

Data Sharing and Research on Peer Review: A Call to Action

Flaminio Squazzoni¹, Ana Marusic², Marco Seeber³, Bahar Menhamei⁴, Michael Willis⁵, Phil Hurst⁶, Aliaksandr Birukou⁷, Tony Ross-Hellauer⁹, Catriona MacCallum⁹, Rachael Lammey¹⁰, Peter Rodgers¹¹, Jason Hoyt¹², Duncan Nicholas¹³, Lynsey Haire¹⁴, Pierpaolo Dondio¹⁵, Hollydrawn Murray¹⁶, Iratxe Puebla¹⁷, Virginia Dignum¹⁸, Tiago Barros¹⁹, Stephen Cowley²⁰, Petra Ahrweiler²¹, Giangiacomo Bravo²², Kalpana Shankar²³, Harry J.J. Blom²⁴, Joris Van Rossum²⁵, Giorgio Pedrazzi²⁶, Federico Bianchi²⁶, Francisco Grimaldo²⁷

¹University of Milan, Italy [corresponding author: flaminio.squazzoni@unimi.it], ²University of Split School of Medicine, Croatia, ³Ghent University, Belgium, ⁴STM Journals, Elsevier, The Netherlands, ⁵John Wiley & Sons, Oxford, United Kingdom, ⁶Royal Society, London, United Kingdom, ⁷Springer Nature, Heidelberg, Germany, ⁸Know-Center, Graz, Austria, ⁹Hindawi, London, United Kingdom, ¹⁰Crossref, Oxford, United Kingdom, ¹¹eLife, Cambridge, United Kingdom, ¹²PeerJ, London, United Kingdom, ¹³DN Journal Publishing Services Ltd, Brighton, United Kingdom, ¹⁴Taylor & Francis, Abingdon, United Kingdom, ¹⁵School of Computer Science, Technological University of Dublin, Ireland, ¹⁶F1000, London, United Kingdom, ¹⁷PLOS ONE, Cambridge, United Kingdom, ¹⁸Umeå University, Sweden, ¹⁹Publons, London, United Kingdom, ²⁰University of Southern Denmark, Denmark, ²¹University of Mainz, Germany, ²²Linnaeus University Vaxjo, Sweden, ²³University College Dublin, Ireland, ²⁴Springer Nature, New York, ²⁵Digital Science, London, United Kingdom, ²⁶University of Brescia, Italy, ²⁶University of Valencia, Spain

Abstract. While recent surveys show that most stakeholders recognise the importance of peer review to the publication process, there is a lack of systematic research on the topic. In a period of hyper-competition for resources, with perverse incentives that lead to academic capitalism and a “publish or perish” mentality, the lack of robust and cumulative research on approaches, models and practices of peer review can slow down efforts towards fostering research integrity and the credibility of scholarly communication. A major challenge in studying peer review systematically is the lack of available data. While data sharing in scientific research has made relevant progress in certain fields, the lack of infrastructures to promote the sharing of peer review data among publishers, journals and academic scholars, the challenges posed by privacy and data protection legislation, and the perceived lack of incentives for publishers, learned societies and journals to share data, have all hampered efforts in this important domain. While public authorities, learned societies and publishers may face different priorities, incentives and obstacles regarding data sharing, the time has come to call to action all stakeholders who play a part in this field. In this paper, we argue that an infrastructure for data sharing is needed to stimulate independent, collaborative, public research on peer review and we suggest measures and initiatives to set up a collaborative effort towards this goal.

Keywords: Data sharing; infrastructure; journals; peer review; scholarly communication.

Peer review is the bedrock of scholarly communication. Peer review subjects manuscripts to scrutiny by independent experts and contributes to the credibility of the scientific literature by improving papers and keeping flawed work out of journals (Tennant et al. 2017). Peer review is key to establishing domains of value and worth in science while preserving the autonomy and self-regulation of the scientific community from external and contingent economic and political factors. It has a central role in establishing and maintaining reputation within the scientific community and protecting socially shared values of genuine curiosity, methodological rigour and empirical argumentation.

While recent surveys show that most stakeholders recognise the importance of peer review to the publication process (Mulligan et al. 2013; Publons 2018), there is a lack of systematic research on the topic (Lee & Mohler 2017; Squazzoni, Grimaldo and Marusic 2018; Batagelj, Ferligoj and Squazzoni 2017). Research on the pros and cons of certain peer review models or practices, as well as on definitional issues concerning the quality and value of the process, has been mostly case-based, often only anecdotal and sometimes even ideological (Cowley 2015; Squazzoni, Brezis & Marusic 2017). This has even led some observers to question the “experimental” validity of peer review (Smith 2006) and call for making peer review “more scientific” (Rennie 2018). Studies on peer review are rarely at the top of the research portfolio of academic scholars, who mostly perform this type of research in their spare time (Malicki et al. 2014).

In a period of hyper-competition for resources, with perverse incentives that lead to academic capitalism and a “publish or perish” mentality, the lack of robust and cumulative research on approaches, models and practices of peer review can slow down efforts towards fostering research integrity and the credibility of scholarly communication. This can have detrimental implications for the legitimacy of the so-called “social contract” that implicitly regulates the relationship between science, public bodies, government and society (Guston & Kenniston 1994).

A major challenge in studying peer review systematically is the lack of available data (Helmer et al. 2017; Ross-Helleauer 2017; Squazzoni, Grimaldo and Marusic 2017). While data sharing in scientific research has made relevant progress in certain fields, the lack of infrastructures to promote the sharing of peer review data among publishers, journals and academic scholars, the challenges posed by privacy and data protection legislation, and the perceived lack of incentives for publishers, learned societies and journals to share data, have all hampered efforts in this important domain. A recent quantitative analysis looking at the last 50 years showed that research on peer review has recently increased but is still fragmented and rarely interdisciplinary (Grimaldo, Marusic and Squazzoni 2018). For any such field of research to be developed, it is imperative that there be systematic and regular access to a large and representative set of data. This is the only way to achieve robust insights on how to improve the value and rigour of peer review (Bravo et al. 2019).

This is the challenge that a large network of scientists and professionals who partnered around a COST Action funded by the European Commission in 2014-2018 (PEERE: www.peere.org) - including most of the authors of this paper - have started to tackle. By fostering community building between publishers and scholars and piloting data-sharing initiatives (Squazzoni, Grimaldo & Marusic 2017), we realised that collective priority had to be given to data sharing from scholarly journals in order to stimulate research on peer review. The recent important efforts to increase the openness of science and research output made by public agencies and learned societies in both EU and the US, need to be accompanied by initiatives to promote independent research on peer review based on access to internal journal data. In our opinion, creating conditions to release internal journal data to independent research teams is an important approach towards transparency and accountability and also a means to establish a more systematic view on editorial practices in scholarly journals.

While public authorities, learned societies and publishers may face different priorities, incentives and obstacles regarding data sharing, the time has come to call to action all stakeholders who play a part in this field. First, we need to establish an **infrastructure for data sharing** to stimulate independent, collaborative, public research on peer review. By collecting and systematizing data on peer review from various journals and publishers on a large scale and making these available for research, such infrastructure could take research on peer review to the next level and stimulate interdisciplinary, trans-domain and cross-methodological analyses. Taking into account the need to protect the legitimate interests of each stakeholder, this infrastructure must be built on protocols and standards of data sharing which truly reflect principles of responsible data management, confidentiality and accountability. The PEERE protocol is a milestone in this direction (PEERE 2017), which would require only incremental improvements, including full compliance with the latest EU general data protection regulation (GDPR).

This infrastructure needs to be independent and ‘neutral’, a community public good maintained by academics and professionals to serve their community rather than any private interest. Such an infrastructure would reduce transaction costs associated with the current need for individual agreements on data access and release between journals and academic scholars, while avoiding that only teams with direct connections to journal editors or publishers have access to data. It would help make research on peer review more replicable, credible and transparent, with potential spillover effects on trust and collaboration around other activities, such as the development of common technical standards for data management across publishers. Publishers could reduce reputational risk and costs of data sharing by relying on shared protocols of data anonymization and management. At the same time, they would benefit from systematic research on editorial standards, while demonstrating their

legitimacy as ‘custodians’ of scholarly records and responsible editorial practices, for example, in the context of concerns regarding predatory journals.

The infrastructure would be a key enabling factor to further by-products and innovations, such as: (a) the possibility of standardizing and streamlining training for reviewers and editors, (b) the exploration of common frameworks for reviewer recognition and credit across publishers, (c) the development of certification systems for reviewers, (d) new standards for journal management and (e) an evidence-based framework to promote cultural change in research assessment.

Secondly, we must magnify this effort by **public funding schemes** that stimulate independent, focused research on peer review. Without data, there is no research but without dedicated financial support, research cannot reach a critical mass to improve stakeholders’ practices and inform innovations and reforms. Opening funding schemes to support more research on peer review would also stimulate education and training, while investments would have high long-term returns on the legitimacy and credibility of science. Such schemes would also complement the investment effort already made by many stakeholders on transparency initiatives under the open science framework. Indeed, opening research and journal data without creating an infrastructure for systematic research could cause a learning gap.

One of the first benefits of this effort would be to study peer review quantitatively and comparatively across domains. This is key to providing analysis on a variety of interrelated factors, such as sources of bias (e.g., lack of gender, ethnic, and methodological diversity), economic efficiency of different models, community engagement and long-term sustainability. This led us to address this call to action especially to public authorities, funding agencies and publishers.

We propose that public authorities and funding agencies:

(i) support the **organisation of a forum** open to a group of experts and representatives of publishers to develop a blueprint on data sharing, which builds upon the protocol on data sharing, recently piloted with Elsevier, the Royal Society, Springer-Nature and Wiley (PEERE 2017).

(ii) commission a **proof-of-concept of an infrastructure** for data sharing on peer review, which integrates the workshop’s findings and complies with GDPR while being open to further input by any interested stakeholder;

(iii) support **the establishment of a data sharing infrastructure and governance framework**;

(iv) commit to establishing **funding schemes** with public calls to support research on peer review based on infrastructure data.

We propose that publishers and learned societies:

(i) participate in the forum by **contributing to the blueprint and the proof-of-concept** as part of the mandate of their role as “custodians” of the scholarly record;

(ii) make sincere efforts to **remove cultural and practical obstacles** against data sharing;

(iii) support the establishment of the data sharing infrastructure with **in-kind investments** (e.g., time of IT staff, negotiations with data providers);

(iv) commit to creating an **independent representative body** that participates in the establishment, governance and development of the infrastructure.

Finally, in our opinion, the best way to defend the public credibility of science is promoting a large-scale effort of systematic reflection inside the scientific community on its institutions and practices. The value and rigour of publications, scholarly journals and communication depend on peer review and its capacity of maintaining disinterested, informed and transparent standards of academic judgement. The purposes of this call to action can both stimulate self-reflection among the scientific community and innovations and reforms of practices.

Acknowledgments

This paper is the output of a meeting held at the Stables, Springer-Nature in London on 15-16 October 2018. We thank Elisa De Ranieri for hosting the meeting. The meeting was financially supported by the COST Action TD1306 PEERE (www.peere.org).

References

- Batagelj, V., Ferligoj, A. & Squazzoni, F. (2017) The emergence of a field: a network analysis on peer review. *Scientometrics*, 113(1), 503-532.
- Bravo, G., Grimaldo, F., Lopez-Iniesta, E., Mehmani, B., & Squazzoni, F. (2019) The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, 10: 322.
- Cowley, S. (2015) How peer review constraints cognition: on the frontline in the knowledge sector. *Frontiers in Psychology*, 6, 1706.
- Grimaldo, F., Marusic, A. & Squazzoni, F. (2018) Fragments of peer review: A quantitative analysis of the literature (1969-2015). *PLoS ONE*, 13(2), e0193148.
- Guston, D. H. & Keniston, K. (1994) Introduction: The social contract for science. In (Eds.), *The Fragile Contract*. Cambridge, London: MIT Press, pp. 1-41.
- Helmer, M., Schottdorf, M., Neef, A. & Battaglia, D. (2017) Gender bias in scholarly peer review. *eLife*, 6: e21718.
- Lee C. J. & Moher, D. (2017) Promote scientific integrity via journal peer review data. *Science*, 357(6368), 256-257.
- Malički, M., von Elm, E., Marušić, A. (2014) Study design, publication outcome, and funding of research presented at international congresses on peer review and biomedical publication. *JAMA*. 311(10); 1065-1067.
- Mulligan, A., Hall, L., & Raphael, E. (2013). Peer review in a changing world: An international study measuring the attitudes of researchers. *Journal of the American Society for Information Science and Technology*, 64(1), 132-161.
- PEERE (2017) The PEERE policy on data sharing on peer review. Available at: <http://www.peere.org/wp-content/uploads/2017/03/PEEREDataSharingProtocol.pdf>.
- Publons (2018). 2018 Global State of Peer Review. Clarivate Analytics.
- Rennie, D. (2018). Let's make peer review scientific. *Nature*, 535, 31-33.
- Ross-Hellauer, T. (2017) What is open peer review. *F1000Research*, 6: 588.
- Smith, R. (2006). Peer review. A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, 99, 759-760.
- Squazzoni, F., Brezis, E. & Marusic, A. (2017) Scientometrics of peer review. *Scientometrics*, 113(1), 501-502.
- Squazzoni, F., Grimaldo, F. & Marusic, A. (2018) Publishing: Journals could share peer-review data. *Nature*, 546(7658), 352.
- Tennant, J. et al. (2017) A multi-disciplinary perspective on emergent and future innovations in peer review. *F1000Research*, 6: 1151.