

(Re?)Building trust in research integrity

Sabina Alam^a, Rachel Burley^b, Chris Graf^c, Alice Meadows^{d,*}, Gabriela Mejias^e and Damian Pattinson^f

^a*Publishing Ethics and Research Integrity, Taylor and Francis Group, Journals, UK*

^b*Chief Publications Officer, American Physical Society, USA*

^c*Research Integrity Director, Springer Nature, UK*

^d*Cofounder, MoreBrains Cooperative, USA*

^e*Community & Program Manager, DataCite, Germany*

^f*Executive Director, eLife, UK*

Abstract. This article is based on a session with the same title from the 2023 APE (Academic Publishing in Europe) conference, in which the authors discussed the challenges and opportunities for the scholarly communications community to improve trust in the quality of research it funds, publishes, and uses. They provide here a brief summary of the discussion, including: how much trouble is research integrity in? Who should be responsible for addressing it? Do we need less technology or more? Is better peer review the answer; if so, what does that look like? What other aspects of the research process should we be considering? Does the existing research infrastructure support these changes, or is further investment needed?

Keywords: Research integrity, persistent identifiers, community, metadata, trust, transparency

Article

The volume of research outputs being published globally continues to increase [1] and, partly because of the ongoing pressure to ‘publish or perish’, this growth has in recent years been accompanied by the development of predatory publishers, paper mills, citation rings, and other rogue players. As a result, we are seeing a significant increase in reports of fraud and, subsequently, retractions — sometimes en masse — from 45 per month in 2010 to around 300 per month in 2021 [2]. This all, in turn, contributes to a declining trust in the integrity of research overall, on the part of both the general public [3] and researchers themselves [4].

According to the World Conferences on Research Integrity Foundation (WCRIF), “Research integrity refers to the principles and standards that have the purpose to ensure validity and trustworthiness of research. Research integrity is vital to realise the societal value and benefits of research. The consistent and coherent adherence to the principles of research integrity, such as honesty, accountability, professional courtesy, fairness, and good stewardship are the hallmarks of research integrity” [5].

Based on this definition, the authors all agree that there is, indeed, a problem with trust in research integrity, but our perceptions of this problem — and what is causing it — differ slightly. For some,

*Corresponding author: Alice Meadows, MoreBrains Cooperative, 21 Kent Street, Brookline, MA 02445, USA. Tel.: +1 857 559 3337; E-mail: alice@morebrains.coop.

technology is the biggest challenge (while also a huge opportunity); others see the lack of public understanding of the scientific process as the major issue; and the lack of consistent, high-quality training and reporting, including around publishing ethics, was also flagged. However, we all agree that, while research integrity is critical for publishers in their role as gatekeepers of research reporting, the whole research community must collectively share responsibility for improving and rebuilding trust in it. Publishers are already collaborating with each other on this, for example, via the International STM Association's newly launched Integrity Hub [6]. Publishers are also working with institutions and vendors, for example, through broader community initiatives such as the National Information Standards Organization (NISO) CREC (Communication of Retractions, Removals, and Expressions of Concern) Working Group [7]. However, collaboration between all players in the research space — funders, institutions, publishers, scholarly societies, infrastructure/service providers, and researchers themselves — is required to effect widespread and lasting change, with the need to find effective ways for publishers to engage with funders and institutions, in particular, being flagged as critical. Unfortunately, opportunities for all stakeholder groups to directly interact with each other are currently few and far between, which can make building trust and collaborating to find solutions challenging.

The necessity for a whole-community approach to address the challenges around research integrity underpinned the other themes that ran through the discussion. These included: a call for more transparency and openness in research processes and publishing; ensuring that all research entities and outputs are connected, and that the links between them have been validated; the threat posed by current perverse incentives that encourage quantity over quality of research publications; and better education and training — for both researchers and the wider community, including the general public.

Persistent identifiers (PIDs), together with their associated metadata, were highlighted by everyone as a way of increasing transparency. They enable trusted connections — including provenance information — between researchers, organizations, grants, outputs, and more [8]. The major PID organizations are themselves typically multi-stakeholder and community-led. However, as Gabriela Mejias pointed out, the benefits of PIDs will not be fully realized until there is widespread global adoption; that goal is still some way off, and will only be achieved if all stakeholders work together to implement PIDs across all research systems. For example, several speakers noted that better validation of researcher identity is also essential to improving transparency and, ultimately, trust in research. ORCID identifiers are an obvious solution; however, although there is continued growth in both the number of IDs registered and integrations [9], they are not yet ubiquitous enough nor widely implemented well enough. Requiring researchers to sign in to their ORCID record whenever they use it, rather than allowing them to key it in manually, would help ensure that validation, and including that information in the associated metadata would enable it to flow seamlessly through to other systems. Similarly, DOIs can help identify research outputs and connect them with researchers, research organizations and funders, other outputs and beyond. But implementing this in all the many systems and platforms in use across the research life cycle will, again, require a cross-community, multi-stakeholder approach.

As Sabina Alam pointed out, trust in research has to be earned, and transparency is one way of doing this. She advocated for a holistic approach, powered by technology that links all elements of a researcher's work — the grant(s) they are awarded, the protocols, reagents and equipment they use, the datasets they develop, their final publication(s), and more. Expanding the connections between different research outputs and elements in this way, as well as validating them, would again be facilitated by the widespread and consistent use of PIDs and their metadata, including the Research Activity Identifier (RAiD). This new PID, developed by the Australian Research Data Commons (ARDC), groups all activities and elements of a research project together [10] — a great example of a community solution to a community challenge.

Another area in need of a shared, cross-community approach is the challenge of perverse incentives for researchers, who are currently often encouraged to publish more, not better. This is driving much of the demand for the fake or fraudulent papers and/or images created by paper mills, predatory publishers, and other bad players in scholarly communication which, in turn, are contributing to the declining trust in research integrity. Damian Pattinson used a football (soccer) analogy to make the point that, just as bad behavior (taking a dive in a football game) is often rewarded more highly (by getting a free kick or a penalty) than the perceived misdemeanor, so getting widely published — by any means — is often rewarded in academia.

Interestingly, an audience member later extended the analogy by noting that, with the introduction of technology to spot real versus fake foul plays in football, players are now less likely to take a dive. Likewise, the authors believe that technology improvements must be part of the solution to spotting and removing fraudulent papers more quickly. For example, Rachel Burley suggested that tools should be used to identify issues upfront, before a paper is even submitted, in order to free up time for reviewers to evaluate the quality of the research itself. At the same time, getting the right balance between technological and human intervention was seen by everyone as an ongoing challenge. In addition, as Chris Graf noted, any solution cannot be purely technical: social changes are also needed. This must include better education and training — another theme that cropped up throughout our discussion — and not just for authors and reviewers, but also for the wider community, many of whom have little to no understanding of how the scientific process actually works. This includes policy-makers, who are increasingly starting to engage with some of these issues [11].

A question from one of the audience members is, perhaps, a good note to end on as a very concrete way of looking at the impact of research integrity issues. They asked whether there is any evidence of publishers' revenue being affected by retractions. The conference took place about three months after Wiley paused production of Hindawi journal special issues and before they retracted 1,200 Hindawi articles as a result of widespread fraud, but we now know that they took a significant financial hit as a result of this - \$9m per their third quarter earnings statement [12]. Money talks and, as a result, we can expect to see more attention paid to addressing the challenges we face around research integrity. So, while trust in research integrity — and, by extension, trust in science — has been shaken, there are many opportunities (and reasons) to address this: ethical, financial, technical, and more. In all cases, taking a cross-community approach, involving all stakeholders in the research ecosystem, will be essential.

References

- [1] L. Bornmann, R. Haunschild and R. Mutz, Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases, *Humanit Soc Sci Commun* **8** (2021), 224. doi:10.1057/s41599-021-00903-w.
- [2] I. Oransky, Nearing 5,000 retractions: A review of 2022. Retraction Watch. <https://retractionwatch.com/2022/12/27/nearing-5000-retractions-a-review-of-2022/> (December 17, 2022).
- [3] B. Kennedy, A. Tyson and C. Funk, Americans' Trust in Scientists, Other Groups Declines. <https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/> (February 22, 2022).
- [4] Trust in Research Results: Researcher Survey Results, Elsevier. https://www.elsevier.com/_data/assets/pdf_file/0011/908435/Trust_evidence_report_summary_Final.pdf (June 2019).
- [5] World Conferences on Research Integrity (home page). <https://www.wcrif.org/foundation/mission>.
- [6] International STM Association website. <https://www.stm-assoc.org/stm-integrity-hub/>.
- [7] CREC (Communication of Retractions, Removals, and Expressions of Concern) Working Group, NISO website. <https://niso.org/standards-committees/crec>.

- [8] A. Meadows, L. Haak and J. Brown, Persistent identifiers: The building blocks of the research information infrastructure, *UKSG Insights* **32** (2019), 9. doi:[10.1629/uksg.457](https://doi.org/10.1629/uksg.457).
- [9] ORCID; C. Shillum, J. Petro, T. Demeranville, I. Wijnbergen, W. Simpson, ORCID 2022 Annual Report, ORCID. Online resource. doi:[10.23640/07243.22250740.v1](https://doi.org/10.23640/07243.22250740.v1) (2023).
- [10] Australian Research Data Commons website. <https://ardc.edu.au/services/ardc-identifier-services/raid-research-activity-identifier-service/>.
- [11] A. Meadows, Why PID strategies are having a moment and why you should care, The Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2023/01/25/why-pid-strategies-are-having-a-moment-and-why-you-should-care/> (January 25, 2023).
- [12] Wiley Reports Third Quarter Fiscal Year 2023 Results, Wiley website (press release). <https://newsroom.wiley.com/press-releases/press-release-details/2023/Wiley-Reports-Third-Quarter-Fiscal-Year-2023-Results/default.aspx> (March 9, 2023).