First Monday, Volume 19, Number 9 - 1 September 2014

PEER-REVIEWED JOURNAL ON THE INTERNET

Putting open science into practice: A social dilemma?

by Kaja Scheliga and Sascha Friesike

Abstract

Digital technologies carry the promise of transforming science and opening up the research process. We interviewed researchers from a variety of backgrounds about their attitudes towards and experiences with openness in their research practices. We observe a considerable discrepancy between the concept of *open science* and scholarly reality. While many researchers support open science in theory, the individual researcher is confronted with various difficulties when putting open science into practice. We analyse the major obstacles to open science and group them into two main categories: *individual obstacles* and *systemic obstacles.* We argue that the phenomenon of open science can be seen through the prism of a social dilemma: what is in the collective best interest of the scientific community is not necessarily in the best interest of the individual scientist. We discuss the possibilities of transferring theoretical solutions for social dilemma problems to the realm of open science.

Contents

1. Introduction

Open science is the concept of making the whole research process as transparent and accessible as possible. It is a topic that concerns scientists, policy–makers, and the general public. Supporting open science is part of the European Commission's Digital Agenda. As Neelie Kroes, Vice–President of the European Commission, put it: 'I'm in no doubt we are entering that phase [the era of open science]: and that the impact will be good for citizens, good for scientists and good for society.' (Kroes, 2013).

Open science in its current form is based on digital technologies. Online tools provide scientists with the technical means to collaborate globally and to share knowledge on an unprecedented scale. 'The Internet and the World Wide Web provide the technical ability to share a much wider range of the evidence, argument and conclusions driving modern research. (...) The potential of online tools to revolutionize scientific communication and their ability to open up the details of the scientific enterprise so that a wider range of people can participate is clear.' (Neylon and Wu, 2009).

Visions centred on harnessing the world's knowledge have occupied intellectuals throughout history. For example, between 1936 and 1938 H.G. Wells developed the idea of a 'world brain' which functions as a universally accessible knowledge resource (Wells, 1938). In 1945 Vannevar Bush described his vision of the 'memex' which functions as a collective memory system that makes knowledge accessible (Bush, 1945). Recently, in 2012, Nielsen also described a concept of a form of collective memory system [1] whereby all forms of knowledge that are of potential scientific value are transferred onto a network in a format that is both human– and machine–readable [2]. Making knowledge resources accessible, searchable, and re–usable increases the chances of discovering previously unseen connections. 'We are reinventing discovery, and the result will be a new area of

networked science that speeds up discovery, not in one small corner of science, but across all science. That reinvention will deepen our understanding of how the universe works and help us address our most critical problems.' [3]. Speeding up the progress of science and providing immediate access to scientific results is of crucial importance. In medical research for example access to knowledge about new findings can be a matter of life and death. In environmental sciences for instance comparing findings from different disciplines allows to draw more accurate conclusions about changes that are observed in the environment.

The potential of using digital technologies to openly share knowledge thereby revolutionising science, however, is not reflected in scholarly practices (Neylon and Wu, 2009; Procter, *et al.*, 2010b). There is a considerable discrepancy between the idea of open science and scholarly reality. While many scientists consider openness in research valuable, few are actually willing to invest the extra time and effort as well as taking potential risks to make their research open and accessible (Procter, *et al.*, 2010b). At the same time, the rapidly changing environment of digital technologies continually shapes the development of open science in its various forms. Thus, we consider it important to examine the current changes and to analyse them in the context of previous research.

In this paper we look at open science through the prism of a social dilemma. Based on the analysis of 22 semi–structured qualitative interviews with researchers from various backgrounds we identified a series of obstacles to open science which we have grouped into two main categories: *individual obstacles* and *systemic obstacles*.

In the subsequent sections of this paper we first describe the concept of open science and the relevant background concerning social dilemmas. Second, we describe our methodological approach and explain how we conducted and analysed the interviews. Third, we describe the identified obstacles to open science. Fourth, we reflect on possible solutions to overcome the described obstacles. And finally, we draw conclusions from the presented findings.

2. The field of open science

At large the concept of open science is nothing new. Historically, one can argue that the emergence of the scientific journal system in the seventeenth century marks a decisive step in the development of open science (Nielsen, 2008). From a contemporary point of view, the open science movement has gained momentum with the popularisation of the internet from the 1990s onwards (Lievrouw, 2010). Even though open science is not a new phenomenon, the recent technological developments have added a new dimension to it. Digital technologies constitute the technical foundation for the contemporary form of open science, allowing scientists to share knowledge and to collaborate in ways that were not possible before (Meyer and Schroeder, 2013). It is worth noting, however, that open science is not tied to the Internet and that scholarly exchange offline retains its importance.

There is no formal definition of open science. We have asked our interview partners about their understanding of open science and have received answers that differ in detail but share a general vision. This vision of open science is well captured by Nielsen's informal definition: 'Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process.' (Nielsen, 2011a).

Open science is often discussed as a phenomenon that has an enormous potential to revolutionise science (Neylon and Wu, 2009; Nielsen, 2012; Waldrop, 2008). Digital technologies are helping to transform scholarship and knowledge sharing (Borgman, 2007; Meyer and Schroeder, 2013). The open access movement (Berlin Declaration, 2003) has contributed to a global dissemination of knowledge. An increasing amount of scientific publications is made available in the form of open access. A recent report released by the European Commission 'estimates that more than 40 percent of scientific peer reviewed articles published worldwide between 2004 and 2011 are now available online in open access form.' (European Commission, 2013). There is also a growing emphasis on sharing data and ways to make them re–useable and citable (Royal Society, 2012). The accessibility and the reusability of research resources as well as the availability of an infrastructure to exchange research materials speeds up the research process (De Roure, *et al.*, 2010; Hannay, 2009). A multitude of online tools can be used to share research materials and exchange knowledge (Lievrouw, 2010; Nentwich and König, 2011; Neylon and Wu, 2009; Procter, *et al.*, 2010b). Scientists can connect with each other on a global scale via social networking sites or exchange knowledge on online platforms, collaboratively write scientific publications, share research materials such as papers, data or code in repositories, disseminate intermediate research insights or updates about current research on blogs and microblogs or in open lab notebooks. Greater transparency throughout the research process opens up opportunities to receive valuable feedback and increases chances for collaborations (Gowers and Nielsen, 2009; Nielsen, 2012). Open science can be seen as a mechanism of cumulative knowledge production whereby scientists draw upon knowledge derived at by 'prior researchers' and make their discoveries available to 'future researchers' (Mukherjee and Stern, 2009).

Opening up the scientific process and making its results accessible to a wider audience plays an important role in disseminating knowledge on a global scale and providing not only scientists but also interested individuals with an insight into the current state of knowledge production (Cribb and Hartomo, 2010; Grand, *et al.*, 2012). The boundaries of science have blurred. 'Science and society have both become transgressive; that is, each has invaded the other's domain, and the lines demarcating the one from the other have all but disappeared.' (Nowotny, *et al.*, 2001).

There are open science initiatives in various areas of science that serve as positive examples of opening up the research process. In mathematics the Fields Medalist Timothy Gowers initiated a collaborative approach to solving a complex problem via a call for action on his blog (Gowers and Nielsen, 2009). In genetics the Human Genome Project and its successor ENCODE (http://www.genome.gov/10005107) made a significant contribution towards establishing the sharing of genetic sequences as the norm within the discipline. In disciplines such as physics, mathematics, and computer science it is common practice to upload papers onto the e–print repository arXiv (http://arxiv.org/). The Directory of Open Access Journals (http://doaj.org/) lists available open access journals. The Open Science Grid (http://www.opensciencegrid.org/) provides scientists with a distributed computing infrastructure (Pordes, *et al.*, 2008). FigShare (http://figshare.com/) provides a platform for sharing research materials, including data, while scientists who work with code can share it on GitHub (https://github.com/). As a social networking site for scientists, ResearchGate (http://www.researchgate.net/) is an example of a platform that fosters the exchange between researchers.

On the surface it seems as though there are numerous open science initiatives. Many of them, however, fail to reach a critical mass (Procter, *et al.*, 2010a) and end up being 'virtual ghost towns' (Nielsen, 2011b). There are not only benefits but also barriers and constraints to openness in research. When it comes to using digital technologies for scholarly communication the major barrier is the lack of benefits, combined with high costs of adoption and the lack of trust (Procter, *et al.*, 2010a; RIN and NESTA, 2010). The willingness to publish work–in–progress as well as to share data decreases with a growing audience. Scientists are more likely to share within a trusted network or research community than on a publicly accessible Web site or blog. Time, credit, and personality as well as the discipline specific culture play an important role in determining attitudes towards sharing knowledge (Acord and Harley, 2013). Many scientists adhere to peer review as the most important, though admittedly imperfect, mechanism to ensure quality and scientific standards (Waldrop, 2008). Therefore some of them are reluctant to share their work in formats that have no formal quality stamp and that are not directly rewarded by the academic system (Procter, *et al.*, 2010a). Even though there are efforts to measure scientists' output on the Web (Priem and Hemminger, 2010) these activities are not officially taken into account when it comes to evaluating scientists' work in terms of career progress. Traditional norms and incentives are still prevailing in the current academic system and are determining scientists' inclinations towards the forms of sharing knowledge. Overall, sharing research materials is context dependent; scientists balance potential future reciprocity against the loss of competitiveness (Haeussler, 2011). There are many variations in sharing practices in open science that are shaped by 'the diversity of scientific purposes, the technical nature of tasks, and the details of organizational structures.' [4]

3. The field of social dilemmas

The concept of a *social dilemma* has been extensively studied in various disciplines (see Kollock, 1998). The notion of a social dilemma can be applied to analysing scientists' behaviour with regards to attitudes towards openness. Generally, '[Social] dilemmas are defined by two simple properties: (a) each individual receives a higher payoff for a socially defecting choice ... than for a socially cooperative choice, no matter what the other individuals in society do, but (b) all individuals are better off if all cooperate than if all defect.' (Dawes, 1980) In other words, social dilemmas are situations where 'individual rationality leads to collective irrationality' (Kollock, 1998). There are different types of social dilemmas, most famously the prisoner's dilemma (Luce and Raiffa, 1957; Poundstone, 1992), the logic of collective action (Olson, 1965), and the tragedy of the commons (Hardin, 1968). These models are typically discussed in cases where direct reward or punishment comes into play for a cooperative or defecting choice of action (prisoner's dilemma), in situations where a collective effort is made to provide a public good and the problem of free–riding on the work of others arises (logic of collective action), and in the context of dealing with exhaustible goods such as natural resources or topics like overpopulation or pollution (tragedy of the commons).

The concept of the commons has been applied to describe social dilemma type problems arising in the process of digitalisation of knowledge and its distribution via the Internet (Ostrom and Hess, 2011). Knowledge can be seen as a 'shared resource, a complex ecosystem that is a commons — a resource shared by a group of people that is subject to social dilemmas.' (Ostrom and Hess, 2011).

Transferring the concept of a social dilemma to the realm of open science one can say that a scientist

who acts driven by individual rationality shares knowledge in a form that drives forward his or her career. If all scientists only share knowledge in situations where they expect to gain individual profit, the common pool of knowledge is fragmented and all scientists are collectively worse off.

4. Materials and method

The findings presented here are based on an extensive literature review on the topic of open science as well as Web research exploring open science initiatives and digital tools that can be used for open science practices. These insights served as the basis for conducting a series of qualitative interviews with researchers, developers of research infrastructure, and a representative of a funding agency. All interviewees were experts in their field and additionally had a connection to open science.

We employed a purposive sampling strategy (Teddlie and Yu, 2007) because our main concern was to make the sample most relevant to the research question. Being able to make generalizations to a population was not a concern in this case. We followed a sequential sampling approach and conducted 22 qualitative interviews between February and June 2013. The total length of the interviews was 12 hours and 41 minutes, the average interview length was 33 minutes. The majority of the interviews were conducted face–to–face during the Science 2.0 workshop of the Leibniz Research Association in Hamburg (February 2013), the International Conference on Internet Science in Brussels (April 2013) and the re:publica 13 conference in Berlin (May 2013). Additional interviews were conducted via Skype. The sample consists of scientists based in Germany, England, Switzerland, and South Africa; 18 of the interviewees were male and four were female. All interviews were recorded with permission and subsequently transcribed verbatim, summing up to a total of about 80 thousand words.

The aim of the interviews was to get a detailed understanding about the current core issues of open science in an academic context. Thus, we focused on researchers and research infrastructure developers familiar with the concepts of open science. We interviewed academics at different stages of their career from the humanities, the social sciences, and the natural sciences in order to cover a broad range of perspectives. The backgrounds of the interviewees are summarised in Table 1.

Researcher		
Background	**PhD Candidate**	**Senior Researcher**
Digital Humanities	2	2
Humanities		1
Social Sciences		1
Computer Science		1
Astronomy		1
Medicine		1
Biology	1	
Biophysics		1
Mediapsychology		1
Media and Communication	1	2
Research Infrastructure		
Information Science	4	
Library and Information Science	1	
Medicine	1	
Research Funding		
German Research Foundation	1	

Table 1. Interviewees' backgrounds.

The interviews were semi–structured and had an exploratory character (Wengraf, 2001). The core questions concerned the perception of the concept of open science, the degree of openness in research practices as well as the boundaries of openness. The impact of technology upon research practices, the use of online tools for research purposes and issues concerning data sharing were also explored. While the core questions remained the same, we allowed ourselves the freedom to follow–up on issues mentioned during the conversation and to adjust the questions according to the expertise of the individual interviewees. The idea was to keep an open mind and not to impose any preconceptions. When conducting the interviews we paid attention to interview guidelines and ethics (Fontana and Frey, 1994).

We analysed the interview material using the software QRS NVivo 10. All transcripts were coded using initial coding characterized by low inference, followed by inferential coding in order to draw out patterns and higher–order concepts inherent in the data (Coffey and Atkinson, 1996). During the initial coding process we identified themes emerging from the interview material. We grouped the statements accordingly and systematically refined the themes into categories. Subsequently, we focussed on those categories that were relevant for discussing the phenomenon of open science in the context of a social dilemma.

4.1. Limitations

It is important to keep in mind that the presented results have limitations. Firstly, the sample size of 22 interviews is relatively small and cannot serve as a basis to draw generalizable conclusions. Secondly, there is a bias towards a European and particularly a German perspective due to the fact that two–thirds of the sample is constituted by researchers based in Germany. Furthermore, the exploratory nature of the interviews as well as the various perspectives that have been taken into account result in covering a wide spectrum but simultaneously keeping insights on a general level.

5. Looking at open science through the prism of a social dilemma

In analysing the interview material with a focus on obstacles to openness we used the notion of a social dilemma as our theoretical framework. We follow the line of argument that what is best for the collective is not necessarily in the best interest of the individual. In other words, what is best for science on a collective level is not in the best interest of the individual scientist.

> If one opens up as an individual one does not clearly, not evidently get the benefit. It is quite possible that one suffers some damage from it.
> (Senior Researcher, Biophysics)

5.1. The vision of open science

Many interviewed researchers have expressed their endorsement for the idea of open science. They describe that sharing knowledge via the Internet provides them with expanded opportunities to exchange feedback and to collaborate with scientists internationally.

> The vision that all of Europe is a huge playground for scientists where there are no borders anymore, no barriers to work together, that is indeed a great vision.
> (Research Infrastructure Developer)

They state that open access to research resources speeds up the research process and makes it more efficient. One proponent of data re–use argued that sharing research materials contributes to the goal of making science a commons.

This optimistic vision of open science, however, was not always reflected in the behaviour of the individual researchers. Only a fraction of the researchers had the intention to make the whole research lifecycle as transparent and accessible as possible. Mentioned examples of specific open science activities mainly referred to open access to publications and presentations, and only in a few cases to data and code. One scholar encouraged to take a critical view on openness and digital ways of working in particular:

> I take notes in a paper book and I write with a fountain pen and I have a very traditional way of working ... I think it is

> interesting to go through the medium of paper to think about digital, because digital tools seem to offer us all these solutions and we are very caught up in the digital as providing answers, and I think we have to pause a little bit and paper forces you to pause ..., it slows you down and it is that slowing down I think is really really important in doing research.
> (Senior Researcher, Media and Communication)

The discrepancy between the idea and practice of open science was also described by researchers who are working on developing research infrastructure. Scientists support the concept of making work available to others but are reluctant when it comes to sharing their own work.

> Many have said, yes yes, on the wave of sharing I am for mutual exchange but always when it comes to it there were good arguments to say, I cannot share my text yet.
> (Research Infrastructure Developer, Information Science)

The same mechanism occurs when it comes to sharing data. Scientists are in favour of the idea of sharing and re–using data but are hesitant to make their own datasets available.

> Everyone thinks data sharing is great but every time it comes to putting it into practice it becomes very difficult. One would like to re–use colleagues' data but sharing one's own data is obviously problematic because often these big challenges are in focus.
> (Ph.D. Candidate, Information Science)

5.2. Obstacles to open science

Based on the interviews we have identified various obstacles which scientists need to overcome in order to practice open science. We have grouped these obstacles into two main categories: *individual* and *systemic* obstacles. Individual obstacles are linked to the personality and attitude of the researcher. Systemic obstacles are determined by the research environment which is subject to a variety of influences. The obstacles are summarised in Figure 1 and are described in more detail in the subsequent sections.

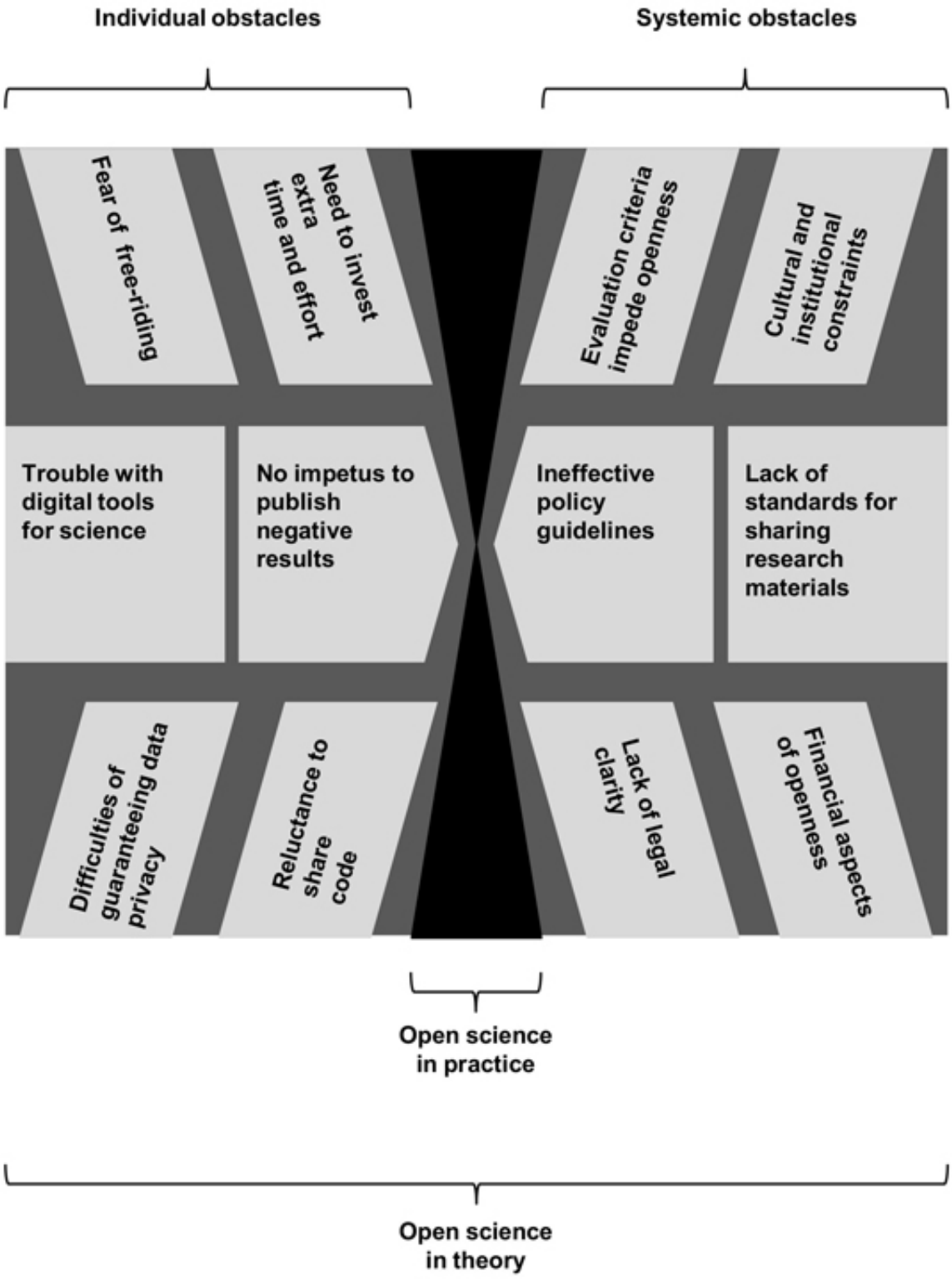

Figure 1. Obstacles to open science.

5.3. Individual obstacles to open science

The fear of free–riding

The fear of free–riding is a recurring topic in the interviews. Scientists fear that if they release their research materials early on in the research process they expose themselves to intellectual property abuse. Sharing initial research ideas carries the risk of someone else taking them over. Releasing intermediate results can lead to other researchers building on that work and reaping the rewards for

it. Publishing data diminishes the chances of using the same dataset for subsequent publications because other scientist can re–use them and already cover that topic. Making source code available can result in a scientist losing his or her competitive advantage. In some areas sharing research insights stands in conflict with the possibility of securing a patent for the findings.

> [Idea theft] plays a role when one makes one's data or one's scientific findings accessible unintelligently. That is, when one does not give much thought to at what stage to make results accessible.
> (Ph.D. Candidate, Information Science)

The need to invest extra time and effort

The individual researcher needs to invest extra time and effort in order to put the idea of open science into practice. While papers or conference contributions are designed with an audience in mind, other research materials are often gathered without the intention of being accessed by a wide audience. A scientist who wants to share data needs to clean up the dataset, add descriptions that are understandable for someone who was not involved in the data collection process and in many cases anonymise the data. Similarly, a scientist who wants to share code needs to make sure that someone else can understand it by for instance adding extended explanatory comments.

> So that is an increased effort, the workflow on the one hand and on the other hand the recognition for this kind of effort are not grounded in scientific processes.
> (Programme Officer, German Research Foundation)

Trouble with digital tools for science

The current form of open science is possible thanks to digital tools and the Internet. There is a multitude of digital tools a scientist can potentially use for research purposes and each tool comes with its own learning curve. Choosing the best tool for a given task, learning how to use it and managing it in accord with an array of other digital tools is challenging for many scientists. Moreover, most digital tools have shortcomings when they are used for scientific tasks. While some scientists develop their own tools in order to have full control over their functionality it is an endeavour that requires not only the skills but also additional time and effort which again is hardly taken into account when assessing the output of a scientist.

> The main problem lies therein that the available tools have not been developed for science and cannot do some things that a scientists absolutely needs.
> (Research Infrastructure Developer)

No impetus to publish negative results

There is no impetus for scientists to publish negative results because it is an activity that requires additional work which is not rewarded by the academic system. Releasing negative results such as findings that show that a seemingly interesting research question leads to a dead end is considered an important aspect of open science because it helps to avoid re–inventing the wheel over and over again, thereby saving resources. Nevertheless, many scientists consider turning such a negative finding into something publishable a waste of time because it does not contribute to their individual profit.

> Taking a result and turning it into something that you can publish is quite a lot of work, you know, formatting, writing, making sure everything is right and so on, ... so if you have a negative result, taking that extra amount of work when you can actually go and look at something else, takes priority I think. So that is probably why people do not do it.
> (Senior Researcher, Astronomy)

Difficulties of guaranteeing data privacy

Providing access to data requires an extra effort on the part of the scientists to ensure data privacy. Even though many scientists work with non–sensitive data which can be shared widely, there are datasets that require scientists to use specific techniques to anonymise sensitive data. In some cases full anonymity cannot be guaranteed and scientists who want to make the data available nevertheless can use access control mechanisms. All of these solutions require extra efforts on the part of those scientists providing access to the data, in addition to tasks such as making the dataset understandable to somebody who was not involved in the data collection process.

> Mak[ing] the data accessible ... put[s] a lot of pressure on

> you to have to anonymise data if it involves any personal identifying information, and that is fine, there are techniques that are reasonable for that, you might control access to the data if you cannot fully anonymise it.
> (Senior Researcher, Computer Science)

Reluctance to share code

A scientist who has invested a lot of time and effort into developing code can be reluctant to make it available to others, especially if her or his competitive advantage relies on the code. While the open source movement propagates making code accessible and reusable, many scientists do not share their code because it conflicts with their personal interest.

> In most academic disciplines that I can think of there is still a bit of reluctance to share codes. ... if you share code you do not necessarily get citations for code and, you know, citations are the currency to get jobs.
> (Senior Researcher, Astronomy)

5.4. Systemic obstacles to open science

Evaluation criteria impede openness

Many researchers stated that open science efforts are not rewarded by the current academic system. Some have argued that established academics can afford contributing to open science while researchers who are still competing for a position in academia are forced to invest all their energies in research activities that are rewarded and can drive forward their career. Since scientists are mainly evaluated on the basis of their publications in high–ranked and peer–reviewed journals, many researchers prioritise on securing as many publications as possible. Posting shorter texts on blogs, exchanging knowledge via social research networks or sharing references to relevant literature in collaborative digital libraries is not taken into account when measuring the scientific output of the individual researcher. There is no standard format for citing data and code which makes it difficult to acknowledge these kinds of contributions. Even though making data and code available is generally appreciated, the lack of incentives prevents the individual scientist from sharing research materials in most cases.

> But I think an important point is that we have an academic system that is extremely shaped by competition. And that the incentives are of course based on publications ... and of course each scholar strives to get the best out of the reputation system for oneself and that clearly stands in potential conflict with the demand for open science, it does not go together well at the moment.
> (Ph.D. Candidate, Information Science)

Cultural and institutional constraints

The culture of a specific discipline has an impact on the attitude of the individual researcher towards openness. In a culture that generally embraces openness, such as for example in astronomy, it is easier for the individual researcher to share. In disciplines where openness is not the cultural norm, it is even more difficult for an individual scientist to overcome various constraints.

> There are a lot of cultural but particularly institutional constraints to that, tenure track ... is still relying on very traditional ways of producing science, and those traditional ways, I am not saying that they discard the possibility of doing open science, but by no mean this is the mainstream. So people tend to go for other directions.
> (Senior Researcher, Digital Humanities)

Ineffective policy guidelines

There are policy guidelines to foster open science such as for example the appeal to make data available after an embargo period. Scientists noted that there are few mechanisms in place to ensure compliance with these guidelines. As a consequence, some scientists perceive existing policy guidelines as ineffective because there are no sanctions for non–compliance.

> The culture of sharing is mentioned by the European Commission but it is not strongly driven forward by science itself because the benefit is frequently not very clear but the burden can thereby increase.

(Programme Officer, German Research Foundation)

Lack of standards for sharing research materials

There is a lack of standards for sharing research materials. Even though there are repositories for publications as well as for data and code or collaborative digital libraries for sharing references, there are multiple standards that exist parallel to each other which impede interoperability. On the one hand knowing which standard to adhere to is difficult for the scientist who wants to share, on the other hand unconventional material is difficult to find. Especially interdisciplinary work can be problematic to fit into existing formats.

> From a technical standpoint that is another big hurdle to share data is that actually the formatting standards. I mean one person's data is not interpretable by another person unless it's in an agreed international standard or good formats.
> (Ph.D. Candidate, Biology)

Lack of legal clarity

In the context of international research projects scientists can be confronted with potential legal difficulties on the grounds of being subject to different jurisdictions. Especially when it comes to data, there are various forms of data ownership. Thus, a scientist who wants to share research outputs often faces a series of legal questions and a lack of clarity as to what complies with the laws, rules, and regulations.

> What I consider to be the main constraint is that in principal not everyone and not at all times is clear about which type of law can be applied to the data.
> (Programme Officer, German Research Foundation)

Financial aspects of openness

A scientist who wants to share her or his research materials needs to take the costs of providing access to it into account. Providing and maintaining the infrastructure is expensive. The costs of publishing in an open access format need to be covered. In case of a tight budget a scientist can be confronted with the lack of means to finance the access to research materials.

> Then there are certainly boundaries which are financially induced. One needs ... infrastructure ... with which one can make those data accessible intelligently.
> (Senior Researcher, Social Sciences)

6. Overcoming the obstacles to open science

In spite of the various difficulties which the individual researcher who wants to practice open science is confronted with, some interviewees also pointed towards potential benefits of engaging in open science activities.

> I am of the opinion that certain questions, probably not all but still certain questions, can indeed profit from being approached in an open way.
> (Senior Researcher, Biophysics)

In some disciplines, such as astronomy, openness is integrated into the discipline specific culture.

> Astronomy is traditionally quite an open science. I mean people are quite open about it, and very happy to share, you know, discuss results and things like that.
> (Senior Researcher, Astronomy)

Some interviewees argued that sharing research output is not only in the collective interest of the scientific community but can also be advantageous for the individual researcher.

> [Sharing data is] actually good for oneself. It's good for science. It's also good for oneself. Because if you share data and someone else uses that data then they'll cite you. Because where they get that data from got be cited.

(Ph.D. Candidate, Biology)

Several scientists argued that making their work accessible increases chances for potential collaborations. Moreover, providing access to research output can create a positive feedback loop.

> [It]is largely about showing the benefit, you know if your work is more visible and you are any good at that then you are going to get famous and you get a career, so there is a selfish reason to be giving your work to other people.
> (Senior Researcher, Computer Science)

Some scientists share their work strategically where the potential benefits outweigh the risks. Others are more intrinsically motivated and consider sharing to be an experimental form that can bring about benefits but where a sense of joy experienced from sharing knowledge with others also plays a role.

> So the first book that I wrote, the first day that we finished the book we decided to make it open. ... And the book now it got one quarter million downloads, [it] is amazing, and we never, never thought that. I mean, not even pornographic books are that successful.
> (Senior Researcher, Digital Humanities)

A few scientists were willing to sacrifice rewards in some areas on the traditional track and engage in experimenting with novel forms of disseminating knowledge.

> Young people are exploring other channels, young people are keen to put things on YouTube or SlideShare or Wikipedia or whatever, even acknowledging that this might not be translated into more score for the tenure track or other things.
> (Senior Researcher, Digital Humanities)

There were divided opinions concerning the development of open science. Some scientists argued that there are no major differences in attitudes towards openness between scientists from different generations because the common denominator that binds them is acting according to the rules of the current academic system. Others argued that the next generation of scientists, who grew up in the technologically determined landscape where an increasing amount of resources can be accessed via the Internet, will have a more open approach to sharing knowledge and incorporate that into their research practices.

Overall, several scientists stressed the significance of sharing research materials intelligently. They argued that the key to sharing knowledge that minimizes the risk of intellectual property abuse is to develop an awareness of what material is shared with whom and at what stage of the research process. Moreover, they raised the point that not only the access to research materials but also their quality is essential.

> Now is rather the time ... to discuss the quality of openness.
> (Senior Researcher, Digital Humanities)

In summary, the interview material reveals the following trends. The majority of the interviewees perceive the idea of open science positively. Some researchers, however, are trapped in a social dilemma situation: they consider putting the idea of open science into practice to be an endeavour that comes with a series of additional difficulties for themselves even though they are aware that overcoming those difficulties is in the collective interest of the scientific community. A few researchers indicate approaches to resolve the social dilemma situation by seeing personal profit in engaging in open science activities and thereby having the motivation to overcome the difficulties of putting open science into practice.

7. Discussion

The results show that researchers who want to practice open science are confronted with a series of obstacles. In some respect the described obstacles echo popular arguments that have been voiced by proponents of open science on the one hand and open science sceptics on the other. The interviews focus on reflecting the current state of open science practices among researchers, highlight the prevailing importance of open science, and simultaneously point to the need of continuing to develop solutions that tackle the obstacles at hand. Grouping the obstacles to open science into systemic and individual ones implies that not only the problems but also the solutions are to be found on both

sides.

When looking at open science through the prism of a social dilemma it seems natural to consider applying approaches to solving social dilemma problems to obstacles encountered in the realm of open science. Various solutions to social dilemmas can be found in the literature (Kollock, 1998). It is important to keep in mind that many observations about social dilemma problems were developed based on designed experiments conducted in isolation. Consequently these solutions hardly match the complexity of real life situations. Nevertheless, existing approaches to overcome social dilemma problems can offer valuable insights for overcoming obstacles to open science and complement existing initiatives that support open science practices. Based on an extensive literature review of social dilemma research Kollock suggests that there are three categories of solutions to social dilemma problems: *motivational, strategic,* and *structural* solutions (Kollock, 1998).

Motivational solutions take the actions of other individuals into account. Social value orientations, communication, and group identity all play a role. Applied to open science one can argue that the fear of free–riding is of less concern if priority of work can be proven by means of a digital timestamp, for instance. Investing extra time and effort stops being a barrier if open science practices are an integral part of the research process and are valued by the research community. Rewarding scientists for developing their own tools and encouraging them to share their solutions means increasing the spectrum of available science specific tools. At the same time, it reduces the conflict between conducting research and developing tools which help to conduct that research. Making negative results publishable and citable increases the incentives to go through the effort of putting the data into an appropriate format. Guaranteeing data privacy in the context of increasingly complex data relationships will remain a challenge but can be faced with increasingly sophisticated solutions in accordance with ethical considerations. The reluctance to share code can be overcome by adopting successful solution from the open source movement as well as rewarding scientists for sharing code and making code citable in a comparable way to other forms of scientific output.

Strategic solutions assume egoistic behaviour of the individual. Reciprocity, the choice of partners, and social learning are important factors. Strategic solutions applied to open science focus on the benefits of making a scientist's work more visible by making it openly accessible and disseminating it broadly which can result in increased levels of citations and more opportunities for collaborations. There are individual scientists who are convinced that making their work accessible is in their best egoistic interest and they do not shy away from sharing their scientific output even if other scientists are hesitant to do so.

In *structural* solutions the rules of the game can be changed. In this case factors such as iteration and identifiability, payoff structure efficacy as well as group size are important. Translated to open science changing the rules of the game refers to overcoming the obstacles on the systemic level. Integrating efforts towards open research practices in evaluation criteria is a transformation that requires time and careful considerations. Solutions towards measuring alternative impacts of scientific output are being developed under the label 'altmetrics' (http://altmetrics.org/manifesto/). In this context it needs to be taken into account that openness needs to be managed and its quality needs to be addressed. Cultural and institutional constraints can be diminished if open research practices are accepted, encouraged and rewarded within a research culture and by research institutions. Policy guidelines are linked to the research culture and can be put into practice more easily if there is no conflict between policy guidelines and institutional guidelines. The lack of standards for sharing research materials is challenging because there is such a variety of them. There are analogue problems concerning software and few solutions to develop universal standards and interoperability between various forms of systems. More clarity, especially with reference to international research projects, can be achieved by working with open licences and agreeing upon a strategy of sharing research materials that all project partners adhere to at the outset of a research project. Finally, more funding opportunities for open initiatives in scientific research would make open research practices more attractive to researchers.

From another point of view Dawes suggests that the 'three important ingredients for enhancing cooperation in social dilemma situations may be: knowledge, morality, and trust.' (Dawes, 1980). These universal ingredients can be considered key factors in open science also. Knowledge about established practices in a specific discipline is important for integrating open science practices into the research process. Morality is the basis for researchers' ethical behaviour within and beyond the scientific community. And trust plays a role in determining the degree of openness and the kinds of materials researchers are willing to share.

Looking at specific successful open science projects serves as a source of inspiration for developing further solutions. As mentioned earlier, there are bottom-up initiatives such as Timothy Gowers' Polymath Project [5], top–down initiatives that establish the rules for a discipline such as the Human Genome Project in genetics and initiatives that provide the infrastructure to integrate citizens into the scientific discovery process such as the citizen science project Galaxy Zoo (http://www.galaxyzoo.org/) in astronomy.

Considering that open science can occur in many different shapes and sizes it is not possible to

identify a universally applicable solution to overcome occurring social dilemma problems. Both self–organising governance structures, such as described by Elinor Ostrom in 'Governing the commons' (Ostrom, 1990) for instance, and top–down regulation are applicable to open science. While it is evident that some form of governance is necessary, we strongly feel that fostering openness in research cannot simply mean forcing scientists towards openness.

8. Conclusion

Digital technologies offer a broad array of tools that can be used to contribute towards making knowledge accessible as a commons. As Ostrom and Hess argue: 'An infinite amount of knowledge is waiting to be unearthed. The discovery of future knowledge is a common good and a treasure we owe to future generations. The challenge of today's generation is to keep the pathways to discovery open.' (Ostrom and Hess, 2011).

In analysing open science through the prism of a social dilemma we have discussed obstacles to open science practices and ideas for overcoming these obstacles. We have transferred solutions suggested for resolving general social dilemma problems to the realm of open science. We hope that our reflections contribute towards developing the concept of open science further. Moreover, we think that the field of open science invites to explore a broad variety of research questions from both discipline specific and interdisciplinary perspectives. In practice, future development of open science depends on the attitudes and activities of individual researchers as well as institutional initiatives to foster openness along the various stages of the research process.

About the authors

Kaja Scheliga studied English and Drama (B.A.) at Royal Holloway, University of London and Computer Science (M.Sc.) at University College London. At the Humboldt Institute for Internet and Society her research topic is open science.
E–mail: scheliga [at] hiig [dot] de

Dr. **Sascha Friesike** works as a postdoctoral researcher at the Alexander von Humboldt Institute for Internet and Society in Berlin, Germany. There he is in charge of the research group "Open Science". He holds an engineering degree from the Technical University in Berlin and a Ph.D. from the University of St. Gallen. Prior to his engagement in Berlin, he worked as a researcher at the Center for Design Research at Stanford University.
E–mail: friesike [at] hiig [dot] de

Acknowledgments

Many special thanks to all researchers who participated in the interviews! The interviews were conducted in collaboration with Benedikt Fecher (Alexander von Humboldt Institute for Internet and Society; German Institute for Economic Research (DIW), Berlin, Germany).

Thanks to Dr. Cornelius Puschmann for his valuable feedback.

This work was supported by the Alexander von Humboldt Institute for Internet and Society.

Notes

1. Nielsen, 2012, p. 175.

2. Nielsen, 2012, p. 183.

3. Nielsen, 2012, p. 207.

4. David, *et al.*, 2008, p. 311.

5. T. Gowers and M. Nielsen, 2009. "Massively collaborative mathematics,"*Nature*, volume 461, number 7266 (15 October 2009), pp. 879–881, doi: http://dx.doi.org/10.1038/461879a.

References

S.K. Acord and D. Harley, 2013. "Credit, time, and personality: The human challenges to sharing scholarly work using Web 2.0," *New Media & Society*, volume 15, number 3, pp. 379–397. doi: http://dx.doi.org/10.1177/1461444812465140, accessed 22 August 2014.

C.L. Borgman, 2007. *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, Mass.: MIT Press.

V. Bush, 1945. "As we may think," *Atlantic* (1 July), at: http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/, accessed 22 November 2013.

A. Coffey and P. Atkinson, 1996. *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks, Calif.: Sage.

J. Cribb and T.S. Hartomo, 2010. *Open science: Sharing knowledge in the global century*. Collingwood, Vic.: CSIRO.

P. David, M. den Besten and R. Schroeder, 2008. "Will e–science be open science?" *Stanford Institute for Economic Policy Research, Discussion Paper*, number 08–10, at http://www-siepr.stanford.edu/repec/sip/08-010.pdf, accessed 14 November 2013.

R.M. Dawes, 1980. "Social dilemmas," *Annual Review of Psychology*, volume 31, pp. 169–193. doi: http://dx.doi.org/10.1146/annurev.ps.31.020180.001125, accessed 22 August 2014.

D. De Roure, C. Goble, S. Aleksejevs, S. Bechhofer, J. Bhagat, D. Cruickshank, P. Fisher, D. Hull, D. Michaelides, D. Newman, R. Procter, Y. Lin and M. Poschen, 2010. "Towards open science: The myExperiment approach," *Concurrency and Computation*, volume 22, number 17, pp. 2,335–2,353. doi: http://dx.doi.org/10.1002/cpe.1601, accessed 22 August 2014.

European Commission, 2013. "Open access to research publications reaching 'tipping point'," *European Commission — IP/13/786* (21 August), at http://europa.eu/rapid/press-release_IP-13-786_en.htm, accessed 15 October 2013.

A. Fontana and J.H. Frey, 1994. "Interviewing: The art of science," In: N.K. Denzin and Y.S. Lincoln (editors). *Sage handbook of qualitative research*. Thousand Oaks, Calif.: Sage, pp. 361–376.

T. Gowers and M. Nielsen, 2009. "Massively collaborative mathematics,"*Nature*, volume 461, number 7266 (15 October 2009), pp. 879–881. doi: http://dx.doi.org/10.1038/461879a, accessed 22 August 2014.

A. Grand, C. Wilkinson, K. Bultitude A.F.T. Winfield, 2012. "Open science: A new 'trust technology'?" *Science Communication*, volume 34, number 5, pp. 679–689. doi: http://dx.doi.org/10.1177/1075547012443021, accessed 22 August 2014.

C. Haeussler, 2011. "Information–sharing in academia and the industry: A comparative study," *Research Policy*, volume 40, number 1, pp. 105–122. doi: http://dx.doi.org/10.1016/j.respol.2010.08.007, accessed 22 August 2014.

T. Hannay, 2009. "From Web 2.0 to the global database," In: A.J.G. Hey, S. Tansley and K.M. Tolle (editors). *The fourth paradigm: Data–intensive scientific discover*. Redmond, Wash.: Microsoft Research, at http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_part4_hannay.pdf, accessed 22 August 2014.

G. Hardin, 1968. "The tragedy of the commons," *Science*, volume 162, number 3859 (13 December), pp. 1,243–1,248, and at http://www.sciencemag.org/content/162/3859/1243.full, accessed 22 August 2014. doi: http://dx.doi.org/10.1126/science.162.3859.1243, accessed 22 August 2014.

P. Kollock, 1998. "Social dilemmas: The anatomy of cooperation," *Annual Review of Sociology*, volume 24, pp. 183–214. doi: http://dx.doi.org/10.1146/annurev.soc.24.1.183, accessed 22 August 2014.

N. Kroes, 2013. "Opening up scientific data," *European Commission — SPEECH/13/236* (18 March), at http://europa.eu/rapid/press-release_SPEECH-13-236_en.htm, accessed 22 August 2014.

L.A. Lievrouw, 2010. "Social media and the production of knowledge: A return to little science?" *Social Epistemology*, volume 24, number 3, pp. 219–237. doi: http://dx.doi.org/10.1080/02691728.2010.499177, accessed 22 August 2014.

R.D. Luce and H. Raiffa, 1957 *Games and decisions: Introduction and critical survey*. New York: Dover.

E.T. Meyer and R. Schroeder, 2013. "Digital transformations of scholarship and knowledge," In: W.H.

Dutton (editor). *Oxford handbook of Internet studies*. Oxford: Oxford University Press, pp. 307–327.

A. Mukherjee and S. Stern, 2009. "Disclosure or secrecy? The dynamics of open science," *International Journal of Industrial Organization*, volume 27, number 3, pp. 449–462. doi: http://dx.doi.org/10.1016/j.ijindorg.2008.11.005, accessed 22 August 2014.

M. Nentwich and R. König, 2011. *Cyberscience 2.0: Research in the age of digital social networks*. Frankfurt: Campus.

C. Neylon and S. Wu, 2009. "Open science: Tools, approaches, and implications," *Pacific Symposium on Biocomputing*, volume 14, pp. 540–544, at http://psb.stanford.edu/psb-online/proceedings/psb09/workshop-opensci.pdf, accessed 22 August 2014. doi: http://dx.doi.org/10.1038/npre.2008.1633.1, accessed 22 August 2014.

M. Nielsen, 2012. *Reinventing discovery: The new era of networked science*. Princeton, N.J.: Princeton University Press.

M. Nielsen, 2011a. "An informal definition of OpenScience," *OpenScience Project* (28 July), at http://www.openscience.org/blog/?p=454, accessed 28 October 2013.

M. Nielsen, 2011b. "Open science: Michael Nielsen at TEDxWaterloo" (6 April), at http://www.youtube.com/watch?v=DnWocYKqvhw&feature=youtube_gdata_player, accessed 14 November 2013.

M. Nielsen, 2008. "The future of science" (17 July), http://michaelnielsen.org/blog/the-future-of-science-2/, accessed 18 October 2013.

H. Nowotny, P. Scott and M. Gibbons, 2001. *Re–thinking science: Knowledge and the public in an age of uncertainty*. Cambridge: Polity.

M. Olson, 1965. *The Logic of collective action: Public goods and the theory of groups*. Cambridge, Mass: Harvard University Press.

E. Ostrom, 1990. *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press.

E. Ostrom and C. Hess, 2011. "Introduction: An overview of the knowledge commons," In: C. Hess and E. Ostrom (editors). *Understanding knowledge as a commons: From theory to practice*. Cambridge, Mass.: MIT Press, pp. 3–26, and at http://mitpress.mit.edu/sites/default/files/titles/content/9780262083577_sch_0001.pdf, accessed 22 August 2014.

R. Pordes, M. Altunay, P. Avery, A. Bejan, K. Blackburn, A. Blatecky, R. Gardner, B. Kramer, M. Livny, J. McGee, M. Potekhin, R. Quick, D. Olson, A. Roy, C. Sehgal, T. Wenaus, M. Wilde and F. Würthwein, 2008. "New science on the Open Science Grid," *Journal of Physics: Conference Series*, volume 125 number 1, at http://iopscience.iop.org/1742-6596/125/1/012070, accessed 22 November 2013. doi: http://dx.doi.org/10.1088/1742-6596/125/1/012070, accessed 22 August 2014.

W. Poundstone, 1992. *Prisoner's dilemma*. New York: Doubleday.

J. Priem and B.H. Hemminger, 2010. "Scientometrics 2.0: New metrics of scholarly impact on the social Web," *First Monday*, volume 15, number 7, at http://firstmonday.org/article/view/2874/2570, accessed 14 November 2013.
doi: http://dx.doi.org/10.5210/fm.v15i7.2874, accessed 22 August 2014.

R. Procter R. Williams and J. Stewart, 2010a. "If you build it, will they come? How researchers perceive and use Web 2.0," at http://wrap.warwick.ac.uk/56246/, accessed 23 September 2013.

R. Procter, R. Williams, J. Stewart, M. Poschen, H. Snee, A. Voss and M. Asgari–Targhi, 2010b. "Adoption and use of Web 2.0 in scholarly communications," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, volume 368, number 1926, pp. 4,039–4,056.
doi: http://dx.doi.org/10.1098/rsta.2010.0155, accessed 22 August 2014.

Research Informatiopn Network (RIN) and National Endowment for Science, Technology and the Arts (NESTA), 2010. "Open to all? Case studies of openness in research," at http://www.rin.ac.uk/system/files/attachments/NESTA-RIN_Open_Science_V01_0.pdf, accessed 11 November 2013.

Royal Society, 2012. "Science as an open enterprise," at http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf, accessed 17 October 2013.

C. Teddlie and F. Yu, 2007. "Mixed methods sampling: A typology with examples," *Journal of Mixed Methods Research*, volume 1, number 1, pp. 77–100.
doi: http://dx.doi.org/10.1177/2345678906292430, accessed 22 August 2014.

M.M. Waldrop, 2008. "Science 2.0: Great new tool, or great risk?" *Scientific American* (9 January), at http://www.scientificamerican.com/article.cfm?id=science-2-point-0-great-new-tool-or-great-risk, accessed 22 November 2013.

H.G. Wells, 1938. *World brain*. London: Methuen, and at http://www.ics.uci.edu/~vid/Readings/Wells_World_Brain.pdf, accessed 25 November 2013.

T. Wengraf, 2001. *Qualitative research interviewing: Biographic narrative and semi–structured methods*. Thousand Oaks, Calif.: Sage.

Editorial history

Received 12 May 2014; accepted 21 August 2014.

Putting open science into practice: A social dilemma?
by Kaja Scheliga and Sascha Friesike.
First Monday, Volume 19, Number 9 - 1 September 2014
http://journals.uic.edu/ojs/index.php/fm/rt/printerFriendly/5381/4110
doi: http://dx.doi.org/10.5210/fm.v19i9.5381