

OECD SCIENCE, TECHNOLOGY AND INDUSTRY WORKING PAPERS

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to OECD Directorate for Science, Technology and Innovation, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France; e-mail: sti.contact@oecd.org.

The release of this working paper has been authorised by Andrew Wyckoff, OECD Director for Science, Technology and Innovation, and by Marcos Bontori, OECD Director for Public Governance. This working paper has been presented and discussed within the Global Science Forum and the Committee for Science and Technology Policy (CSTP). The Framework that is proposed in the paper was presented at a CSTP workshop on Open Science, Open Innovation and the Digitalisation of STI in Paris, October 2016 and a workshop on Open and Inclusive Innovation in Seoul, July 2017.

This publication is a contribution to the OECD Going Digital project, which aims to provide policymakers with the tools they need to help their economies and societies prosper in an increasingly digital and data-driven world.

For more information, visit www.oecd.org/going-digital.

#GoingDigital

© OECD/OCDE 2018

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to rights@oecd.org.

TABLE OF CONTENTS

OPEN AND INCLUSIVE COLLABORATION IN SCIENCE: A FRAMEWORK4

ABSTRACT4

1. OPEN SCIENCE MEANS DIFFERENT THINGS TO DIFFERENT PEOPLE.....5

2. THEORETICAL FOUNDATION AND DEVELOPMENT OF A COMMON CONCEPTUAL
FRAMEWORK OF OPEN SCIENCE7

 2.1. Open research agenda setting.....10

 2.2. Open funding mechanisms.....11

 2.3. Open access to publications12

 2.4. Open research data.....13

 2.5. Open government data14

 2.6. Citizen science14

 2.7. Crowd-sourcing.....15

 2.8. Open research infrastructures15

 2.9. Enabling e-infrastructures16

 2.10. Open science tools17

 2.11. Open peer review18

 2.12. Open licenses and IPR19

 2.13. Public engagement (open science communication)20

 2.14. Knowledge transfer.....20

 2.15. Science repositories21

 2.16. Metrics22

3. CROSS-CUTTING ISSUES23

4. A CAUTIONARY NOTE - RISKS AS WELL AS OPPORTUNITIES24

BIBLIOGRAPHY26

OPEN AND INCLUSIVE COLLABORATION IN SCIENCE: A FRAMEWORK

*Qian Dai**, *Eunjung Shin*[†], and *Carthage Smith**

(*) OECD Directorate for Science, Technology and Innovation (STI), Global Science Forum, Science and Technology Policy Division (STP);

([†]) Science and Technology Policy Institute (STEPI), Korea.

ABSTRACT

Open science can be variously defined. In some communities it is related principally to open access to scientific publications, for others it includes open access to research data and for others still it includes opening up the processes of academic research to engage all interested civil society stakeholders. The absence of a common understanding of what is, and isn't, included in open science creates confusion in discussions across these different communities. It is potentially holding back efforts to develop effective policies for promoting open science at the international level. This paper builds on the limited conceptual work that has been published to date and proposes a broad framework for open science. The framework is not meant to be prescriptive but should help different communities and policy makers to decide on their own priorities within the open science space and to better visualise how these priorities link to different stage of the scientific process and to different actors. Such a framework can be useful also in considering how best to incentivise and measure various aspects of open science. Digitalisation is fundamentally changing science and the paper lays out some of the opportunities, risks and major policy challenges associated with these changes.

1. OPEN SCIENCE MEANS DIFFERENT THINGS TO DIFFERENT PEOPLE

Open science in its broadest sense refers to efforts to make the scientific process more open and inclusive for all relevant actors, within and beyond the scientific community, as enabled by digitalisation. However, there are very significant differences across scientific disciplines, across countries and across different societal stakeholder groups as to what open science means for them and where the focus of policy attention should accordingly be directed.

Most scientists would agree that open access to scientific publications is a desirable aim, although even here there can be reservations from those parts of the science community that tend to publish mainly in books rather than via journal articles. When it comes to what open access publication means in practice and the acceptability of preprint publication, open peer review practices and the use of social media there are big differences between different scientific disciplines. Likewise, with regards to open data - whilst most scientists might agree in principle to the idea of having easier access to the underpinning data associated with scientific publications, many would also argue that their own field of research needs special consideration. Molecular biologists might publish their genomic sequence data immediately, but physicists involved in expensive long-term experiments, will make the case that an embargo period is necessary before their primary data is shared beyond a limited circle of collaborators (and in any case who would want the massive amounts of unprocessed data that their experiments generate?)

When it comes to really opening up science and making it inclusive of other societal stakeholders, the division within the scientific community is even starker. Social scientists and clinical researchers may be used to co-designing and co-producing their research with citizens and patients, chemists may be used to sharing their knowledge in close partnerships with industry, but many fields of science remain wary of any distortion of the academic scientific process by outside interests. The notion that only scientific peers are qualified to judge what areas of research merit funding and what the priorities in these areas should be is indeed still alive.

From the policy-makers perspective, there are many drivers for open science and these do not always align with the perspectives of the scientific community. Whilst promoting scientific excellence and reproducibility (via open data) are widely embraced, stimulating innovation and business sector engagement may be less enthusiastically taken up in some quarters and arguments about democratic engagement in different stages of the scientific process are treated with considerable scepticism by a large part of the science community.

Across different countries the enthusiasm for open and more inclusive science varies considerably reflecting political, societal and scientific structures, cultures and history as discussed in the report Making Open Science a Reality (OECD, 2015a). Whilst many Northern European and Latin American countries have adopted a broad definition of open science and developed national open science strategies that fully integrate innovation and public engagement, other countries are more reticent. In some the discussion on open science is largely restricted to widening access to science publications. In others it may include research data but with varying degrees emphasis on openness. And the connection between open access publications, open data, innovation and public engagement is frequently not made at the national strategic and policy level.

All of these different perspectives are legitimate and valid and deserve to be respected. The aim of the rest of this paper is not to defend or condemn any of them. Rather, the aim is to start from a broad definition of open science and develop a holistic framework in which different actors can identify their potential roles and that policy-makers can use, either as a whole or partially, to identify the links and overlaps between different aspects of open science. The framework is a tool to assist all the different communities with an interest in open science and help to structure the necessary discourse between these various stakeholders.

As already indicated in this brief introduction, open science is a complex and multi-layered concept and there is a danger in trying to capture all of this complexity in detail that one creates confusion rather than clarity. Hence the framework is necessarily a simplification of reality and, although it is a logical extension of previously published conceptual work, some choices have had to be made in terms of what topics to include and in defining their relationship to different actors and different stages of the research process. These choices can be justifiably criticised. The framework makes no pretention at being definitive and the reader is invited to interpret and use it accordingly.

2. THEORETICAL FOUNDATION AND DEVELOPMENT OF A COMMON CONCEPTUAL FRAMEWORK OF OPEN SCIENCE

The term 'Open Science' was first coined by the economist Paul David (2003) in an attempt to describe the properties of scientific goods generated by the public sector (OECD, 2015a). Since then, the term has proven to be a dynamic concept that continues to evolve. For such an important notion it is surprising that the conceptual thinking, as identified in the recent literature, is actually quite limited and there appears to be no universal understanding as to what open science, as commonly used today, does and does not encompass.

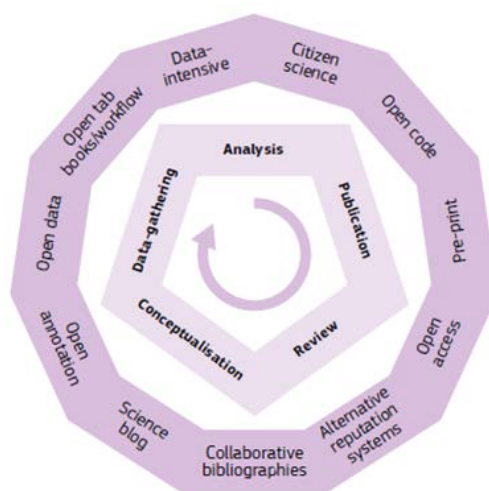
In a recent publication, Fecher and Friesike identified five schools of thought: the infrastructure school, the public school, the measurement school, the democratic school, and the pragmatic school (Fecher and Friesike, 2014) each with its own assumptions, goals and key focuses. Among various topics that are considered central to the different schools, open access to science publications has received the most attention and open research data is also viewed as a fundamental aspect of open science. In the report, *Making Open Science a Reality* (OECD, 2015a), the OECD refers to open science as efforts to make the output of publicly funded research more widely accessible in digital format to the scientific community, the business sector, or society more generally (OECD, 2015a). As such, open science was strongly advocated in the Daejeon Ministerial Declaration on Science, Technology, and Innovation Policies for the Global and Digital Age (OECD, 2015b).

So, many commentators and policy makers agree that open science is more than just open research data and open access to publications for the benefit of the scientific community. The European Commission (EC), has defined open science as “a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools” (EC, 2016a). This broader and more holistic framing (Figure 1) is echoed in some national policy documents. In a recent French white paper open science is defined as “a new horizontal approach to access to scientific work and objectives, and to sharing of scientific results, as well as a new way of doing science, by opening up its processes, codes and methods” (CNRS, 2016). Finland defines open science and research as “efforts to promote open procedures in scientific research activities” pointing out that achieving openness requires an open approach at all stages of the research process (The Open Science and Research Initiative (ATT), 2014).

If we follow the arguments that open science is about opening up the scientific process, the immediate questions are open to whom and how? The OECD 2015 report includes discussion of the key actors and enablers for open science beyond the traditional science enterprise. Indeed, to develop a holistic framework for open science, we need to emphasize the importance of all the actors in open science and digitalisation as the game changer. In so doing, we can more confidently define open science as “the efforts to make the scientific process more open and inclusive to all relevant actors, as enabled by digitalisation”.

The challenge then is to develop a holistic framework for open science based on its 3 basic elements: the scientific process, key actors and digitalisation. Within this framework, we can position specific topics that contribute to, or are affected by, open science, including not only open research data and open access to publications, but also topics such as research agenda-setting and crowd-funding mechanisms, etc. Collecting these topics together in a single framework, helps us to explore interactions, policy targets and the roles of various actors in a more coherent way. Such a framework can also be useful in considering how best to incentivise and measure various aspects of open science.

Figure 1. Open science and the scientific process



Source: European Commission, 2016a.

So let us further consider the 3 basic framework elements...

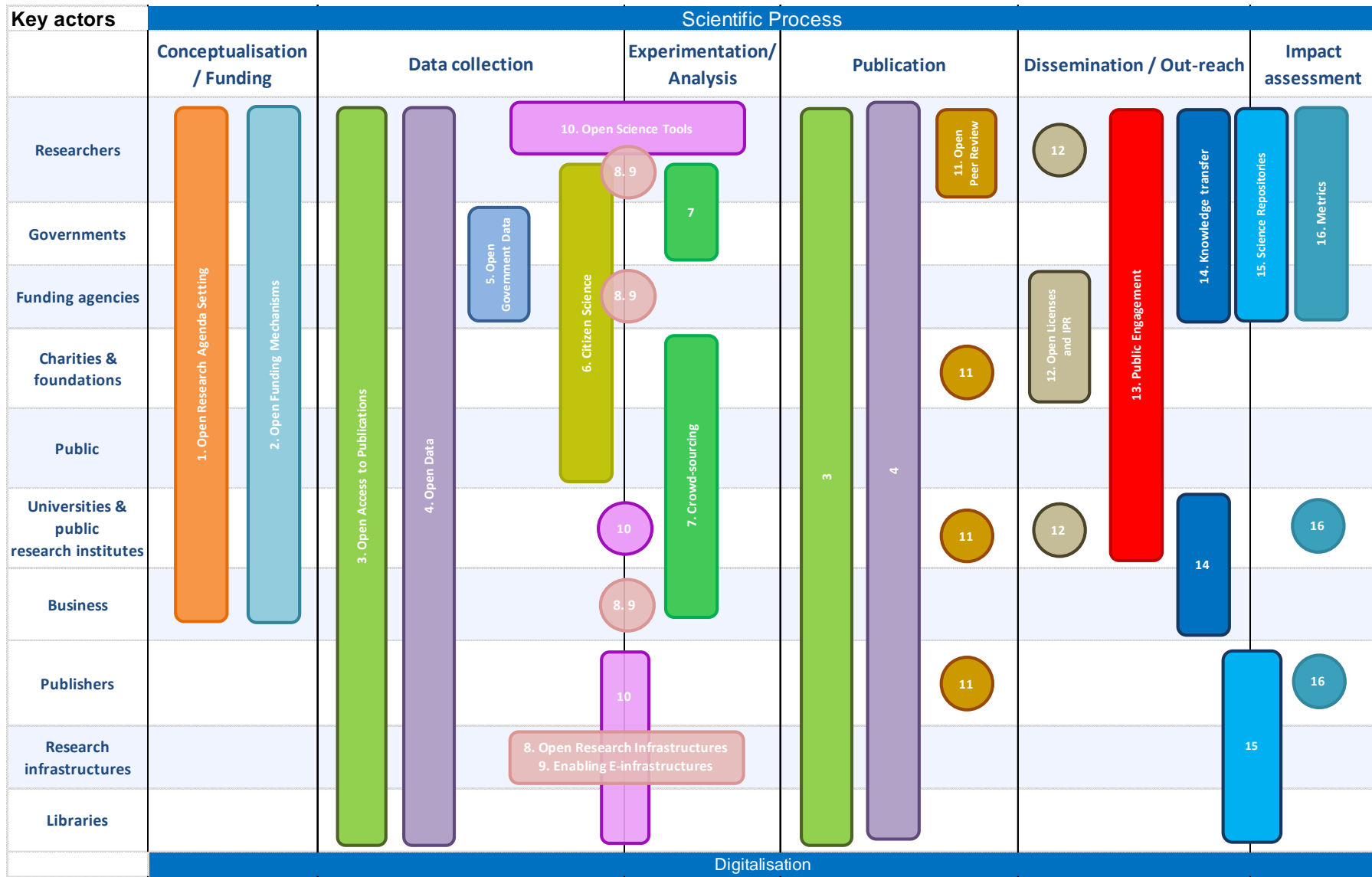
Element 1: the scientific process, includes the key steps in the research life cycle. With reference to Fig 1, we include in the scientific process: the procedures of conceptualisation and funding, data collection, experimentation and analysis, publication, dissemination and out-reach, and the additional step of impact assessment. Open science emphasises aspects that might not appear in traditional definitions of research conduct, especially the initial process of conceptualisation (including planning and funding) and the extended end processes of outreach and impact assessment. In many ways these are the areas, at the interface between science and society and business, where digitalisation has the most potential to radically alter the scientific enterprise, although thus far they are the areas that have received the least attention in relation to open science.

Element 2: the key actors can be extrapolated from those identified in the 2015 OECD report and includes: researchers, governments, funding agencies, charities and foundations, the public, universities and public research institutes, business, publishers, research infrastructures and libraries. While the public is considered mainly in the context of citizen science in this previous OECD work, it is recognised that the public is a growing actor in multiple aspects of the (open) scientific process. Research infrastructures also merit inclusion as specific actors as they play a vital role in data generation and storage, and themselves can be a key focus for open science in terms of providing access to critical facilities within and beyond national science communities (e.g. European Charter for Access to Research Infrastructures).

Element 3: Digitalisation can be broadly defined as the use of digital technologies to change a process, business or domain or in this case, to change science. The development and almost ubiquitous adoption of new information and communication technologies (ICTs) is transforming the scientific process. Digitalisation enables the scientific process to be more open and inclusive for all actors. As ICTs continue to evolve, the opportunities for exploiting open science will evolve in parallel.

Figure 2 proposes a preliminary framework that positions specific open science topics in relation to the scientific process and key actors (elements 1 and 2) with digitalisation (element 3) represented as a critical constant variable (although in practice ICTs are continuously evolving).

Figure2. Open science framework



This framework can certainly be expanded upon but even in this rudimentary form it helps to illustrate the complexity of the concept of open science and the involvement of different actors in different aspects. Whilst policy action to date has been largely focussed on open access to publications and, to a slightly lesser degree, open research data, it is important to understand that for many stakeholders open science is much more than these two topics and extends beyond the traditional scientific community. Consideration of how digitalisation can make scientific processes more open and inclusive to all relevant actors whilst maintaining or improving the quality of scientific outputs needs to take account of the complexity of different interests and objectives.

What follows is a more detailed review of each of the specific topics that have been placed within the proposed Open science framework that is presented in figure 2. (The principle actors involved in each of these areas are indicated in the framework.)

2.1. Open research agenda setting

Defining research priorities and setting science agendas, almost always involves funding agencies and researchers as the leading actors who historically have set their own priorities (Great Britain: Parliament: House of Lords: Science and Technology Committee, 2010; OECD, 2003). These processes have been criticised as having insufficient public engagement (Turney, 2011) although there has been a growing trend of involvement of a diverse range of actors (OECD, 2003), especially in research sectors such as health, environment and urban planning (Mitton et al., 2009). Responses to the 2016 OECD Science, Technology and Innovation Policy (STIP) questionnaire indicate that the involvement of multiple actors is being promoted in a variety of national science initiatives.

Health research is often cited as a good example of public involvement in research priority and agenda setting. A study for public involvement in Health Technology Assessment in the UK found that public input at the agenda setting stage often influenced the research plans; this influence included making patient and carer perspectives explicit, changing the focus of the research, and adding new outcomes (Oliver, Armes and Gyte, 2009). This positive impact of public engagement is echoed in many other studies (Mitton et al., 2009). Different methodologies, including surveys, citizen juries (Gooberman-Hill, Horwood and Calnan, 2008) and institutionalized associations, have been employed. And experiences (AMRC, 2009), toolkits (e.g. the G-I-N Public toolkit), process models and quality criteria checklists are openly available for consultation (Thiel and Stolk, 2013). In one example, th

e James Lind Alliance (JLA) has developed a formal process, involving multiple iterative steps, through which patients and clinicians identify their shared research priorities (Chalmers et al., 2013). More generally, the health sector could serve as a good reference for studying open research agenda setting more systematically (Mitton et al., 2009).

Governing bodies for science are recognizing the benefits in consulting multiple actors and engaging them in agenda setting at national levels. For instance, Korea has started a government-funded research programme in which the public first proposes research questions to be answered and then scientists develop research proposals to address these questions. The Dutch National Research agenda is being developed on the basis of 11,700 questions submitted by the public (De Graaf, Rinnooy Kan and Molenaar, 2017). Both of these public consultation exercises have been enabled by digitalisation and there is an increasing use of e-platforms for engaging a variety of actors.

Though efforts have been made to classify different formats and mechanisms of open agenda setting (Rowe and Frewer, 2005; Mitton et al., 2009; Engage2020, 2017), it is a challenge to develop an overall framework for such processes and identify good practices or consistent approaches for setting priorities (Great Britain: Parliament: House of Lords: Science and Technology Committee, 2010). Formal evaluation

of such processes remains also a key challenge, due to the absence of any optimal benchmark against which they might be compared and measured (Rowe and Frewer, 2004; Abelson and Gauvin, 2006). As a result, it is also hard to see whether a more open input to agenda setting differs systematically from that of experts, organisations or charities (Oliver, Armes and Gyte, 2009). There is also the need to develop skills and capacities for effectively implementing such processes (e.g. the capability to interpret and summarise views accurately, as well as to manage expectations of how these agendas and priorities will be addressed). Multi-stakeholder consultation processes can be expensive even when they are facilitated with on-line tools.

From the policy perspective there are a number of issues that merit consideration in relation to the inclusion of citizens in research agenda setting processes. These include defining the overall emphasis and priority to put on co-design in different areas of science under the broader umbrella of balancing fundamental with more applied, solutions oriented research. Whilst digital tools can make consultation and co-design processes more efficient there are also significant resource implications if co-design is to be done effectively

2.2. Open funding mechanisms

Governments provide the largest share of the funding for universities and other public research institutions. This funding is allocated mainly in forms of block funding and project funding, the latter usually involving peer review evaluation. There is also a very substantial amount of R&D financed and performed by business (OECD, 2003) although this is mainly focussed on applied research and experimental development. More recently, and enabled with the development of web 2.0 and new generation ICT technologies, we have seen the emergence of a plethora of new funding mechanisms that involve new actors - the public, charities and foundations - in new kinds of projects. These include online crowdfunding platforms, novel philanthropic investments, inducement prizes and social payments (Eisfeld-Reschke, Herb and Wenzlaff, 2014; Osimo, Pujol and Porcu, 2015).

A recent EC report defines alternative funding mechanisms for research as competitive research-funding mechanisms that: are led by non-governmental organisations; or, set research priorities in an open way without strong identification of research priorities; or, select proposals through other means than peer review of the projects (Osimo, Pujol and Porcu, 2015). Common features include, direct involvement of citizens, use of flexible vehicles for funding niche projects and transparent decision-making processes (Eisfeld-Reschke, Herb and Wenzlaff, 2014). There have been a number of recent efforts to review the research landscape and assess the potentials of these funding mechanisms (Osimo, Pujol and Porcu, 2015; Jakimowicz and Osimo, 2016).

Crowd-funding is one of the most prevalent open funding mechanisms that can be used to support research, and is deeply linked to digitalisation. It involves an open call, normally through the Internet, for the provision of financial resources either in the form of donations or in exchange for the future product or some form of reward, to support initiatives for specific purposes (Belleflamme, Lambert and Schwienbacher, 2014). Crowd-funding is growing strongly across many different sectors and is expected to overtake traditional venture capital investment in 2016 (Jakimowicz and Osimo, 2016). With respect to support for science, success has been enjoyed by scientific projects on general crowd-funding sites such as Kickstarter and IndieGoGo. In one well-documented example, a project called iCancer on IndieGoGo raised in excess of 2 Million USD for cancer research (Cullina, Morgan and Conboy, 2014). There are also dedicated crowd-funding websites for scientific research, including [the OSSP](#), [MyProjects-Cancer Research UK](#), [StartNext](#), and [Experiment.com](#). The way these platform work so far is largely similar to the way the more generic crowd-funding platforms operate, although there are discussions as to how to tailor them to the specificities of science, as well as to make the project processes and results more open and accessible to the funding donors.

Besides crowd-funding, innovation tournaments or inducement prizes are another mechanism frequently employed by government agencies, philanthropic bodies or business to promote challenge-oriented research and development. Examples include the DARPA Grand Challenge for autonomous cars, the Horizon prizes of the European Commission, the K-Crowd project in Korea, and the Google Lunar XPrize. Such prizes are designed to involve non-traditional actors, as well as to bridge the gap between research and application.

These funding mechanisms are considered as complementary to the traditional funding mechanisms and help increase diversity and creativity. At the same time, they can present challenges for existing regulatory or normative frameworks, and come up against barriers related to money and data flow (Jakimowicz and Osimo, 2016). Management of risks is also a key challenge, as new mechanisms may face problems such as high rate of failure, platform closure, fraud, cyber-attack and lack of transparency (Kirby and Worner 2014). There is a need for a better understanding of the impact of these mechanisms on overall research systems. They are often focussed on applied research, may promote or exclude niche research and make new demands on scientists' communication skills (Cullina, Morgan and Conboy, 2014; Osimo, Pujol and Porcu, 2015).

From the policy perspective there are several issues that come to the fore when considering open and alternative funding mechanisms. The balance between quality assurance and risk taking is important in this context. The relative roles of the public sector versus philanthropy and other research donors and the relationship between different funders and research institutions is also area that is likely to become increasingly complex offering both new opportunities and challenges

2.3. Open access to publications

Among the various topics that fit under open science, open access to publications has received the most attention and is universally viewed as a fundamental aspect of open science. Over the past decade, international organisations such as OECD, UNESCO, and World Bank have published guidelines on scientific publication access. At the same time, national governments and public funding agencies have gradually adopted policies that promote open access to publications. Most recently, the Amsterdam Call for Action on Open Science sets 'full open access for all scientific publications' an important pan-European goal for 2020 (EU, 2016).

The issues relating to open access to science publications have been discussed extensively in the report - Making Open Science a Reality (OECD, 2015a). In summary there are currently 2 main approaches to providing access to science publications openly and free of charge at the point of delivery on-line: the Green route, which involves delaying open access for an initial period during which subscription only access is provided and the Gold route, in which the author pays for publication and open access is immediate. Hybrid models are also being tested and all of these different approaches have their advantages and inconveniences as well as their proponents and opponents.

Whilst open access to publications is strongly advocated by the majority of the scientific community, there are growing concerns about the increase in predatory 'pay to publish' on-line journals that have very little, or no, quality control processes (Butler, 2013). This is part of a more general challenge of information overload, whereby the number of scientific publications in any particular scientific field exceeds the capacity of individual scientists to read them. Digitalisation, and the ease of on-line publishing and dissemination, both accentuate this problem and potentially provide solutions. Initiatives such as sceinceopen.com enable scientists share their own selected collections of publications and openly comment on other publications. Journal publishers are developing algorithms to recommend articles to customers based on their previous 'on-line' search histories. However, the attraction of such tools will need to be

balanced against their potential for introducing selective bias into the process of scientific information exchange.

Most importantly there is a need for sustainable business models for new publishing and dissemination paradigms. The whole area of the dissemination of scientific information is rapidly evolving (see section 13 ahead) and the role of formal peer reviewed scientific publications is only one part of this dynamic landscape. It is important as new models and new actors find their place in this evolving landscape that the long term stewardship of the (past and future) scientific record is ensured.

From the policy perspective, there is a challenge to establish the appropriate mix(es) of mandates, funding mechanisms, digital infrastructure and incentives to promote and sustain open access publishing. Different public and private sector actors can play a role but ensuring the quality and sustainability of scientific information will require careful attention.

2.4. Open research data

The Amsterdam Call for Action on Open Science sets the ‘optimal reuse of research data’ as a second major pan-European goal for 2020 (EU, 2016). An overview of the challenges and opportunities for making research data openly available and re-useable is provided in the 2015 OECD report on Making Open Science a Reality. The report i) reviews the policy rationale behind open access and open research data; ii) discusses and presents evidence on the impacts of policies to promote open access and open data; iii) explores the legal barriers and solutions to enable greater access to research data; iv) provides a description of the key actors involved and their roles; and finally v) assesses progress in OECD and selected non-member countries based on a survey of recent policy trends. It proposes a number of areas for policy intervention to promote open research data: better incentive mechanisms, skills development, data quality assurance, legal and ethical frameworks, business models for data repositories and international coordination.

Encouraging the more open use of human subject data (or personal data) for research is a particular challenge that is complicated by both legal and ethical considerations. These issues are addressed in recent OECD-GSF work on the use of new forms of data for social sciences research (OECD, 2013b) and the ethical issues relating to the use of such data (OECD, 2016).

In many fields a major obstacle to making scientific data more open is the behaviour and attitude of scientists. Good data management or the development of high quality data sets that can be re-used are not incentivised or rewarded and career paths for data managers are unclear in an academic research system that is focussed on publications in scientific journals. OECD has done some work on the challenges for sharing clinical research data related to Alzheimer's disease and issues such as recognition, credit and reward for researchers were seen as more of an obstacle than technical or legal barriers. Of course, incentives can only be effective if there is somewhere to store the data and make it continuously available (see sections 8, 9 and 15 below).

As with scientific publications policy makers need to consider carefully how mandates and incentives - including funding - can be used to promote open data. This includes the development of appropriate measures and indicators to incentivise and monitor open data practices. New legal, ethical and governance frameworks may be required to encourage the more open use of human subject data (from and for both the public and private sector). New procedures for triaging data and frameworks for making decisions on what data should be maintained in the longer term and what can be discarded will also be necessary.

2.5. Open government data

Open Government Data (OGD) refers to government or public sector data (i.e. any “raw” data produced or commissioned by public sector) made available through open access regimes so that it can be freely used, re-used and distributed by anyone (OECD, 2015a). Digitalisation plays a pivotal role in the development of OGD. Digital infrastructure enables OGD by hosting and publishing content and also improves efficiency, quality and equality of data access (Ubaldi, 2013). The development of standards for data exchange and interoperability are an important part of the OGD agenda. A 2013 OECD study looked at the added-value of OGD for public governance (accountability, transparency, better government services and civil servants), society (citizen self-empowerment, social participation and engagement) and the economy (Ubaldi, 2013). Many initiatives and portals have recently been developed by national governments not only to improve accessibility to, and availability of, their data, but also to foster their re-use by the entire open government data ecosystem to create public. According to the 2017 OECD OURData Index Korea, France, Japan, the UK and Mexico are leading in this policy area (OECD, 2017b).

From a science perspective, OGD is important for two main reasons: 1. The data itself can be of enormous value for research, particularly in the social sciences and humanities as well as health and environmental sciences; 2. The legal and ethical frameworks that apply to OGD can often be directly applied to research data (in fact some research data qualifies as government data). Work on indicators and monitoring and impact assessment of OGD can also provide insights for equivalent issues relating to open science data.

A series of challenges have been identified for implementing OGD (OECD, 2015a), relating to policy, technical, economic, organisational, cultural and legal aspects. For instance, disclosure policies may limit data transparency and copyrights may result in lack of clarity of data ownership. Agreed common data standards are lacking, which creates barriers for data access and utilisation. Feasible business cases are needed taking into account the costs of data collection, processing and provision. A culture of OGD needs to be nurtured in both the government and across the entire OGD ecosystem of users, including researchers.

There are lessons to be learned at the policy level from OGD experiences that have implications for the open data for science. New policies to maximise the potential value of OGD for scientific research, e.g. in relation to complex societal challenges, may also be required

2.6. Citizen science

Citizen science is a broad term that encompasses many actors and spans a range of levels of engagement: from being better informed about science, to participating in the scientific process itself by observing, gathering or processing data (EC, 2016b). It can be more narrowly understood as people who are not professional scientists taking part in one or more aspects of science (UNEP, 2014) or as the involvement of the public in scientific research (Citizen Science Organisation, 2016). It involves collaborations between the public and researchers/institutes, but also engages governments and funding agencies. Digitalisation has enabled the emergence of online collaborative platforms and analytical tools that are central to many productive and successful examples of citizen science (OECD, 2015a; Franzonia, and Sauermann, 2013).

In practice, starting with the narrower definition, three main domains for citizen science can be identified. The largest is composed of research on biology, conservation and ecology, and utilizes citizen science mainly for collecting and classifying data, a second is in geographic information research where citizens participate in the collection of geographic data, and the third is in social sciences and epidemiology (Kullenberg and Kasperowski, 2016). There are also different types of citizen science that have been

classified as Action, Conservation, Investigation, Virtual, and Education, and which differ in goals and organisation (Wiggins and Crowston, 2011).

Although in absolute numbers, the amount of publications generated by citizen science remains small, digitalisation in the past decade has led to the creation of many online citizen science platforms with a growth in the number and variety of projects (Newman et al., 2012; Kullenberg and Kasperowski, 2016). Platforms such as Zooniverse, FoldIT and eBird are engaging thousands of citizens on a daily basis in rigorously designed scientific research projects. Citizen science is rapidly evolving from being an extension of traditional scientific methods, engaging volunteer citizens to collect data under the direction of scientists, to a powerful mechanism for co-designing experiments and co-producing scientific knowledge with large communities of interested actors. This process has been enabled by combination of digital tools including online platforms, mobile apps and social media.

There are a number of familiar challenges for public engagement in science, relating to targeting the required audience and promoting effective participation (Kaufman, 2014). In practice, citizen science also faces organisational challenges (governance: matching the right people, division of labour, project leadership), motivational challenges (sustainable involvement and support, incentive mechanisms), technical challenges (technology literacy for both the scientists and public participants), as well as quality and evaluation challenges (Franzonia, and Sauermannb, 2013; Zhao and Zhu, 2014). There is a need for better coordination among stakeholders, wider recognition of the value of data generated by citizen science, and stronger international coordination to aggregate and analyse this data (UNEP, 2014).

From the policy perspective defining where citizen science approaches might best be adopted in different contexts and how to enable this will require careful consideration. Exploiting the potential of citizen science, while maintaining quality and promoting trust in science, are ongoing issues for policy makers and the scientific community more broadly.

2.7. Crowd-sourcing

Crowd-sourcing is a term frequently used in combination with citizen science, or considered as a specific form of citizen science. Unlike the other types of citizen science, in which the public is engaged mainly in relation to data collection, crowd-sourcing is characterised by an online, distributed problem-solving and production model (Brabham, 2008), positioning it more towards the analysis end of the research process. The business sector is frequently involved in crowd-sourcing initiatives, which can also be incentivised with prizes (see section 2). For instance, InnoCentive allows business to raise questions and interact with more than 300 000 registered users, or “problem solvers”. Another example is Kaggle (www.kaggle.com), where private companies and research teams publish unsolved problems related to specific data sets, and people from all over the world compete to find the best solutions (OECD, 2015a).

There is an important distinction to be made between citizen science, which involves citizens working together with professional scientists, and 'do it yourself' science, which involves non-scientists in designing and conducting scientific experiments. Crowd-sourcing can be used to encourage either or both of these approaches. Digitalisation and open science enable both approaches.

Whilst many of the policy issues raised by citizen science as a whole, such as assuring quality and accreditation, are relevant to crowd sourcing, 'do it yourself' research, which is conducted outside the normal governance mechanisms for science, raises new questions related to safety, security and ethics?

2.8. Open research infrastructures

Research infrastructures (or shared facilities) play an important role in open science. They are not only essential for carrying out experiments and analysis, but many large facilities also generate huge

amounts of data and support open science through associated e-infrastructures. Since the data issue is largely covered in topics 3 and 15, the term open research infrastructures as applied here mainly refers to open access to large experimental facilities, e.g. synchrotrons, or shared material resources, e.g. biobanks. For these facilities, better access to non-academic actors, notably the business community, and to overseas users, are areas of increased emphasis in relation to open science.

In Europe, the ESFRI roadmap requires that infrastructures must apply an “Open Access” policy for basic research, i.e. be open to all interested researchers and select proposals in open competition on the basis of scientific excellence as judged by international peer review (ESFRI, 2011). In the 2016 European Charter for Access to Research Infrastructures, ‘access’ refers to the legitimate and authorised physical, remote and virtual admission to, interactions with and use of Research Infrastructures and to services offered by Research Infrastructures to users, who include academia, business, industry and public services. It identifies three different modes of access: the excellence-driven mode, the market-driven mode and the wide access mode (EC, 2016c).

Digitalisation is facilitating the open access to research infrastructures. For example, the People's Republic of China has built an open online platform which lists all publicly financed research infrastructure and facilities, as part of a national initiative to promote access beyond the host institutes to other public research institutes, universities, businesses and the public (State Council of China, 2014). In a similar but more focussed initiative in Korea, the Institute of Lighting Technology provides an online platform for LED research. These on-line initiatives may be complemented by targeted resources to support access by specific user groups. For example, in New South Wales, Australia, a system of TechVouchers is providing a direct subsidy to encourage use of infrastructure and participation from the broader innovation system (Australian Government, 2010)

A large diversity is found in policies that determine researchers’ abilities to access infrastructures. Even when it is claimed that access is entirely merit-based (i.e., does not depend on whether the proposing scientist is affiliated with the facility or comes from a country that funds it), there are various unspoken restrictions (OECD, 2008), influenced by considerations such as return of investments, national security, privacy and confidentiality, commercial sensitivity and intellectual property rights (Group of Senior Officials on global research infrastructures, 2015; EC, 2016c). The increasing emphasis on more open access, especially for business communities and overseas users, is posing new challenges for established access policies and practices.

Exploiting the full potential of research infrastructures as a catalyst for open science requires policy action. Under appropriate conditions, public research infrastructures can play an important role in promoting business engagement in open science and innovation.

2.9. Enabling e-infrastructures

Enabling e-infrastructure ensures that the scientific information and data that is available through science repositories (see 15 ahead) can be exchanged and analysed. It includes the high level services that are embedded within some of these repositories and it also includes super-computing facilities and distributed computing networks, e.g. the Open Science Grid (Fecher and Friesike, 2014). Large investments are being made across OECD and non-member countries into e-infrastructures, with the development of cloud-computing, big data analytics and virtual technologies. These e-infrastructures are being combined with Open Science Tools (see 10 ahead) and science repositories in online collaborative platforms, such as the J-Stage platform in Japan and the European Open Science Cloud initiative.

Apart from the challenges noted for open research infrastructures (section 8), the necessary link between e-infrastructures and data means that the costs of data curation, provision and storage, as well as

challenges to the international flow of data are particularly important issues for e-infrastructures. The OECD-GSF has recently joined with international partners to start 2 projects on data infrastructures for open science. The first of these is looking the international coordination of data networks, including issues such as governance, trans-national data flows and funding. The second project is exploring sustainable business models for data repositories (see section 15 below). However, the main focus of these studies is data access and there are a number of important policy issues relating to shared data services and data analysis facilities that are not being addressed.

Funding and business models for e-infrastructures are an important area for policy attention. Linked with this, are questions related to the organisation of e-infrastructures in relation to science repositories and shared access to computing power and analytical expertise. (see also section 15 ahead)

2.10. Open science tools

A number of digital tools for researchers are now available to support open science. These include open lab notebooks, collaborative writing tools, open workflows, collaborative bibliographies, etc. (Hampton et al., 2015) [see Table 1]. Shared tools for text and data mining are an important area also for future development (OECD, 2015a).

Beside tools that target individual aspects of the research workflow, there are also efforts to create platforms that cater for multiple needs in an integrated manner. Some address specific domains, for example, the Virtual Biodiversity Research and Access Network for Taxonomy (ViBRANT) spans networking, data collection, analysis and publishing. While some, such as the Open Science Framework provided by the non-profit Centre for Open Science (USA), are generic project-oriented platforms that are able to connect to various 3rd party services and tools.

A convergence is also observed between open science and open source (Willinsky, 2005), where the codes of software and tools are open for dissemination and adaptation, and open source hardware is used to build research equipment (Pearce, 2012). The development of 3D printing and new materials, combined with open source modelling open and design software, provides powerful tools for open science (Pearce, 2013).

One very practical challenge for the use of open science tools is that established researchers must be willing to invest time in learning and re-learning how to use the tools as they evolve (Hampton et al., 2015). Meanwhile, new researchers need to be prepared with necessary skills and data literacy. Also, the development of robust and reliable mechanisms for community-based quality control is considered important for managing projects through online tools (Ponte and Simon, 2011).

Both public and private sector actors have a role to play in providing the underpinning digital tools and platforms that are essential for open science. Policies to establish mutually beneficial public-private partnerships will be required. In this context a specific policy issue is the extent to which software tools, computer codes and algorithms, which may be essential to test and reproduce scientific results, should be made openly available.

Table 1. A list of online tools for open science processes

Concept	Names of tool or service
Open discussion	Twitter, Blogs, Open lab notebooks, GitHub comments
Technical support	StackOverflow
Hypotheses/Design	Data management planning tool
Date Life-cycle support	Data repositories: KNB, Dryad, GBIF; Open office, MySQL, OpenRefine, Morpho
Version Control	Git and GitHub
Visualisation of geospatial data	GRASS, QGIS
Workflow tools	Kepler, VisTrails
Reproducibility	R, Rstudio, Python, Pycharm, IPython notebook/Project Jupyter, Sweave, markdown, pandoc, knitr, Babel, Rmarkdown
Writing collaboration	Google docs, Etherpad, ShareLateX, WriteLaTeX, Authorea
Reference management	Zotero, Mendeley
Distribution of figures and talks	Figshare, Slideshare, Speakerdeck
Distribution of preprints	bioRxiv, arXiv, PeerJ Preprints
Pre-publication peer review	Peerage of Science, Axios Review
Publication	DOI for code, DOI for data, 'Green' open access, 'Gold' open access, Licenses
Discovery of published data	DataONE, re3data
Social networking	ResearchGate, Academia.edu, Facebook
Tracking research product impact	ORCID, ImpactStory, Altmetric
Informal discussions	Discussion groups, personal blog/website

Source: Hampton, 2015.

2.11. Open peer review

One definition of open peer review is that it is where one or both of the parties involved are aware of each other's identities (Tattersall, 2016), as opposed to the single-blind or double-blind peer reviews frequently used for publications. This form of (limited and controlled) open review has tended to be resisted by researchers (Ware, 2008).

With the development of digitalisation and open access, it is argued that open peer review should be re-defined as on-line reviewing by all interested members of the scientific community, whether this be carried out anonymously or not. It has been proposed that this could be done in a post-publication (Kriegeskorte, 2012) or multistage manner, and may be one way to address the dilemma between the desire for speedy publication and the need to assure scientific quality (Pöschl, 2012). Together with altmetrics (see 16 ahead), open peer review is considered to be a core part of open science evaluation. Online platforms, often linked with open access publication portals, are now available for open peer review. Examples include the Open Review on Research Gate, and the F1000Research of Faculty of 1000.

However, it is notable that the use of these platforms is currently minimal (Tattersall, 2016) and challenges remain as to how to provide incentives and proper guidance to reviewers. This involves not only the science community, but also funding agencies, libraries/repositories, and research institutes. It is

also a concern that as the number of choices available for open online comment grows, it may be difficult for commenters to maximize the impact of their comments (Noorden, 2014). The validity and quality of some formats of open peer review require further analysis, e.g. comments as post-publication review (Anderson, 2014).

Exploiting the potential of digital tools to improve peer review processes and ensure the quality of openly accessible scientific information is likely to require policy mandates and incentives. At the same time online platforms that integrate the inputs and outputs of science have the potential to better inform scientific planning and policy-making.

2.12. Open licenses and IPR

A number of mechanisms can be used to assert Intellectual Property Rights on scientific materials, methods, publications and data. These primarily include patents, copyright and trade-marks. Ownership in itself is not an obstacle to openness and sharing. On the contrary, the patent system was designed to make information on inventions open and being clear about where ownership lies can be an essential prerequisite to sharing. However, the way that something is licensed by its owner can have important implications for open science. Restrictive licences, by both public and private entities, can seriously inhibit open science.

An open licence is one that imposes very few restrictions on what users can do with the material or product that is being licensed (Open Data Institute, 2016). In relation to the digitalisation of science, open licensing can relate to content (if it is subject to copyright), open source software and hardware, and data.

Creative commons (CC) licenses are frequently adopted by open access publications. CC offers six basic model clauses, two of which are considered a free licence: CC BY and CC BY-SA (OECD, 2015a). Another popular open license for content is the GNU Free Documentation License (GFDL), which was designed for manuals, textbooks, other reference and instructional materials. It is also used by Wikipedia (coupled with CC BY-SA).

Data ownership and licensing is more complex, especially in terms of the extent to which data/a database is copyrightable. There is a need for caution in applying CC licenses, other than CCO, to data/databases (Open Data Commons, 2016). Besides CC, a set of licences has been created specifically for databases by the Open Data Commons. The UK has launched an Open Government Licence, which is useful for the reuse of government and other public sector data (Korn, N. and Oppenheim, 2011).

Open source licenses allow software to be freely used, modified, and shared. For open source hardware, seven licenses are suggested by the Open Source Hardware association, from general ones such as CC BY-SA to more specific ones such as CERN Open Hardware License and the TAPR Open Hardware License (Open Source Hardware Association, 2016).

At a higher level, there is an ongoing debate about the implication of a number of patents for open science, particularly where these patents cover 'foundational' research information or materials. Some research institutions, e.g. the Montreal Neurological Institute, are exploring a "no patent" and open use approach to all their scientific outputs, effectively disowning all their IPR (Owens, 2016). They see this as a fundamental pillar of open science - allowing all users to exploit the outcomes of scientific research that has been publically funded without any restrictions other than acknowledging attribution. Such institutional policies are currently rare but may gain more leverage as digitalisation and open science continue to challenge established norms and practices.

Establishing an optimal balance between IPR protection and openness is an ongoing policy challenge. For example, appropriate policies could help to promote the use of open licensing agreements on essential research material, computer code, data and information.

2.13. Public engagement (open science communication)

Public engagement at the early stages of the scientific process, i.e. agenda-setting, and in co-production or citizen science has been discussed earlier (see section 1 and 6). At the downstream end, digitalisation has also provided new possibilities for researchers to communicate about their research in a more interactive manner, using tools such as science blogs and social media.

Science blogging is viewed as having the potential to become a new model for science journalism and a powerful tool that can be used by academic institutions to disseminate scientific information and facilitate conversations about science (Kouper, 2010). Studies have demonstrated the ability of blogs to improve science communication (Sublet, Spring and Howard, 2011), involve unconventional actors and provide timely information in times of crisis or risk (Mahrt and Puschmann, 2014). Researchers are increasingly active in the sphere of social media, e.g. the Science journal produces a list of top 50 science stars based on Twitter followers. Fenda, a mobile app run by the Chinese popular science website Guokr.com, enables the public to ask questions to leading researchers, with the questioner paying a price that he/she considers worthy for a 1 minute response.

With open science communication there are also challenges, including how to make sense of the large amounts of available data and information, and how to increase non-scientist participation in a sphere where the majority of both authors and readers are professional scientists or future professional scientists (Kouper, 2010). Effective public engagement is not simply a translation of facts but more importantly a negotiation of meaning, which address an intended audience's values, interests, and worldviews (Nisbet, 2009). A change of the mind-sets of researchers and proper steering from policy makers are called for (Gerber, 2014), with assistance from appropriate indicators and incentives (Neresini and Bucchi, 2011).

At the level of education policy, scientific literacy is an important aspect of public engagement in science. Ensuring the quality and appropriateness of on-line science education and communication materials has implications for science policy. Effective public engagement by scientists, including the promotion of science literacy, needs to be incentivised, measured and rewarded.

2.14. Knowledge transfer

The majority of scientific research is carried out in the private sector and the relationship between public and private research and innovation has been extensively studied. However, digitalisation and open science are shifting the boundary conditions between academia and the private sector and opening up new possibilities for productive collaboration and innovation. By engaging business in open science, it is argued that the potential benefits from open access to publications and open data in developing new products will be realised more effectively (Jong and Slavovaba, 2014). It is notable that an increase in knowledge transfer from public research institutes to business, as well as patent donating from business to research institutes are identified among the emerging trends in open science (Friesike et al., 2015). The development of joint public-private partnerships for the delivery of open science-related services is also growing and being actively promoted by government in some countries, such as the USA and Finland (OECD, 2015a).

It is important to note, with reference to the proposed open science framework, that business involvement in open science can be distinguished from open innovation, with the former oriented more towards knowledge creation and basic science rather than product development and experimental or

applied research. And insight studies on the former are lacking (Friesike et al., 2015). It is reported that external sources of knowledge play a much smaller role than internal or market sources; generally, less than 10% of innovating firms rank them as “highly important” for innovation activities (OECD, 2013a). It is not yet clear how digitalisation is really influencing business’ involvement in open science other than access to public research results, and how this can translate to open innovation (Chesbrough, 2015).

From the policy perspective further work is required on how digitalisation is affecting public-private partnerships in basic science. The effective management and attribution of IPR in public-private open science partnerships is an important area for policy attention

2.15. Science repositories

In line with open science, data itself is increasingly viewed as an essential part of research infrastructure and one which carries benefits for society and the economy. Sustainable data infrastructure is required to curate rapidly increasing volumes of data, which are becoming more complex over time. With the volume and variety of data increasing exponentially, budgets for data stewardship will struggle to keep pace despite falling storage costs.

Data stewardship is performed by a wide variety of data repositories, national and international in scope, some of which may also manage scientific publications. Although many established national and international data repositories have reliable sources of income from research funders, these sources of income are generally inelastic and may be vulnerable (whether to short-termism, ill-considered re-prioritisation or attempts to pass responsibility to other budgets). Some data repositories are exploring means of diversifying their income streams to increase sustainability and the economics of such a service – including value proposition, cost recovery, stakeholders’ willingness to pay and overall business model – will need to be robust in order to be sustainable. There is a pressing need, therefore, to explore new and innovative income streams and to establish sustainable business models for science repositories (OECD, 2017a).

Individual data repositories often function together in networks, which may be disciplinary or domain focused and frequently cross national boundaries. Indeed greater international cooperation is often a stated policy aim for open science and this starts with more effective sharing of data and information. International networks of data repositories have a critical role to play in underpinning open science, linking with other enabling e-infrastructures (see section 9 above) and managing the scientific data that will be exploited with e-tools in integrated digital platforms, re the " Open Science Cloud". In the absence of effective and sustainable international coordination of data repositories, which requires attention to issues such as shared standards, effective network governance, human skills and finance, open science will be seriously impeded. The OECD-Global Science Forum has an ongoing project, working with international partners, such as the Research Data Alliance, to address some of these data network issues from a policy perspective.

In addition to data and scientific publications, there is increasing awareness that computer code can, in many cases, be an important output of science. Access to specific code can be essential for testing and reproducing scientific results and the further development of such code by secondary users, where permitted, can lead to new research applications. Some online repositories, such as runmycode.org, are now providing public access to the code and data that underlie research publications

The development of sustainable business models for long-term open provision of good quality data is an important area for policy attention. This relates to the relative roles of public and private actors in the stewardship of research data. Frameworks and incentives to ensure effective international cooperation of data infrastructures for open science are also required.

2.16. Metrics

The use of new online tools and open availability of relevant data and information can enable the development and use of new metrics to capture different types of scientific outputs and their impacts (OECD, 2015a). These new metrics include usage-based bibliometrics, and measures of other research outputs such as data, and software, which can be linked to individual scientists (Fenner, 2014). The development of such metrics is known as Altmetrics - the creation and study of new metrics based on the Social Web for analysing and informing scholarship (Priem et al. ,2010). The OECD 2015 report gives an introduction to altmetrics.

Article level bibliometrics and author level metrics are being actively developed. For example, the open access publisher PLOS provides article level metrics that include altmetrics, citations and usage-based metrics such as number of reviews and downloads. Author levels metrics measure the impact of individual authors. As researchers engage more and more in open science, unique digital identifiers, such as Open Researcher and Contributor IDs (ORCID) for individuals or Digital Object identifiers (DOIs) for data-sets, are enabling interoperability and tracking across different platforms (ORCID, 2012).

The application of altmetrics, many of which can be linked to bibliometrics, is increasing and many publishers, such as Elsevier and the Nature Publishing Group, are providing such measures. The information is aggregated and analysed by non-profit projects such as Impactstory (which is also open source) or by commercial services such altmetric.com and Plum Analytics. Altmetric.com now is tracking over 5 million research outcomes, while Impactstory monitors over 1 million.

Although altmetrics have evident potential benefits, it is also acknowledged that many of these alternative metrics require further investigation to clearly assess what kind of impact they are measuring and how complete and representative the data sources are (OECD, 2015a). Whilst considerable efforts have gone into developing altmetrics based on finer analysis of traditional scientific publications and to a lesser extent social media, there has been only limited progress in developing rigorous indicators and metrics for other aspects of open-science. Most importantly, the provision of well-structured, open data-sets, which are considered to be essential for open science, is not well captured in current metrics.

There is an urgent need to clarify the expectations of different stakeholders and explore how these can these be translated into useful indicators and measures that can be used for policy analysis. New metrics need to be developed to measure, monitor and incentivise open science, including open data practices.

3. CROSS-CUTTING ISSUES

There are a limited number of cross-cutting issues, including scientific skills, legal frameworks, international coordination, and the development of integrated digital platforms that apply across the scientific process as a whole, involve multiple actors and thus cannot be readily mapped onto a specific area of the proposed Open Science Framework (Fig 2). Whilst these issues can be considered in relation to each specific topic (1-16 above), from a policy perspective they also need to be considered as topics in their own right.

Skills are a key enabler to allow the actors to be properly involved in open science. Previous studies have identified a shortage or mismatch of data-related skills for science (OECD, 2015a). Considering the wide spectrum of open science, skills are not only needed for researchers to use digital tools, but also for the public, business, etc., to engage efficiently in open science. And the skills required for enabling open science for different actors need to be identified and strengthened as necessary. A brief view of national initiatives for skills training is given in the OECD 2015 report.

Issues such the flow and management of data and access to large research infrastructures, require **coordination at the international level**. International governmental organisations, such as the OECD, the EU, the World Bank and UNESCO, can play a critical role in promoting intergovernmental co-ordination at international level and in shaping the political agenda through the development of guidelines and principles around specific themes, such as access to publicly funded research data (OECD, 2015a).

Open science-friendly **legal and ethical frameworks** are another overarching measure that is required to promote open science. Currently the focus is mainly on intellectual property and privacy issues, some of which are discussed above in relation to specific open science topics, but legal and ethical considerations go beyond these two areas (OECD, 2015c). In the European context many of these issues are captured within the concept of Responsible Research and Innovation (RRI).

At the present stage, online **platforms** mainly target different stages of the scientific process, e.g. providing services for data analysis, for publishing, or for evaluation. As technologies, standards and protocols develop, an integrated on-line platform that provides services across the whole scientific process (Förstner et al., 2011; Sweeney and Crosas, 2013), or integrates various other platform services (Carp, 2014), can be envisioned. This raises important policy questions related to the ownership, control and access to such platforms.

4. A CAUTIONARY NOTE - RISKS AS WELL AS OPPORTUNITIES

Science is undergoing a paradigm shift with digitalisation affecting all stages of the scientific process. This is providing exciting new opportunities for advancing scientific understanding, engaging with society and promoting knowledge transfer and innovation. It can also help to bring different disciplines and actors from different nations together to address complex global challenges. At a more systemic level, open science can contribute to making the conduct of science more rigorous, more transparent and more accountable and could help address some of the concerns about current incentive and reward systems for individuals and institutions. An optimistic, but not unrealistic, view of Open Science is it will make science more efficient, more effective and more attractive.

However, there are also many challenges associated with the transition to Open Science and these need to be addressed if the promise is to be realised. Some of these challenges have been referred to above in relation to specific topics (1-16). At a more generic level there are also concerns that cut across many topics and relate to the evolution of the science system as a whole. Science has traditionally been relatively closed to non-scientists, with expert peer review being used to ensure quality and validate communications; science has been its own gate-keeper and prided itself on being self-policing. Whilst recent cases of scientific fraud and problems with reproducibility have raised serious questions about the internal governance processes of science, it is not immediately obvious that greater openness and broader access to scientific tools, information and data will immediately lead to improvement. The established norms of scientific conduct, including critical appraisal and objectivity, are not necessarily shared by all stakeholders with an interest in science. Particularly in areas where societal values, economic interests and science may be in conflict, it is important that the necessary dialogues are informed by science that is as rigorous and trustworthy as possible. Whilst openness and transparency are positive in that they promote re-examination and testing of published evidence, there is concern that in an 'open market-place' sloppy science or pseudo-science will displace good science. As recent events have demonstrated, the world wide web and social networks are efficient propagators of mis-information as well as honest information and do not necessarily discriminate between the two. This is well illustrated in the area of climate science, where genuine scientific debate around the uncertainties and limits of human-induced global warming has been confounded by the wide dissemination of pseudo-scientific information.

There are particular challenges related to Big Data and the use of complex algorithms and models to mine and analyse this data (Boyd and Crawford, 2012; O'Neil, 2016). Data-driven science has huge potential but the old adage that if you have enough data you can prove anything is not without foundation. As Big Data becomes ubiquitous across science there is a growing need for the careful application of new statistical techniques and a critical analysis of the strengths and limitations of mathematical models (Spiegelhalter, 2014). This requires that the algorithms that are used to analyse big data are themselves open and available to critical analysis and improvement.

As mentioned above in relation to on-line platforms, there is also a concern that in the transition towards Open Science critical parts of the scientific process become dependent on a small number of commercial companies. Whilst scientific publishers and dot.com companies have contributed hugely to the development and dissemination of science, most commentators would agree that it is not desirable for the public scientific enterprise to become overly dependent on individual firms that have effective monopolies on major aspects of the scientific process. Similarly, it is not desirable that the huge amounts of data that are held by dot.com companies are withheld from academic researchers nor that important areas of

research, e.g. in health and social sciences, become the unique provenance of these companies and a small number of privileged academic institutions. New regulatory frameworks, partnerships and governance arrangements will need to be developed to ensure that science and society truly benefit from the potential of big data and the opportunities presented by open science (OECD, 2016).

BIBLIOGRAPHY

- Abelson, J. and F.-P. Gauvin (2006), *Assessing the Impacts of Public Participation: Concepts, Evidence and Policy Implications*, Canadian Policy Research Networks Inc., Ottawa.
- Association of Medical Research Charities (AMRC) (2009), *Paths to patient and public involvement for medical research charities*, https://www.amrc.org.uk/sites/default/files/doc_lib/2009-10%20Natural%20ground.pdf.
- Anderson, K. (2014), “Stick to Your Ribs: The Problems With Calling Comments ‘Post-Publication Peer-Review’”, *The Scholarly Kitchen*, <https://scholarlykitchen.sspnet.org/2014/04/15/stick-to-your-ribs-the-problems-with-calling-comments-post-publication-peer-review/>.
- Australian Government. (2010), *Strategic Framework for Research Infrastructure Investment*.
- Belleflamme, P., T. Lambert. and A. Schwienbacher (2014), “Crowdfunding: Tapping the right crowd”, *Journal of Business Venturing*, Vol. 29/5, Elsevier, Amsterdam, pp. 585-609.
- Boyd, D. and K. Crawford. (2012), “Critical Questions for Big Data” *Information, Communication and Society*, Vol. 25/5, pp. 662-679, <https://doi.org/10.1080/1369118X.2012.678878>.
- Brabham, D. C. (2008), “Crowdsourcing as a Model for Problem Solving”, *The International Journal of Research into New Media Technologies*, Vol. 14/1, SAGE Publications, pp.75-90, <https://doi.org/10.1177/1354856507084420>.
- Butler, D. (2013), “Investigating journals: The dark side of publishing”, *Nature*, Vol 495, Springer Nature, pp. 433-435.
- Carp, J. (2014), “A web platform for streamlining scientific workflows”, *Opensource.com*, <https://opensource.com/life/14/6/center-open-science-framework>.
- Chalmers, I. et al. (2013). “Tackling treatment uncertainties together: the evolution of the James Lind Initiative, 2003–2013”, *Journal of the Royal Society of Medicine*, Vol. 106/12, Royal Society of Medicine, London, pp.482-491, <http://jrs.sagepub.com/content/106/12/482.long>.
- Chesbrough, H. (2015), *From Open Science to Open Innovation*, Science | Business Publishing 2015.
- Citizen Science Organisation. (1 August 2016), “Citizen Science”, Citizen Science, <http://citizenscience.org/>.
- CNRS (2016), *Open Science in a digital republic*, <http://www.cnrs.fr/dist/z-outils/documents/2016%2003%2024%20White%20Paper%20Open%20Science.pdf>.
- Cullina, E., L. Morgan. and K. Conboy (2014), *The Development of a Public/Private Model for the Crowdfunding and Crowdsourcing of Scientific Research Projects*.
- De Graaf, B., A. Rinnooy Kan and H. Molenaar (2017), *The Dutch National Research Agenda in Perspective*, Amsterdam University Press, Amsterdam.
- Eisfeld-Reschke, J., U. Herb and K. Wenzlaff (2 March 2014), “Research Funding in Open Science”, *Opening Science*, http://book.openingscience.org/vision/research_funding.html.
- Engage2020 (2017), Engage2020 website, <http://engage2020.eu/> (accessed 7 July 2017).
- ESFRI (2011), *Strategy Report on Research Infrastructures*, Publications Office of the European Union, Luxembourg.
- European Union (EU) (2016), *Amsterdam Call for Action on Open Science*.
- EC (2016a), *Open Innovation Open Science Open to the World*, European Commission, Brussels..
- EC (2016b), “Citizen science”, <https://ec.europa.eu/digital-single-market/en/citizen-science> (accessed 4 March 2016).
- EC (2016c), *European Charter for Access to Research Infrastructures*, European Commission, Brussels.
- Fecher, B. and S. Friesike (2 March 2014), “Open Science: One Term, Five Schools of Thought”, *Opening Science*, <http://book.openingscience.org/>.

- Fenner, M. (2014), “Altmetrics and Other Novel Measures for Scientific Impact”, in *Opening Science*, Springer Nature, https://link.springer.com/chapter/10.1007/978-3-319-00026-8_12.
- Förstner, K. U. et al. (2011), *Collaborative platforms for streamlining workflows in Open Science*, Proceedings of the 6th Open Knowledge Conference, Berlin.
- Franzonia, C. and H. Sauermann (2013). “Crowd science: The organisation of scientific research in open collaborative projects”, *Research Policy*, Vol. 43/1, Elsevier, Amsterdam, pp. 1-20.
- Friesike, S et al. (2015), “Opening science: towards an agenda of open science in academia and industry”, *The Journal of Technology Transfer*, Vol 40/4, Springer Nature, pp. 581-601, <https://link.springer.com/article/10.1007/s10961-014-9375-6>.
- Gerber, A. (2014), “Science Caught Flat-Footed:How Academia Struggles with Open Science Communication”, in ”, in *Opening Science*, Springer Nature, https://link.springer.com/chapter/10.1007%2F978-3-319-00026-8_4.
- Gooberman-Hill, R., J. Horwood and M. Calnan (2008), “Citizens' juries in planning research priorities: process, engagement and outcome”, *Health Expect*, Vol. 11/3, John Wiley & Sons Ltd, pp. 272-281.
- Great Britain: Parliament: House of Lords: Science and Technology Committee (2010), *Setting priorities for publicly funded Research, 3rd report of session 2009-10, Vol. II: Evidence*, The Stationery Office Limited, London.
- Group of Senior Officials on global research infrastructures (2015), *Group of Senior Officials on Global Research Infrastructures Progress Report 2015*.
- Hampton, S. E. et al. (2015), “The Tao of open science for ecology”, *Ecosphere*, Vol. 6/7, Ecological Society of America, Washington, D.C., pp. 1-13, <http://dx.doi.org/10.1890/ES14-00402.1>.
- Jakimowicz, K. and D. Osimo (2016), *The potential of crowdfunding for Research and Innovation*.
- Jong, S., and K. Slavovaba (2014). “When publications lead to products: The open science conundrum in new product development”, *Research Policy*, Elsevier, Amsterdam, Vol. 43/4, pp. 645-654.
- Kaufman, A. B. (2014), “Let’s Talk about Citizen Science: What Doesn’t Work”, *Animal Behavior and Cognition*, Vol. 1/4, Sciknow Publications Ltd., pp. 446-451, http://animalbehaviorandcognition.org/uploads/journals/4/02.Kaufman_Final.pdf.
- Kirby, E. and S. Worner (2014), “Crowd-funding: An Infant Industry Growing Fast”, *Staff Working Paper of the IOSCO Research Department*, No. 3/2014, International Organisation of Securities Commissions, Madrid.
- Korn, N. and C. Oppenheim (2011), *Licensing Open Data: A Practical Guide*, http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf.
- Kouper, I (2010), “Science blogs and public engagement with science: Practices, challenges, and opportunities”, *Journal of Science Communication*, <https://jcom.sissa.it/archive/09/01/Jcom0901%282010%29A02>.
- Kriegeskorte, N (2012), “Open evaluation: a vision for entirely transparent post-publication peer review and rating for science”, *Frontiers in Computational Neuroscience*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3473231/>.
- Kullenberg, C. and D. Kasperowski (2016), “What Is Citizen Science? – A Scientometric Meta-Analysis”, *PLoS ONE*, Vol. 11/1, PLOS, San Francisco, <https://doi.org/10.1371/journal.pone.0147152>.
- Mahrt, M., and C. Puschmann (2014), “Science blogging: an exploratory study of motives, styles, and audience reactions”, *Journal of Science Communication*, https://jcom.sissa.it/archive/13/03/JCOM_1303_2014_A05.
- Mitton, C. et al. (2009). “Public participation in health care priority setting: A scoping review”, *Health Policy*, Vol. 91/3, Elsevier, Amsterdam, pp. 219–228.
- Neresini, F. and M. Bucchi (2011). “Which indicators for the new public engagement activities? An exploratory study of European research institutions”, *Public Understanding of Science*, Vol. 20/1, SAGE Publications, pp. 64–79, <http://journals.sagepub.com/doi/pdf/10.1177/0963662510388363>.
- Newman, G. et al. (2012), “The future of citizen science: emerging technologies and shifting paradigms”, *Frontiers in Ecology and the Environment*, Vol. 10/6, John Wiley & Sons Ltd, pp. 298–304, <http://onlinelibrary.wiley.com/doi/10.1890/110294/abstract>.

- Nisbet, M. C. (2009), “Framing Science: A New Paradigm in Public Engagement”, in *New Agendas in Science Communication*, Taylor & Francis Publishers, New York.
- Noorden, R. V. (14 March 2014), “The new dilemma of online peer review: too many places to post?”, Nature newsblog, <http://blogs.nature.com/news/2014/03/the-new-dilemma-of-online-peer-review-too-many-places-to-post.html>.
- OECD (2017a), “Business Models for Sustainable Research Data Repositories”, *OECD Science, Technology and Industry Policy Papers*, No. 47, OECD Publishing, Paris, <http://www.oecd-ilibrary.org/docserver/download/302b12bb-en.pdf?expires=1517998291&id=id&accname=guest&checksum=F55562A3B03CF2ED3D4EC0F30B66554A>.
- OECD (2017b), *Government at a Glance 2017*, OECD Publishing, Paris, http://dx.doi.org/10.1787/gov_glance-2017-en.
- OECD (2016), “Research Ethics and New Forms of Data for Social and Economic Research”, *OECD Science, Technology and Industry Policy Papers*, No. 34, OECD Publishing, Paris, <http://www.oecd-ilibrary.org/docserver/download/5jln7vnpxs32-en.pdf?expires=1517998515&id=id&accname=guest&checksum=61144EB2487868BE8102BEE9E5AA9F71>.
- OECD (2015a), “Making Open Science a Reality”, *Science, Technology and Industry Policy Papers*, No. 25, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5jrs2f963zs1-en>.
- OECD (2015b), *Daejeon Declaration on Science, Technology, and Innovation Policies for the Global and Digital Age*, <http://www.oecd.org/fr/sti/daejeon-declaration-2015.htm>.
- OECD (2015c), *Inquiries into Intellectual Property’s Economic Impact*. OECD Publishing, Paris, <http://www.oecd.org/sti/ieconomy/intellectual-property-economic-impact.htm>.
- OECD (2013a), *Commercialising Public Research: New Trends and Strategies*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264193321-en>.
- OECD (2013b), *New Data for Understanding the Human Condition*, <http://www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.htm>.
- OECD (2008), *Report on Roadmapping of Large Research Infrastructures*, <https://www.oecd.org/sti/sci-tech/41929340.pdf>.
- OECD (2003), *Governance of Public Research*, OECD Publishing, Paris, <http://www.oecd-ilibrary.org/docserver/download/9203051e.pdf?expires=1518008767&id=id&accname=ocid84004878&checksum=A04708A65E7C838E651279C77CFC293B>.
- Oliver, S., Armes, D. G. and G. Gyte (2009), “Public involvement in setting a national research agenda: a mixed methods evaluation”, *The Patient*, Vol. 2/3, Springer Nature, pp. 179-90.
- O’Neil, C. (2016), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown.
- Open Data Commons. (2016), “Licenses FAQ. Retrieved from Open data commons”, http://opendatacommons.org/faq/licenses/#Why_Not_Use_a_Creative_Commons_or_FreeOpen_Source_Software_License_for_Databases (accessed 6 July 2016).
- Open Data Institute. (1 August 2016). “Publisher’s Guide to Open Data Licensing”, <https://theodi.org/guides/publishers-guide-open-data-licensing> (accessed 1 August 2016).
- Open Source Hardware Association (2016), “Open-Source Hardware FAQ”, <http://www.oshwa.org/faq/> (accessed 8 August 2016).
- ORCID (2012) “What is ORCID”, <http://about.orcid.org/about/what-is-orcid/mission>.
- Osimo, D., L. Pujol. and F. Porcu (2015), *Analysis of Emerging Reputation and Funding Mechanisms in the Context of Open Science 2.0, Part 2*, Publications Office of the European Union, Luxembourg.
- Owens, B. (2016), “Montreal institute going 'open' to accelerate science”, <http://www.sciencemag.org/news/2016/01/montreal-institute-going-open-accelerate-science>.
- Pearce, J. M. (18 November 2013), “3D-printing your lab equipment”, Elsevier Connect, <https://www.elsevier.com/connect/3d-printing-your-lab-equipment-its-cheaper-than-you-think>.

- Pearce, J. M. (2012), “Building Research Equipment with Free, Open-Source Hardware”, *Science*, Vol. 337/6100, AAAS, Washington, D.C., pp. 1303-1304.
- Ponte, D. and J. Simon (2011), “Scholarly Communication 2.0: Exploring Researchers' Opinions on Web 2.0 for Scientific Knowledge Creation, Evaluation and Dissemination”, *Serials Review*, Vol. 37/3, Taylor & Francis Publishers, New York, pp. 149–156.
- Pöschl, U. (2012), “Multi-stage open peer review: scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation”, *Frontiers in Computational Neuroscience*, <https://doi.org/10.3389/fncom.2012.00033>.
- Priem, J. et al. (2010), “Altmetrics: A manifesto”, <http://altmetrics.org/manifesto> (accessed 26 October 2010).
- Rowe, G. and L. J. Frewer (2005), “A Typology of Public Engagement Mechanisms”, *Science, Technology, & Human Values*, Vol. 30 /2, SAGE publications, pp. 251-290.
- Rowe, G. and L. J. Frewer (2004), “Evaluating Public-Participation Exercises: A Research Agenda”, *Science, Technology, & Human Values*, Vol. 29/4, SAGE publications, pp. 512-556.
- Spiegelhalter, D. J. (2014), “The future lies in uncertainty”, *Science*, Vol. 345/6194, AAAS, Washington, D.C., pp. 264-265.
- State Council of China (2014), “Opinions of the State Council on Opening up Major National Scientific Research Infrastructure and Large-scale Scientific Research Instruments to the Public”, http://www.gov.cn/zhengce/content/2015-01/26/content_9431.htm.
- Sublet, V., C. Spring and J. Howard (2011), Does Social Media Improve Communication? Evaluating the NIOSH Science Blog”, *American Journal of Industrial Medicine*, Vol. 54/5, John Wiley & Sons Ltd, pp. 384–394.
- Sweeney, L. and M. Crosas (2013), *An Open Science Platform for the Next Generation of Data*, <http://dx.doi.org/10.2139/ssrn.2302084>.
- Tattersall, A. (2016), “Open Peer Review” in *Altmetrics: A practical guide for librarians, researchers and academics*, Facet publishing, London.
- The Open Science and Research Initiative (ATT) (2014), “The open science and research handbook”, <https://openscience.fi/handbook>.
- Thiel, G. v. and P. Stolk (2013), “Patient and Citizen Involvement”, http://www.who.int/medicines/areas/priority_medicines/BP8_5Stakeholder.pdf.
- Turney, J. (2011), “Democratic experiments”, Times Higher Education, <https://www.timeshighereducation.com/features/democratic-experiments/417989.article> (accessed 1 August 2016).
- Ubaldi, B. (2013), “Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives”, *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5k46bj4f03s7-en>.
- United Nations Environment Programme (UNEP) (2014), “Realizing the Potential of Citizen Science”, in *UNEP Year Book 2014 emerging issues update*, <http://staging.unep.org/yearbook/2014/PDF/chapt6.pdf>.
- Ware, M. (2008), “Peer review: benefits, perceptions and alternatives”, Publishing Research Consortium, London.
- Wiggins, A. and K. Crowston (2011), “From Conservation to Crowdsourcing: A Typology of Citizen Science”, proceeding of the 2011 44th Hawaii International Conference on System Sciences.
- Willinsky, J. (2005), “The unacknowledged convergence of open source, open access, and open science”, *First Monday*, Vol. 10/8, <http://dx.doi.org/10.5210/fm.v10i8.1265>.
- Zhao, Y. and Q. Zhu (2014), “Evaluation on crowdsourcing research: Current status and future direction”, *Information Systems Frontiers*, Vol. 16/3, Springer Nature, pp 417–434, <https://doi.org/10.1007/s10796-012-9350-4>.