A framework to monitor open science trends in the EU

Elta Smith (eltas@rand.org), Salil Gunashekar (sgunashe@rand.org), Catherine Lichten, Sarah Parks, Joanna Chataway

RAND Europe, Cambridge, United Kingdom

Abstract

Open science refers to ongoing changes in the way research is conducted, with a move towards increased transparency, collaboration, communication and participation. Open science involves changes for scientists themselves, such as open access to scientific papers and open research data, as well as increased understanding of and participation in science by citizens. These changes are expected to enable greater transparency, collaboration and research integrity in the short term and improve scientific quality in the long term. Open science is one of three priority areas for European research, science and innovation policy. A challenge for policymaking in this emerging area, however, is a lack of evidence regarding open science. In this context, RAND Europe and our partners—Deloitte, Observatoire des Sciences et des Technologies, Altmetric and Digital Science—are developing a monitor on behalf of the European Commission, Directorate-General for Research and Innovation. The monitor will track open science trends in Europe and identify the main drivers, incentives and constraints on its evolution. RAND Europe will also develop a website to host the monitor that will provide policymakers and stakeholders with access to data on open science.

Introduction

In the past two centuries, a picture has emerged of scientists toiling away in labs and offices. Trading amongst themselves in a currency of highly specialised publications, they remain obscured from the general public until they emerge to announce a major breakthrough in medicine or technology…

This description, though obviously exaggerated, highlights several of the characteristics of scientific research that are being challenged through initiatives and activities that together constitute what is increasingly labelled as 'open science'. Broadly, open science has come to represent a set of changes taking place in how scientists interact and collaborate with one another, the engagement of the public in science, and expectations about the imperative to share results – particularly those obtained through publicly-funded research. Linking all of these changes is an evolution in our understanding of the role of science in society that, some would argue, is bringing back values that have been overshadowed in modern science, such as a spirit of exploration[1] and an appreciation for sharing knowledge.[2]

More specifically, doing 'open science' relies not only on being more transparent about the research process, its outputs and outcomes, but also specifically engaging in particular kinds of activities, such as making research results available through open access platforms, publishing the data (and

[1] Lichten, C.A., M. Hafner and S. Wooding (2014) Venture Research: Fostering trust and freedom in research funding. Santa Monica, CA: RAND Corporation, http://www.rand.org/pubs/research_reports/RR506.

[2] Könneker and Lugger (2013) '*Public Science 2.0—Back to the Future*' Science 4 October: Vol. 342 no. 6154 pp. 49-50.

metadata) supporting publications, and even making software, code, labbooks and research protocols freely available – usually facilitated by digital technologies and particularly the Internet.

Aside from the availability of digital technology, which enables many of these developments, two main drivers for these open science-related changes are: a strong belief in the value of free circulation and criticism of ideas, and a re-appreciation of the role of data by researchers.[3] Researchers are perceived to be seeking new ways to disseminate their outputs and collaborate, becoming increasing critical of existing peer-review systems, and finding new sources for data. Specific developments include researchers sharing protocols and tips online through sites like OpenWetWare.org; crowd-sourced computer game players have been named as authors on scientific publications,[4] while the programme GalaxyZoo sources contributions from online visitors classifying images of galaxies.[5] Researchers are also identifying new sources of data, for instance mining Twitter for information on public health and epidemiology,[6] and on political trends.[7]

The implications for science are significant and important to understand. Potential benefits include science becoming more inclusive, democratic and relevant to society.[8] At its best, open science supporters believe that it could help break down barriers between disciplines, and between science and the public, and speed up the scientific process by tapping into the critical mass necessary to generate ideas and facilitate falsification of theories.[9]

Open science is developing of its own accord, and will likely continue to do so, whether policy measures are implemented or not. But the potential benefits of open science connect strongly with European research and innovation priorities related to the free flow of knowledge and inclusivity. Policies could be put in place whilst open science is evolving, to encourage developments to occur in a coordinated fashion, nationally, regionally and internationally. This requires a forward looking view, anticipating and preparing for the future, rather than reacting to developments as they occur, and exactly this kind of approach has recently been encouraged by the European Commission's Science and Technology Advisory Council.[10] Within the EU, and aligned with the move from the Lisbon Strategy to Europe2020, there has also been a general push for increased monitoring, aiming for both increased transparency to enable each country to see what others are doing, and timely

[3] EC (2015), p. 8.

[4] Cooper, S., F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović. (2010) "Predicting protein structures with a multiplayer online game." Nature 466, no. 7307: 756-760.
Eiben, C.B., J.B. Siegel, J.B. Bale, S. Cooper, F. Khatib, B.W. Shen, F. Players, B.L. Stoddard, Z. Popovic, D. Baker (2012) 'Increased Diels-Alderase activity through backbone remodeling guided by Foldit players,' Nature biotechnology 30, no. 2: 190-192.

[5] Zooniverse (2016) 'GalaxyZoo', www.galaxyzoo.org

[6] Brownstein, J. S., C.C. Freifeld, & L.C. Madoff (2009) 'Digital disease detection—harnessing the Web for public health surveillance,' New England Journal of Medicine, 360(21), 2153-2157.

[7] Larsson, A.O., & H. Moe (2012) "Studying political microblogging: Twitter users in the 2010 Swedish election campaign." New Media & Society, 14(5), 729-747.

[8] Kroes, N. (2014) 'The Transformative power of digital science.' As of 26th of March: http://ec.europa.eu/programmes/horizon2020/en/news/neelie-kroes-transformative-power-open-digital-science

[9] Popper, K. (1963) Conjectures and refutation. London: Routledge.

[10] 'The Future of Europe is Science,' a 2014 report published by the Commission's Science and Technology Advisory Council

delivery by having clear and measurable targets that are regularly surveyed.[11] Monitoring the emerging trends of open science can help towards achieving these aims.

As such, we are supporting the European Commission in developing an open science monitor, which focuses on trends in Europe, but also assesses how open science is evolving in other parts of the world.[12] It will be accessible to the public through an online platform and will present as far as possible trends by country and scientific discipline. The monitor will serve as a pilot to test the potential for analysing open science trends as the movement evolves over time and will also be used as a tool for the Commission's Open Science Policy Platform to develop policies to support open science where appropriate.[13]

Conceptual approach

When we began developing the monitor, there was no overarching monitor for open science already in existence, although several EU-level science and technology monitors do assess particular aspects of open science. These include the European Innovation Scoreboard (formerly Innovation Union Scoreboard),[14] ERA progress reports,[15] the Digital Scoreboard,[16] and Responsible Research and Innovation trends.[17] The OpenAIRE platform is an EU-funded infrastructure programme specifically focused on monitoring open access data for research publications.[18]

The success of these approaches has varied, with the Innovation Union Scorecard being considered amongst the most successful. OpenAIRE is widely viewed as an excellent source of information on open access data. The ERA progress reports, on the other hand, have proven to be more challenging to produce, in part due to the use of surveys which have garnered few responses, making comprehensive EU analysis difficult. One learning point from these systems is that good indicator choice is crucial for designing successful monitoring systems.

Jaffe (1999) identified a good indicator as ideally being as precise as possible and unbiased. [19] Furthermore, the relationship between the underlying concept and indicator should be stable over time, comparable, not susceptible to manipulation and subject to aggregation. It is generally not possible to meet all of these criteria as there are trade-offs in choosing indicators which fulfil the

[11] What is the difference between Europe 2020 and its predecessor the Lisbon Strategy? http://ec.europa.eu/europe2020/services/faqs/index_en.htm

[12] RAND Europe (2016) 'Monitoring Open Science Trends in Europe',

http://www.rand.org/randeurope/research/projects/open-science-monitor.html

[13] European Commission (2016) 'Research & Innovation, Open Science: Europena Open Science Policy Platform', http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform

[14] European Commission (2016) 'Growth: Internal Market, Industry, Entrepreneurship and SMEs, European Innovation Scoreboard', http://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_en

[15] European Commission (2014) 'European Research Area, ERA Progress: ERA Progress Report 2014', http://ec.europa.eu/research/era/eraprogress_en.htm

[16] European Commission (2016) 'Digital Single Market, Digital Economy & Society: Digital Scoreboard', https://ec.europa.eu/digital-single-market/en/digital-scoreboard

[17] ResAGorA RRI Trends (2016) 'What is RRITrends?', http://www.rritrends.res-agora.eu/

[18] Rettburg, N and B Schmidt (2015) 'OpenAIRE: Supporting a European open access mandate' *College and Research Libraries News*, Association of College and Research Libraries, http://crln.acrl.org/content/76/6/306.full.

[19] Jaffe, Adam B. (1999) 'Measuring knowledge in the health sector,' *OECD/NSF high-level forum, NSF, Washington, DC*. Vol. 264.

most important objectives of the monitoring system, but it is important to consider all of these aspects in the choice of indicators.

In general, therefore, the monitor we are building takes as a first principle that successful monitoring of open science requires that the indicators used need to be easily linked to the trend being assessed so that they can be readily understood, and ideally built from well-established data sources.

Characteristics of open science

Open science can be regarded as a systemic change in the way research is being conducted, affecting steps throughout the research process, from upstream components like idea generation, planning and design, through to the outputs and impacts of research, further downstream. We are attempting to monitor the diversity of the different components making up this broader open science system. We initially undertook a comprehensive literature review to identify core defining characteristics of open science. These were the presented in an online consultation to invited stakeholders for further validation.[20] Open access publications, open research data, and 'open' communication activities were ranked most important by the majority of respondents. This helped us to prioritise and focus on three main characteristics to feature in the monitor.

Based on further feedback, we decided that another important characteristic in the open science ecosystem--citizen science or civic engagement--will be included in our assessment of open science trends, but will not be monitored in the same way as the other three characteristics. In a way, citizen science is part of both the supply and the demand side of open science, and for this reason should be analysed separately. It relates more to 'who' is taking part in the scientific process. Moreover, 'measuring' citizen science is much more difficult than for other aspects of open science for reasons that we describe below.

Open access research publications

Among the open science trends, arguably the most fully developed from a policy perspective is open access to research outputs – particularly publications. Open access relates to the accessibility of the research approach undertaken as well as to the outputs of the research. In the preceding Seventh Framework Programme for research and innovation (FP7), which ran over the period 2007-2013, the Commission enabled open access to FP7-funded research outputs with the reasoning that doing so helps optimise the impact of publicly-funded research in the EU.[21]

Horizon 2020, the EU research and innovation framework programme for 2014-2020 – in the spirit of the EU's strategy to boost international co-operation in research and innovation – promotes 'borderless research' that is open to researchers from all over the world, and openness in the sharing of results. It is envisaged that all resulting publications will become openly and freely available, and researchers are also encouraged to make available the underlying data.[22]

[20] The consultation involved individuals working in areas related to open science in the EU and third countries, and primarily included the following two groups: (a) policy-makers in the European Commission, and from the Member States and outside the EU, as well as funders, librarians and publishers; and (b) researchers, and representatives of academies of science, open data and open access platforms, crowdfunding and crowdsourcing platforms, and citizen science platforms.

[21] European Commission (2014) 'Policy Initiatives'

[22] European Commission (2014) 'Horizon 2020 in brief. The EU Framework Programme for Research & Innovation'.

A study by the European Commission (2013) found that in the EU-28 Member States, 45% of the papers published from 2008-2011 were open access. [23] Related studies found that among 48 major science funders, the majority considered both 'gold' and 'green' forms of open access acceptable forms of open access publication.[24,25]

Open access was the characteristic ranked highest amongst open science characteristics in an online consultation we conducted with policy-makers and researchers and was also the primary focus of discussions at the 2016 Dutch Presidency Conference on open science. Open access is viewed as being integral and specific to open science: integral in that it is important for achieving a major goal of open science which is to accelerate scientific progress and impact for societal benefit, and specific because it is a necessary condition for open science. Open access allows both researchers and citizens to access research findings, extending the flow and use of knowledge within the 'academy' and beyond.

Open access is important to measure in part because this aspect of open science is the most developed, with established policies and momentum behind it to achieve targets for open access to become universal by 2020.[26] Articles also remain the primary research output, so measuring the degree of openness in this area provides insight into the extent to which open science is becoming a 'normal' part of research activities.

While this characteristic is important to monitor, we also believe that it should not be over-emphasised, as open science is much more than simply open access, despite the fact that open access tends to dominate discussions at present. Additionally, open access may be achieved in a relatively short period of time--potentially in fewer than five years–therefore it may quickly become 'obsolete' as an area to monitor. This consideration is built into the monitor as a form of 'future-proofing'.

Open research data

Open access to research data is less well developed than open access to publications, but it is viewed by stakeholders as a critical component of open science. Horizon 2020 therefore encourages researchers to make available their underlying research data.[27] The Commission also launched an Open Research Data Pilot within Horizon 2020, as a step towards increasing the practice of data sharing and reuse of data, and to support the realisation of the societal and economic benefits those

[23] Archambault, E., D. Amyot, P. Deschamps, A. Nicol, L. Rebout & G. Roberge (2013) 'Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011.' Brussels: ScienceMetrix for European Commission DG Research & Innovation. As of August 2014: http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf

[24] Caruso et al. (2013) 'Open Access Strategies in the European Research Area.' ScienceMetrix. http://www.science-metrix.com/pdf/SM_EC_OA_Policies.pdf

[25] In line with wider policy approaches in this area, the EC identifies two primary routes for open access publication: 'Gold' open access: scientific publishers provide open access to the article immediately upon publication (in this case, the publication costs are borne by the authors, their institution or the funding agency sponsoring the research. 'Green' open access or 'self-archiving': the researcher (or a representative of the researcher) archives the published article (or the final version of the peer-reviewed manuscript) in an online repository before, alongside or after the article's publication. European Commission (2013) 'Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, Version 1.0'

[26] The Netherlands EU Presidency (2016) Amsterdam Call for Action on Open Science. http://english.eu2016.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science.

[27] European Commission 2014, 'Horizon 2020 in brief. The EU Framework Programme for Research & Innovation'.

data could bring.[28] Under the pilot, researchers provide a draft Data Management Plan as part of their proposal, to be finalised in the first six months of the project. Although development of a Data Management Plan can be prepared voluntarily by all participants, it is mandatory in the Horizon 2020 Work Programme (2014-15) for a number of areas of Horizon 2020 that together account for about 20 per cent of the total budget for that period.[29]

Open research data relates both to the transparency of data collection and to enabling the replication of research results. Open data was ranked very highly in our online consultation and also discussed widely at the 2016 Dutch Presidency Conference. Like open access, this characteristic was considered to be integral to open science, although with more caveats. For example, compared to open access it is more relevant to specific disciplines in terms of the opportunities it provides (at least currently) and requires more meta-data to make it useful, such as guidance on where it came from and how it was used.

Furthermore, we would wish to distinguish between 'open data' and 'open *useful* data.' That is, simply posting data on a website, for example, with no description of it, or even putting sequences in a sequence database but not explaining how they were generated, is not particularly helpful from the perspective of evaluating or re-using data. Accessibility is important to open science (a 'first step'), but it was clear from the consultation that this quality is necessary but not sufficient; the ability to understand, use and reuse data is essential.

Communication activities related to open science

Communication activities are an important part of open science. Open science communication relates to the engagement of different stakeholders and activities at various stages of the research cycle and touches upon aspects of collaboration, accessibility and transparency. Communication is 'enabled' in some sense by open access and open research data: that is, as research data and findings become more open, there is more for people to communicate about.

Communication activities related to open science help to characterise the way in which 'conversations' about scientific research take shape. Unlike open access and to some extent open research data, communication activities related to open science are a more 'bottom-up' characteristic of the movement, driven largely by the research community rather than by policy. Communication spans discussion on social media (e.g. Twitter, ResearchGate, and Academia.edu), comments on published papers (e.g. PubPeer), discussion of preprints (e.g. bioRxiv), communication while projects are ongoing, and 'live' project or proposal work (e.g. the 'opening science' book[30]), or proposals for the open data prize.[31]

---

[28] Europa website 2013, 'Commission launches pilot to open up publicly funded research data'

[29] These include future and emerging technologies, research infrastructures (e-Infrastructures), leadership in enabling and industrial technologies (Information and Communication Technologies), several societal challenges, and science with and for society.

[30] Bartling, Sonke & Sascha Friesike eds. (2014) *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Berlin: SpringerOpen. http://www.openingscience.org/get-the-book/ (printed and dynamic books available).

[31] Martone M, Murray-Rust P, Molloy J, Arrow T, MacGillivray M, Kittel C, Kasberger S, Steel G, Oppenheim C, Ranganathan A, Tennant J, Udell J. (2016) 'ContentMine/Hypothes.is Proposal.' *Research Ideas and Outcomes* 2: e8424. doi: 10.3897/rio.2.e8424

The definition of what constitutes communication activities and collaboration was blurred in the online consultation. On reflection, it is difficult to draw the boundary between these two concepts. For example, platforms that enable more open communication help promote collaboration. And while there are aspects of communication that are more readily 'measureable' as being 'open', collaboration is harder to characterise in this way.

Citizen science and public engagement in research

While many aspects of open science are about new ways that researchers are interacting with one another and sharing information, another aspect relates to how a different set of research players – citizens – are taking a more active role in science. This involvement aligns with EU values for citizen engagement. For instance, 'The Future of Europe is Science' a 2014 report published by the Commission's Science and Technology Advisory Council, emphasises the importance of involving citizens in decision-making, not just 'small elites', and incorporating citizens' views in science and technology policies.[32]

While citizen involvement in science is both an aim and enabler of open science (as are the other three characteristics), there are many dimensions of citizen science to consider in the open science 'ecosystem'. The biggest issue is definitional: what exactly does 'citizen science' mean? It can refer to citizens 'doing science', for example, through crowdsourcing. Or it can mean greater understanding of science by the public by virtue of greater access to information about the research process (e.g. ability to use open research data or download open access articles) or the ability to understand science and engage with scientists (i.e. more 'open' communication through blogs, social media, priority setting for funding etc.). The public is also engaging in policy-making through, for example, agenda-setting for research systems.

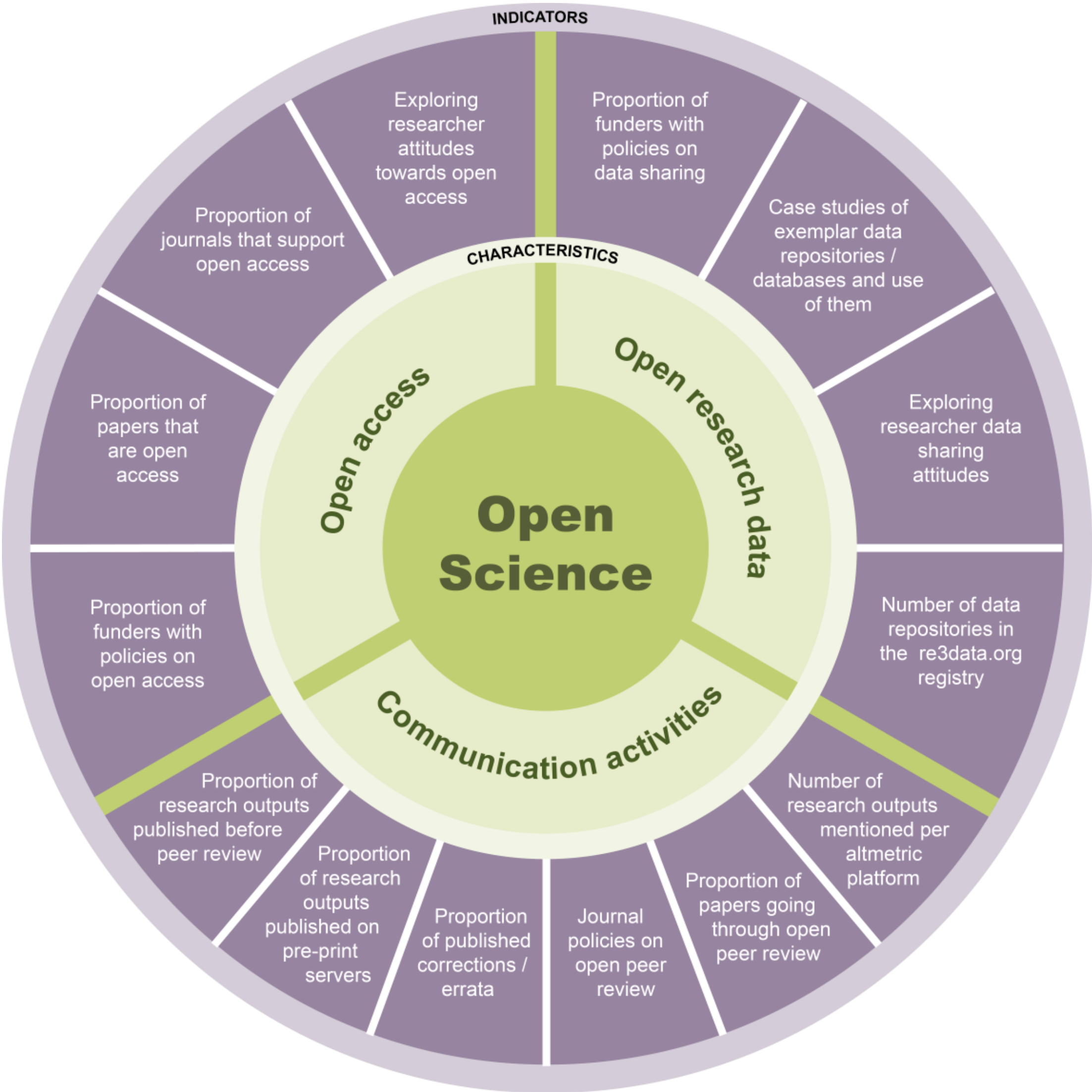
More importantly, there is no way to clearly distinguish between 'scientists' and 'citizens'. Considering its importance to open science and the issues described here, the study team believes that including citizen science in the monitor would by necessity limit the number of dimensions that can be presented, which is unduly restrictive and potentially problematic for long-term monitoring and understanding of the open science system. Therefore, we decided that citizen science should be presented more qualitatively than the other characteristics, with information highlighting developments across several of the dimensions discussed here.

Choosing indicators to monitor trends for each characteristic

The monitor we have designed will bring together these characteristics as a way to demonstrate how the open science 'system' is evolving in Europe. To do so, we include indicators for each characteristic that adhere to a set of core principles described below. Figure 1 illustrates how the proposed indicators (depicted in the outside ring) align with each of the three core open science characteristics (depicted in the inner ring).

[32] European Commission Science and Technology Advisory Council (2014) 'The Future of Europe is Science', http://ec.europa.eu/archives/commission_2010-2014/president/advisory-council/documents/the_future_of_europe_is_science_october_2014.pdf

Figure 1: Alignment of proposed indicators with core open science characteristics to be monitored during the project



***Source**: RAND Europe (2016)*

First, indicators must be normalised so that they can be compared and have meaning. Raw counts will generally not be useful unless they are placed on a scale which shows whether the count is relatively big or small in its context. For this reason, where possible we focus on proportions rather than counts. For example, in the context of open access, we are monitoring the proportion of papers that are available open access, rather than the absolute number of papers.

The indicators must also be contextualised in terms of the quality and importance of data. For some indicators, simple counts will not be sufficient and multiple ways of measuring openness will be required. For example, in the context of open research data, the issue is not about tracking how much data is open, but how much is usefully open. While it would be interesting to show information on the downloads and views of data from open data repositories such as Figshare and

Genbank, we think it would more helpful to demonstrate usefulness of such types of platforms on the monitor by complementing quantitative data with qualitative information through the use of exemplar case studies. The case studies could help answer questions related to how these data repositories and databases fit into the broader context of open science, their coverage and future plans within the evolving open science ecosystem.

As far as possible, we have proposed indicators that align with the three core open science characteristics in terms of the number of indicators per characteristic, the types of indicator and the levels of disaggregation. In general, there are roughly the same numbers of indicators for each characteristic. It is inevitable, however, that there is some overlap of indicators across characteristics. For example, to illustrate open science communication activities, we have included the proportion of research outputs published on pre-print servers as one of the indicators (as this relates to researchers disseminating their research online in an open and transparent way prior to publication). However, we also feel that this indicator is very relevant to open access, and would therefore aim to link it to both characteristics on the monitor.

In addition, we have proposed indicators across the three characteristics that cover a range of similar 'types'. Under open access and open research data, for example, we propose to show information derived from surveys that illustrate researcher attitudes towards open access and data sharing, respectively. We expect to show data related to funders with policies on open access and data sharing. Finally, where possible, we demonstrate open science trends across all three characteristics by disaggregating the indicators by subject/discipline, country and time.

We have selected indicators that provide insight into open science across the research process. That is, we look across the research process from inputs (e.g. funding policies) through to activities and outputs (e.g. publications) to develop a picture of how open science is evolving for each characteristic. Critically, however, at this stage there is no way assess any causal or even correlative aspects of the trends from inputs through to outputs due to a lack of established datasets across the different features of open science. It is still too early in the evolution of open science to demonstrate these kinds of connections; rather our aim is to show the way open science is taking shape across different stages of the research process and for different types of activity.

Transparency in monitoring

Despite their centrality to the system, open science is more than open access, open research data, associated communication activities and citizen science. These characteristics were chosen/ as the focus of our analysis for the first monitoring effort because they represent the core features of open science at present and they are the most well-developed and well-understood aspects of the system. For these reasons, the chosen characteristics also lend themselves more readily to measurement in a monitor, or in the case of citizen science, there is enough information available to present a compelling picture of its importance.

There are many other characteristics of open science that could be included in a monitor. These include: open code, open software and research infrastructure, open peer review and evaluation, open educational resources and open innovation. Each of these is to a large extent enabled by open access, open data and 'open' communication activities. But while all of them are important to consider as part of the wider open science ecosystem, they are not as prominently featured.

We also identified indicators that would in principle be better at monitoring open science trends than those that we ultimately chose for this monitoring exercise. Identifying appropriate indicators for the project required making trade-offs between indicators that were based on already existing data that were feasible to collect and relevant to the characteristics that were chosen for monitoring at this stage.

The aspects of open science that were not chosen for this pilot of the monitor should not be excluded entirely from the monitoring effort. Instead, we will highlight them on the website, including the reasons why they were not chosen for monitoring. We also propose that the website includes a tool to (effectively) crowdsource ideas for future monitoring approaches. This could include a feature that allows for comment on the importance of these other characteristics, which is likely to change over time, and therefore identify whether and when they should be monitored and how.

What are the potential impacts of open science?

There is seldom a clearly-defined path from underpinning research to outcomes and impacts, but impact is ultimately what most people wishing to understand the open science system would want to know about. Open science may be beneficial for many reasons including reducing duplication of research efforts and equally enabling the replication of research results where needed. It may also result in faster more timely research results, and ultimately greater overall efficiency of the research process.

Some of the most prominent examples of 'open science' in action have emerged from large-scale public health crises. For example, data sharing of genome analyses to tackle the Ebola epidemic was widely seen to have enabled geneticists and evolutionary virologists to work together to confirm the origin and transmission mode of the virus, as well as estimated routes of infection and predicted rates of mutation. This information supported crisis management efforts by local and international public health organisations through showing them where to focus their relief efforts and enabling them to develop practical advice to limit the spread of infections. Data sharing was also considered to be helpful for both the public and private sectors to more quickly design new therapies, diagnostics kits and vaccines. Similar efforts have been undertaken to support the Zika virus outbreak and to combat malaria.[33]

Evidence has emerged, at least anecdotally, that other benefits of open science are beginning to be realised. Scientists have used online forums and social media to rapidly debunk high-profile studies found to have flaws, such as the 2010 announcement of a discovery of arsenic-based life[34] and the more recent claim that views on same-sex marriage can be changed by a brief conversation with a gay canvasser.[35]

[33] For example: Ebola (www.eboladata.org) and Malaria (www.opensourcemalaria.org).

[34] Hayden, E. (2012) 'Study challenges existence of arsenic-based life'. Nature News. 20 January.

[35] Bohannon, J. (2015) 'Science retracts gay marriage paper without agreement of lead author LaCour'. ScienceInsider. 28 May.

The Joint Research Centre (JRC) completed reports reviewing ongoing citizen science and smart cities activities,[36] and citizen science and 'do-it-yourself' science projects.[37] The JRC found that citizen science projects go beyond data collection, producing impacts such as awareness-raising and community-building, and that it is important that such initiatives emphasise the importance of trust and of giving something back to the community as opposed to focusing solely on taking away information. Citizen science activities can be educational and provide a means of gathering valuable observations or computational contributions (in the case of online games). As citizens take more active roles in science, however, issues relating to bias and conflicts of interest arise that need to be addressed as they do for 'traditional' scientists.[38]

There are many indications of changes in the way science is being done, who is doing it, and what they are studying. These examples suggest that these changes may have widely beneficial effects on science and ultimately on society, although there will be challenges and potentially negative effects as well. But it is too early in the evolution of open science to measure impact in a comprehensive way for any of the trends we monitor in this project. At this stage, a qualitative, case study approach has been chosen to illuminate specific examples of open science achievements and also to explore potential negative impacts of open science as well as barriers to achieving the most desired impacts.

Open science policies

The potential benefits of open science connect strongly with global research and innovation priorities related to the free flow of knowledge and inclusivity. Although open science can be thought of as a grassroots movement, led by researchers themselves, as well as publishers, media platforms, industry, and citizens, it has been increasingly receiving serious attention from governments and other institutions worldwide. For example, open science is one of three main priority areas for the European Commission's science, research and innovation policy.[39] In the U.S., the White House Office of Science and Technology Policy developed policies to increase public access to federally-funded research results.[40]

In some areas, such as open access to research publications, open science policies have already been put in place at EU and Member State levels. Open data policies and infrastructure development are under discussion, while little policy development has occurred in other areas such as the use of alternative metrics, citizen science and the role of open science in scientists' careers.

The potential benefits of open science align with European Research Area (ERA) priorities. The overarching aim of the ERA is to create a 'unified research area open to the world, based on the Internal Market, in which researchers, scientific knowledge and technology circulate freely and through which the Union and its Member States strengthen their scientific and technological bases,

[36] Craglia M, Granell Canut C. (2014) 'Citizen Science and Smart Cities'. JRC Technical Report.

[37] Nascimento, Guimarães Pereira and Ghezzi (2014) 'From Citizen Science to Do It Yourself Science'. JRC Science and Policy report.

[38] Anon. (2015) 'Rise of the citizen scientist'. Nature News. Editorial. 18 August. http://www.nature.com/news/rise-of-the-citizen-scientist-1.18192

[39] European Commission (2015) 'Open Innovation, Open Science, Open to the World', http://europa.eu/rapid/press-release_SPEECH-15-5243_en.htm

[40] U.S. OSTP (ND) 'OSTP Public Access Policy Forum', https://www.whitehouse.gov/administration/eop/ostp/library/publicaccesspolicy

their competitiveness and their capacity to collectively address grand challenges.'[41] The 2012 Communication on the ERA envisaged significant progress in Europe's research performance, which would entail growth promotion and job creation.[42] The ERA would play a crucial role in opening up and connecting research systems across the EU, including by optimising access to, and circulation and transfer of scientific knowledge.

The most recent ERA progress report[43] identified ongoing areas for improvement related to open science. For instance, national roadmaps for developing research infrastructures (which have been adopted by 22 Member States) are not clearly aligned to the European Strategy Forum on Research Infrastructures roadmap. Alignment across Member States would contribute to linking up data infrastructures and promoting data sharing and collaboration across Europe. The progress report identifies a need for Member States to do more to work together on developing data infrastructures and repositories. Twenty Member States have taken measures to support open access to publications but only five have done so for open research data. A range of barriers exist, related to legal issues, trust, culture, and technical and financial challenges.

Open science is also largely seen as a grassroots phenomenon driven by researchers at present, through the actions of a range of actors – publishers and technology platforms, industrial actors, and, to a lesser extent, citizens – who all play roles in driving change. With that in mind, respondents to a recent consultation on Science 2.0[44] stressed that any policy developments should enable researchers to continue driving open science.

Monitoring to improve policy development

Respondents to the Science 2.0 consultation also made it clear that while researchers need to be rewarded for engaging with open science, there is a risk that over-reliance on open science-related metrics could have adverse consequences. This echoed considerable concern raised in the Science 2.0 consultation about whether and how researchers would be recognised and rewarded for doing open science, and that creating additional expectations for researchers related to open science would mean added burdens for researchers who already face very high workloads and stiff competition in career progression. Participants warned that over-reliance on new metrics for tracking open science-related activities can be risky, a view consistent with findings from a recent UK review that assessed the role of quantitative metrics in research evaluation.[45] Among other findings, the review concluded that inappropriate indicators can lead to unintended, undesirable consequences, and that metrics cannot replace expert qualitative assessment.

With that in mind, our monitor has used indicators that were improved upon and validated by policy-makers, researchers, funders, and others involved in open science activities through an online focus group consultation. The group representing policymakers, funders and libraries strongly

[41] European Commission (2012) 'A Reinforced European Research Area Partnership for Excellence and Growth'

[42] *Ibid.*

[43] European Commission (2014) 'Communication from the Commission to the Council and European Parliament: European Research Area Progress Report 2014'.

[44] RAND Europe (2014) 'Science 2.0: Science in Transition', http://scienceintransition.eu/

[45] Wilsdon, J., et al. (2015). 'The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management.' Hefce.

agreed (82%) that a monitor would be useful to them; similarly, 71% of researchers also felt a monitor would be beneficial. Amongst both groups, the most common benefit cited was the ability to keep up to date with progress in open science. Respondents also hoped the monitor would allow them to demonstrate that open science practices are increasing, potentially helping to encourage uptake by others, as well as to promote and show the utility of open science. The monitor could help stakeholders understand the needs of others in the community and what infrastructure or skills may be required. Both groups felt that for policymakers, the monitor could help in developing policy targets, better tailor policy instruments to needs in the area and identify policy gaps. Policymakers also specifically mentioned that the monitor would give them the information that would help them remain active stakeholders in Open Science discussions.

Conclusion

Open science is still in the early stages of its evolution and our attempt to monitor its progress is by necessity in its infancy as well. The monitor is intended to serve as a guide to policy development while open science is in the formative stages of development, providing the Commission and its Open Science Policy Platform with consolidated information on trends and practices to show where policy making may be beneficial, and where participants in open science initiatives have already been able to establish their own beneficial practices. Equally, open science may have negative effects on the scientific research system, and these should also be highlighted and explored. The monitor can help to identify problems with open science as well as its positive results. And we are designing the monitor so that it can be useful for a range of other stakeholders, particularly the researchers who are closely involved in ‘operationalising’ open science principles on a day-to-day basis.