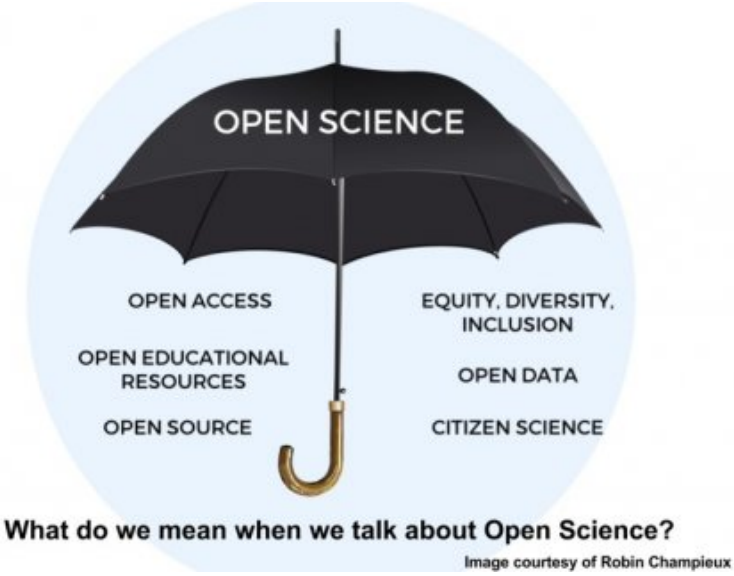
Sharing is caring, but is privacy theft?

February 1 2018



Credit: PLOS Blogs

Open Science (OS) is a movement toward increased sharing among scientists of their data, their materials, their computer code, their papers, and their peer reviews. The ultimate goal of this movement is to boost collaborative progress and bring greater transparency. Scientists might more rapidly build on results of others and help each other spot errors.

In the wake of the replication crisis in science and medicine, OS practices are widely seen as a path toward greater robustness and reliability of science. With full availability of materials, data, and code,

researchers could easily reproduce results reported by others. They could just download the data and code and rerun the analyses. They could then make adjustments to the analyses or replicate the experiment to find out how reliable, robust, and generalizable the results are.

Additional benefits that can result from OS practices include establishing more efficient work pipelines and increased exchange and collaboration with the scientific community. For example, in the field of human brain mapping, developments such as OpenfMRI, NeuroVault, and the Human Connectome Project are prominent examples of successful large-scale open data collaborations.

Another element of OS practices is "preregistration" of studies. Researchers publicly declare the hypothesis and analysis plans before collecting the data. This can help prevent researchers from trying many different analyses, a widespread practice which can inflate the rate of false-positive findings. Like in clinical trials, preregistrations and registered reports have been suggested as new publishing formats to make a clear distinction between a priori declared confirmatory and post-hoc exploratory testing.

Dave Eggers's book "The Circle" presents a satirical vision of the ultimate future social network to connect people and explores the tension between the positive and negative aspects of greater transparency and community cohesion ("Sharing is caring" and "Privacy is theft"). OS, similarly, faces the challenge of finding a balance between transparency and collaboration on the one hand, and privacy and individual freedom on the other.

As we are actively figuring out the balance between transparency and collaboration in research, we thought it was worth reaching out to six of our colleagues who have thought extensively about OS. We hope that additional scientists will weigh in with further insight regarding this

balance not only in human brain mapping, but also in other scientific fields.

Specifically, we asked them: What are the main challenges in moving toward Open Science and how can we meet them? Here are their responses:

Kirstie Whitaker (University of Cambridge, Department of Psychiatry):

"Change is coming. Before we continue, let's define some terms for potential readers: Open Science is an umbrella term that can mean different things to different people. Open access research allows everyone to learn from scientific work (particularly that paid for by the tax payer). Open educational resources mean we don't re-invent the wheel when we teach others about our work. Open source materials are ones that allow you to see inside, and improve, the black box. Open data allows researchers to verify our work, and conduct analyses that could not be carried out by one group alone.

Open Science also means open to everyone. We can use the power of curious non-experts through Citizen Science projects. The Open Neuroimaging Laboratory was a finalist for the Open Science Prize and sought to "lower the barriers for researchers, students, and citizen scientists to help scientific discovery". We can look to other neuroscience projects such as Eye Wire and FoldIt for inspiration in the future.

Finally, Open Science means open for all. Whose voices are not currently represented well in our field of study? Who is not advancing to tenured positions? How do we ensure that researchers in the developing world are able to contribute to our quest to understand the human brain? All of the open practices above facilitate the inclusion of under-represented minorities, but it will require ongoing focus and

consideration to create an equitable community. That's my biggest challenge: addressing my implicit (and explicit) biases to ensure we have bigger, better and more diverse ideas in the future.

I would like to live in a world where helping to advance the boundary of scientific knowledge is rewarded through new findings and by confirming (or not) already published results irrespective of who owns the data."

Lara Boyd (University of British Columbia, Faculty of Medicine):

"Open science will truly transform scientific enquiry. It will enable the assemblage of large data sets mined from multiple sources, from different populations, labs, and countries. This type of big data is essential to address complex questions that no single lab or sole study can tackle. Open science also de-mystifies the peer review process, making feedback transparent and increasing accountability across the entire research community (i.e., authors, reviewers and consumers).

All of my work is published as open source and our data are always available to anyone; this is an ethical decision for me. My research is funded by public sources. Thus, the tax-payers of Canada and the US have already "bought" this knowledge. It is only right to make it free for them to consume the information they already paid for.

I do think that a full transformation to open science will be hard for the research community. We need to develop new safeguards that protect the identity of the individuals in our datasets, make sure that we collect common data elements and take the time and make the effort to ensure our data are open accessible. However, once these steps are taken we will all reap the benefit of closer scrutiny of our ideas and broader dissemination of our findings."

Nikolaus Kriegeskorte (Columbia University, Zuckerman Institute):

"Open Science is an unstoppable movement that will transform the entire endeavor of science. It's a great thing! We are reinventing the way we work together in the context of the web. Open Science is open data, open code, open papers (open access), and open evaluation (post publication peer review). In all of these areas, boosting information flow can improve our collective cognition. However, we must not underestimate how fundamentally different science will be after the transformation. The current tectonic shifts will create a different reality, and the transition will be as uncomfortable as it is inspiring. If I tried to tackle all aspects of opening my workflow at once, it would stop me in my tracks as a scientist. I therefore try to ease into new ways of doing science. We have been sharing key data sets and analysis tools for a while. However, sharing all data and code for all papers is still a challenge, as it requires much overhead. We publish all our papers as preprints, ensuring instant and permanent open access. I also share all my peer reviews on my blog as soon as I write them. As we share more, we need to build a new culture of constructive criticism. When someone publishes a bad paper, or makes a mistake in analysis, we need to point this out. However, we also need to acknowledge that errors are normal and prevent occasional missteps from harming a good scientist's career. I hope we can motivate people to make the transition using many carrots and only a few sticks."

Jeanette Mumford (University of Wisconsin-Madison, Waisman Centre):

"Getting started with practicing Open Science can be a bit daunting and I worry there is more talking and less doing out of fear of making mistakes. For example, I'm reminded of a time I was helping somebody run some computer code. We worked out how to properly type out the commands and then had a long discussion about what the outcome might look like with a finger hovering over the return button. I remember

thinking – just hit the button! That being said, I think one of the biggest obstacles is getting started. I hope folks can at least pick one component and try it out. It may not go perfectly, but that's okay in most cases. I've been sharing my simulation code for my last few publications and it wasn't the greatest code on the first try, but folks were able to answer their questions and I'm improving. There are various degrees of openness as well. Just today I was reading a paper and I have a really simple question that I could answer in 20 minutes if I had the numbers of 8 matrices shown in a figure! Not only could I answer my question fast with the summary measures, but it would have been much easier for the authors to share the summaries compared to the raw fMRI data, which would take me a day to analyse. So, I encourage people to dive in and try something. You can build upon and improve your efforts in the future."

Russell Poldrack (Stanford University, Department of Psychology & Stanford Centre for Reproducible Neuroscience):

"It's really heartening to see how Open Science has finally hit the mainstream within the human brain mapping community. Five years ago, our challenge was to convince people that open science was a good idea. Now our main challenge is to convince junior researchers that they can succeed doing open science in an environment where the incentive structures have yet to catch up with our ideals. The competition for academic positions is fierce, and still largely revolves around having lots of publications that report novel, positive results in high-impact journals. Some practices promoted by open science advocates, such as pre-registration of study design and analysis plans, make it much more likely that one will end up with null results (we have already seen this in my lab), or at least with complicated results that make it difficult to tell the kind of tidy story that it takes to publish in these high-impact outlets. Given the current incentives around hiring and promotion, this penalizes researchers who do the right thing. Fixing this is going to require changes from the top, since most hiring and tenure decisions are made

by senior researchers who (in my experience) are generally the least enthusiastic about open science practices. I think that the best thing that junior researchers can do is to highlight their open science activities in their CV, as suggested by Lucina Uddin in a relatively recent Tweet, and also to publicize their activities via social media. This will help make those activities salient to search committees and review panels."

Adam Thomas (National Institute of Mental Health, Laboratory of Brain and Cognition):

"The rapid growth of the open science movement represents a fundamental shift in the way society approaches producing and disseminating discovery and new knowledge. The movement can be grouped into three broad themes: open code, open data, and open papers, each of which faces unique challenges.

The primary challenge for the movement towards open code, methods, and data, is the additional burden researchers face documenting methods and sharing their datasets. This has been recently highlighted by the outcry in response to the NIH's attempt to declare all research a clinical trial. This move would force researchers to adopt much more transparent and open methodologies in all their studies, but it would also add an immense level of paperwork and bureaucracy to scientists who are accustomed to running small, nimble laboratories on budgets that are orders of magnitude smaller than what is typically spent on a clinical trial. Funders should work to provide and support lightweight, low cost means of practicing open science. Both the Open Science Framework and the BIDS standard address aspects of this problem.

The open papers movement has grown from frustration with the small group of editors who control which new studies are worthy of attention as well as frustration with the publishers who are making a considerable profit from the enterprise. Alternative publishing platforms are

advancing rapidly and should be supported. However, we must ensure that the democratization of publication does not blur the lines between legitimate scientific literature and articles produced with predatory, commercial, or political motives."

A main challenge consistent across many responses is the need for a better incentive structure that supports scientists in becoming more open and transparent in their work. We are hopeful that these six experts are representative of a growing movement in the brain mapping community because after all, we are more likely to learn to understand how the brain works as a community that collaborates with an open spirit, sharing our data, code, material and ideas.

This story is republished courtesy of PLOS Blogs: blogs.plos.org.

Provided by Public Library of Science