D. (2014): The Transformative Nature of Transparency in Research Funding. Published: December 30, DOI: 10.1371/journal.pbio.1002027


AUTHORS

corresponding author:

DIETMAR LAMPERT
ZSI — Centre for Social Innovation, Austria
Linke Wienzeile 246, 1150 Vienna
T: +43 1 4950 442
E: lampert@zsi.at

MARTINA LINDORFER
ZSI — Centre for Social Innovation, Austria

ERICH PREM
eutema GmbH, Austrian

JÖRG IRRAN
eutema GmbH, Austrian

FERMÍN SERRANO SANZ
Universidad de Zaragoza, Spain