



Published in final edited form as:

*Neuroinformatics*. 2014 October ; 12(4): 509–511. doi:10.1007/s12021-014-9246-1.

## A Community Spring For Neuroscience Data Sharing

**Giorgio A. Ascoli**

Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA, USA

Giorgio A. Ascoli: [ascoli@gmu.edu](mailto:ascoli@gmu.edu)

I consider 2014 up to now a good year for neuroscience data sharing. This positive feeling is rooted in my personal experience with three interesting events that took place in the short span of five months during the first half of this year. While all three events occurred on US soil and were organized by American institutions, I believe they reflect an international trend. The world-wide reach of the implications for the field should be of interest to the readers of our journal and to all neuroinformatics stakeholders at large.

The first event I attended was a public round-table panel conveyed on February 20–21 by the National Science Foundation (NSF) in their Arlington, VA headquarters. The topic was research robustness and replicability in cognitive science, and the organizer was the Advisory Committee for the Social, Behavioral, and Economic Sciences Directorate<sup>1</sup>. The ~30 panelists included (in no particular order and among many others eminent personalities) the Director of the White House Office of Science Technology Policy, the Editor-in-Chief of Science, the Director of the Center for Scientific Review of the National Institutes of Health (NIH), the directors of the American Psychological Association and of the American Psychological Society, and numerous deans and professors of leading academic institutions. Many journalists and reporters were also in attendance. Part of the two-day meeting focused on statistical analysis and experimental design, but a sizeable portion of the agenda was dedicated to data sharing. The discussion did not tackle the questions of “whether” or “how” data sharing could be beneficial to scientific reproducibility. Instead, the entirety of the time in these sessions was spent on brain-storming on how to promote data sharing. The underlying assumption was that data sharing “automatically” improves research reproducibility, an argument long sustained by many exponents of neuroinformatics that everyone in that room seemed to agree with as a matter of fact.

Although I attended the whole meeting and found nearly all content relevant to neuroinformatics, this is particularly true of the session dedicated to editorial policies for maximizing data sharing. The fact that I was specifically invited as a founding co-editor of *Neuroinformatics* (the only neuroscience journal represented at the meeting) to offer my remarks in this session is a testimony of the high esteem our journal commands as a leader on the theme of data sharing. Panelists ranged widely in opinions, from the conservative belief that it should not be the responsibility of journals to enforce data sharing, to the radical views that data deposition in public databases ought to become a condition for publication. I presented the *Neuroinformatics* policy of simply requesting an “Information Sharing Statement” for every article<sup>2</sup>. Authors can decline not to share their data, tools or resources, but they have to state their explanation publicly. While this has happened in a few

cases (e.g. for reasons of commercial interest or subject anonymity), data from the vast majority of papers published in our journal are indeed shared.

Other panelists and attendees, including a few other editors (mostly of journals in behavioral, cognitive, and biological sciences) were at first expectedly skeptical of such a strange new system (of which they were seemingly unaware). However, as they started perusing online article after article in our journal in real time during my talk, they grew progressively more enthusiastic and most of them appeared totally sold by the end. A few weeks after the panel, at least one of those editors (from one among the longest-standing biology journals) brought a resolution to their editorial board to adopt a similar “Information Sharing Statement” policy. Unfortunately, the proposal was watered down through an extensive discussion by the board and eventually was not voted upon (I know because I am a member of that editorial board). At any rate, I was pleasantly surprised to see an (apparently unsolicited) Information Sharing Statement in an article recently published in a different journal<sup>3</sup>.

The second event I attended was a symposium sponsored by the American Association for the Advancement of Science (AAAS, the publishers of Science magazine) and the Potomac Institute for Policy Studies on March 21 in Washington DC, specifically aimed to advancing data sharing in neuroscience. The symposium (an audio-video recording of which is freely available in its entirety<sup>4</sup>) was again attended by an eminent crowd, including (once more in no particular order) the AAAS president (also former director of the National Institute of Drug Abuse, one of the NIH institutes in the “neuroscience” domain), the former director of NSF, the head of the NIH neuroscience international office, and the Associate Director of the National Library of Medicine (formerly in charge of the US “original” Human Brain Project<sup>5</sup>), among several others. A comprehensive report of the meeting with full transcripts is also available online<sup>6</sup>.

I was honored to be invited to deliver the keynote talk in recognition of the data sharing achievements of NeuroMorpho.Org, which the organizers considered an exemplary success story in neuroscience. The key question I was asked to answer is: how do we convince data owners to deposit their data? I wrote about that challenge on the pages of this journal some eight years ago<sup>7</sup>. Indeed, in the last few releases “success rate” of NeuroMorpho.Org data sharing hovered around 50%. That is, about half of the neuronal reconstructions we identify in the published literature and request from the authors become available for public sharing through this web site. This figure constitutes a considerable climb-up from my initial estimates of approximately 25% in the early days, when only about one quarter of data owners agreed to deposit their morphologies.

Part of the difference is undoubtedly due to the growth of the project and the gradually diminishing diffidence to entrust one’s data to an “unknown” entity. Yet I’m convinced that another major reason is our choice to release the information of whether each and every data set we identify is publicly available or not, as well as the reason if it is not<sup>8</sup> (most commonly, lack of response from the owners). In countless cases, it is only after the third reminder with no reply, when we send an email that relays our regrets and communicates

that their names and paper will be posted online on the “data not available (request denied or unanswered)” list, that we receive a willing response (at long last) to share the data.

The effectiveness of the “information sharing statement” policy pioneered by Neuroinformatics and of the data sharing open disclosure of NeuroMorpho.Org in fostering data sharing reveals an interesting phenomenon. It is not necessary to coercively mandate authors to share data. It is sufficient to let the World know whether they do. With this realization, I found myself ready when I was granted the unique privilege and opportunity to testify before the White House Presidential Commission on Bioethics in their meeting dedicated to the BRAIN initiative. This third event was held in Atlanta, GA, on June 9–10. The Commission tackled many important issues in the course of their public two-day information gathering summit prior to closed-door deliberation, ranging from personal de-identification to cognitive enhancement, from animal experimentation to prospective nanotechnologies. Nevertheless, the agenda also called for one full session on the topic of data sharing (you guessed it, that’s what they deposed me as an expert scientific witness for). The meeting agenda, webcast, transcripts, blog coverage, presentation materials, and press releases are all public<sup>9</sup>.

In their introductory remarks on this issue, none of the commissioners ever doubted or questioned “if” sharing neuroscience data is the ethical thing to do. The only questions in their minds related to mitigating the potential risks related to human subject protection. When I reported that half of the neuronal reconstructions from rodents are *not* shared, and still digital morphologies are considered a success story of data sharing in basic neuroscience research, I saw jaws dropping all around the table in disbelief. Then within seconds outrage ensued. How dare government-funded scientists not deposit their data in public archives? Why do NIH, NSF, and the Department of Defense *not* mandate their grantees to share data? Which peer-reviewers in their right state of mind would recommend acceptance of a manuscript describing data that are not going to be available to the community? I suggested the practical potential of a putative broad policy not quite mandating data sharing, but at least systematically collecting and publicly posting the information of who shares what, who does not, and why. This information (eventually augmented by metrics on data re-use) could then be employed in evaluation of academic performance, promotions and tenures, grant reviews, scientific awards, and philanthropic decisions.

In the years prior to the founding of the Neuroinformatics journal, the proportion of neuroscientists advocating for data sharing was small and perhaps similar to the numbers of those who felt the opposite way (because of cost, quality issues, intellectual properties, academic competition etc.). The majority of our peers simply did not care either way or was otherwise silent and passive. In the early days of this journal we spent time and energy convincing our colleagues to share data, tools, models, and resources<sup>10</sup>, but the scientific community had not yet taken a clear position on this matter. The most remarkable commonality among the three events I described from the first half of this year is not just their emphasis on data sharing: that is rather the selection criterion for inclusion in this narrative. Instead, what I find striking is the uniform and unchallenged starting assumption of all parties involved that sharing data is good and necessary, much like other basic moral

tenets of scientific research such as rigorous integrity, societal utility, and open communication. From those experiences, I learned that society is now overwhelmingly in support of neuroscience data and resource sharing.

With the tide of peer pressure now in favor of data sharing, I am optimistic that journal-mandated Information Sharing Statements and publication of data (non-)sharing records from curated databases will become ever more effective to encourage neuroscientists to make their data, models, and tools publicly available. After all, nobody wants to be considered a “bad person”. Eventually, this position might penetrate everyone’s conscience and we will all start wanting to share data. Perhaps we can learn from a 35-year study on the relationship between self-awareness and transgressive behavior<sup>11</sup> in Halloween “trick-or-treaters.” Children were instructed to only take one piece of candy from a bowl left in the room. Nobody else was present, but in one condition a mirror was behind the bowl. Subjects transgressed significantly less with this simple set-up than in the control (without mirror but otherwise identical) condition. Much like a world-wide reflecting mirror, systematic public disclosure of data sharing is likely to reduce the fraction of neuroscience data and resources that are publicly reported yet not publicly available.

## References

1. <http://www.nsf.gov/dir/index.jsp?org=SBE>
2. Kennedy DN. Barriers to the socialization of information. *Neuroinformatics*. 2004; 2(4):367–368. [PubMed: 15800368]
3. Pyka M, Klatt S, Cheng S. Parametric Anatomical Modeling: A method for modeling the anatomical layout of neurons and their projections. *Front. Neuroanat*. 2014; 8:91. [PubMed: 25309338]
4. [potomacinstitute.org/past-events/2673-neuroscience-and-data-sharing.html](http://potomacinstitute.org/past-events/2673-neuroscience-and-data-sharing.html)
5. De Schutter E, Ascoli GA, Kennedy DN. On the future of the human brain project. *Neuroinformatics*. 2006; 4(2):129–130. [PubMed: 16845164]
6. [potomacinstitute.org/attachments/article/2988/webDataSharing16July2014.pdf](http://potomacinstitute.org/attachments/article/2988/webDataSharing16July2014.pdf)
7. Ascoli GA. The ups and downs of neuroscience shares. *Neuroinformatics*. 2006; 4(3):213–216. [PubMed: 16943627]
8. From [NeuroMorpho.Org](http://NeuroMorpho.Org), click on “Literature Coverage”
9. <http://bioethics.gov/node/2876>
10. Gardner D, Toga AW, Ascoli GA, Beatty JT, Brinkley JF, Dale AM, Fox PT, Gardner EP, George JS, Goddard N, Harris KM, Herskovits EH, Hines ML, Jacobs GA, Jacobs RE, Jones EG, Kennedy DN, Kimberg DY, Mazziotta JC, Miller PL, Mori S, Mountain DC, Reiss AL, Rosen GD, Rottenberg DA, Shepherd GM, Smalheiser NR, Smith KP, Strachan T, Van Essen DC, Williams RW, Wong ST. Towards effective and rewarding data sharing. *Neuroinformatics*. 2003; 1(3):289–295. [PubMed: 15046250]
11. Beaman AL, Klentz B, Diener E, Svanum S. Self-awareness and transgression in children: Two field studies. *J Pers Soc Psychol*. 1979; 37:1835–1846. [PubMed: 512839]