Beyond Open Access - The Changing Culture of Producing and Disseminating Scientific Knowledge

Heidi Laine
University of Helsinki
+358405139593
heidi.laine@helsinki.fi

Leo Lahti
University of Turku
+358 40 565 5872
leo.lahti@iki.fi

Anne Lehto
Tampere University Library
+358401904262
anne.m.lehto@staff.uta.fi

Samuli Ollila
Aalto University School of Science
+358 503746963
samuli.ollila@aalto.fi

Markus Miettinen
Freie Universität Berlin
+49 30 838 52226
markus.miettinen@iki.fi

ABSTRACT
This is an extended abstract for the workshop "Beyond Open Access - The Changing Culture of Producing and Disseminating Scientific Knowledge", organized by the Open Knowledge Foundation Finland Open Science Working Group at the Academic Mindtrek Conference, 24 September 2015. The workshop organizers felt that the traditional model for disseminating scientific knowledge, through pay-walled peer-reviewed journal articles, has become both inefficient and unfair, and that the Open Access to publishing movement solves only part of the problem. The workshop took the four main functions of the academic article as a starting point for the discussion; a) dissemination of scientific knowledge, b) a forum for academic discussion, c) maintaining and monitoring the quality of research and d) determining academic merit. The aim was to reflect on alternative ways of meeting those functions, such that would support the principles of open science (transparency, accessibility, integrity). These alternatives included open research processes, altmetrics and open peer review. The effects of open practices on research integrity were also discussed. Recordings of the workshop presentations are available for viewing at bit.ly/beyond-open-access.

CCS Concepts
• **Social and Professional Topics**

Keywords
Altmetrics; Metrics; Open peer review; Open science; Research Integrity.

1. INTRODUCTION
The global scientific community is in the midst of a cultural change. The traditional model for disseminating scientific knowledge, through peer-reviewed journal articles, has become both inefficient and unfair.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
AcademicMindTrek'15, September 22-24, 2015, Tampere, Finland
ACM 978-1-4503-3948-3/15/09.
http://dx.doi.org/10.1145/2818187.2818282

While digitalization has caused the added value provided by academic journals to become less significant, the grip of the big publishing houses on scientific content and through that on taxpayer euros has not loosened. Some see a solution in the Open Access to Publishing movement. Others argue that it is a compromise at best and at worst merely a new source of revenue for publishers. Many of the challenges remain the same regardless the financial logic: slow publishing processes, opaque peer-review practices, no incentives and mechanisms for a real-time dialogue between the publishing researcher and his/her audience, not to mention the impact factor, often accused of being the root for the prevailing "publish or perish" mentality in academia.

It seems that many of the roadblocks on the way to genuinely open science have to do with the article: the need to keep quiet about new discoveries until they have been published, so that the novelty value doesn't dissipate and make the research uninteresting for journals; or the practice of withholding research data in case it could be one day used as the basis for a new article; not to mention the legal obstacles in the way of data mining. One can't help but wonder, whether there is a revolution lurking behind the reform that is Open Access, a way that could solve all the above mentioned issues and make scientific knowledge truly the property of everyone, as it should be.

The goal of the workshop was to discuss and showcase the possibilities offered to science by the digital environment. The workshop's core question was how to do science in a way that is truly open, not just at the end of the process and with a significant delay, as in many of the Open Access models. [1] This question was be approached from angles rising from the main functions of the academic journal: a) dissemination of scientific knowledge, b) a forum for academic discussion, c) maintaining and monitoring the quality of research and d) determining academic merit. In this extended abstract we will describe alternative ways of serving these functions, such that are more open and resource efficient than the current ones, while also allowing more ownership of the process to the research community.

2. OPENING THE RESEARCH PROCESS: THE CASE OF THE NMRLIPIDS BLOG
Since 2013, ultra-open computational biophysics research has been conducted at *nmrlipids.blogspot.fi*, a public blog through which anyone can both follow and participate in the research process. To credit the participants, traditional peer reviewed articles are published; coauthorship is offered to everyone who participates via

the blog. Below I share our experiences on opening the research process; I do this by going through **eight problems we initially expected to meet**, but (so far) did not.

1: No one participates? NMRlipids gets 1000 pageviews per month, and to date our 27 participants across the globe have made over 350 scientific contributions through the blog.

2: The project soon grinds to a halt as people lose interest. A key feature in the continued success of NMRlipids has been that the scientific research taking place is actively and passionately lead by Dr. Samuli Ollila. This seems to be crucial for the success of an ultra-open research project in general: It must have a leader or leaders, who at all times have an overview of the project, keep it focused, and by their own efforts ensure that the project moves forward. It seems that we humans are happy to help to accelerate a process that we see is already moving, but that we are considerably less keen to start setting in motion a process that appears to have stopped.

3: No tools for open collaboration? Just the simple blog as a discussion forum was ok, but not optimal, as it got difficult to follow as the number of contributions grew. We have shifted a significant amount of the collaboration to GitHub (github.com), and started using Zenodo (zenodo.org) to share (and get DOIs for) data.

4: Trolls, spammers, or people babbling nonsense drown the scientific content? All the discussions have been to the point. Possibly because 1) we require the contributors to give their name and affiliation and 2) the technical nature of the discussions, making it difficult for non-specialists to participate.

5: Personal conflicts escalate into useless fighting? This has not happened, but this could be partly because we have, unfortunately, not received very many critical comments in general. Indeed, it is seems likely that those scientists who do not agree with our findings are simply choosing not to participate in NMRlipids; this naturally cuts down the number of personal conflicts, but also limits the efficiency of the scientific process.

6: Powerful scientists get angry and start to sabotage our careers? So far the feedback we have been receiving has been ranging from 'you are crazy' to 'you are brilliant'. No one has given us the impression that they would think us as a threat and thus someone who should be smoked out of science.

7: Someone steals our results? This is unlikely, as all our data is public with well-documented publication dates. Thus should someone try to publish our findings under their own names, our priority can be easily demonstrated.

8: Peer reviewed journals will not publish results that are already public? Many publishers, such as American Chemical Society (acs.org) state that their journals do not publish findings that have been made previously public. To test this statement in practise, we sent our first manuscript to an ACS journal, stating explicitly in the cover letter that we are summarizing work based on open reseach. This turned out not to be a problem, but manuscript was sent to peer review by the editor.

To sum up, the experiences gained on open research during the NMRlipids project have been extremely positive, and none of the major fears we had before starting the project have actualized. Quite the contrary, the open research approach has proven to be an extremely fruitful as well as rewarding way to do research.

The text should be in two 8.45 cm (3.33") columns with a .83 cm (.33") gutter.

3. ALTERNATIVE WAYS OF DETERMINING ACADEMIC MERIT

3.1 Citation based metrics

Traditionally citations have been considered as a proxy for scientific impact. The interest for measuring impact has risen from the need of universities to show the return on investment, caused by the spread of New Public Management ideals, and as a result of funding models based on publication output. The most influential citation based impact tools are the Impact Factor and Scopus, which calculate the impact of journals. Hirsch index, or the H-index was created in 2005 to determine the impact of individuals. It takes into account the number of articles and citations. Since it is cumulative, it favors senior scientists over early career researchers. Finland has created its own tool called the Publication Forum in a quest to find a more balanced solution that would take the humanities and social sciences better into account. Instead of citations it uses expert panels for determining impact. Some of the challenges of citation based metrics, which the Publication Forum panels also rely on to a certain extent, are f. e. that they give every article equal value and that they can be manipulated. [2][3]

3.2 Altmetrics

According to some, societies are moving from monetary based systems to attention based ones. Altmetrics are a solution for taking attention into account in measuring scientific impact. In addition to the articles they can be applied to both journals and authors. They track activities in web based environments, usually social media. The pioneer of the field is the Public Library of Science (PLOS). The field is developing fast, with new players emerging frequently. [4] Altmetrics.com, Plum Analytics by EBSCO and ImpactStory. Most of the new metrics tools are commercial, just like some of the more traditional citation based ones, and charge fees from institutions using their services. They broaden the concept of academic impact and facilitate rewarding researchers' societal engagement, but they are not advocates of open science per se.

3.3 How to increase the research impact?

It is likely that the use of altmetrics will increase and the altmetrics tools will themselves increase their impact. From an individual researcher's point of view they are certainly something to be aware of. For a researcher who is looking to get the maximum career benefit from the changing landscape of scientific impact the following advice could be given: a) create an identification profile, e. g., ORCID, b) co-author articles, write review articles, since they get cited the most, and utilize Open Access publishing channels when possible, c) make sure that multiple channels are used to disseminate information about your publications, and c) discuss and share your work in social media networking sites.[5] At the same time it is good to keep in mind that it is hard to predict which of the emerging applications and services will be relevant in the future.

4. OPEN PEER REVIEW

4.1 New technical tools and opportunities

Peer review is a cornerstone of the scientific process and crucial for controlling the quality of published research. In standard peer review, authors submit a manuscript to a journal, the editor invites reviewers to give their feedback and recommendations on the manuscript, and the final decision to publish is contingent on the reviewer feedback. Traditionally, the peer review process has taken

place as closed correspondence between the author, the editor, and the reviewers. This system was established for print media during the 18th century when print space was remarkably scarce, and publishing details of the review process would have come with remarkable costs.

The ultimate aim of peer review is to support the scientific process: to control the quality of publications, and to support the authors in improving their work. The revolution in online publishing and modern communication technologies have provided new technical tools and opportunities to revise traditional peer review practices and potentially take advantage of increased transparency to support these objectives of peer review.

It has become technically possible to publish the full peer review history with relatively low added costs. Journals such as F1000 Research and PeerJ have pioneered this, proposing best practices and technical platforms to implement open peer review. Open peer review covers many aspects, including publishing the reviewer identities, open access to the full review history, possibility for external parties such as other researchers or the general public to contribute in the peer review process, and making intermediate versions of the article available to the public. Widely accepted definition of open peer review does not exist at the moment, and open peer review can refer to any combination of these different elements. The openness could be optional or mandatory.

4.2 Advantages of increased openness in reviewing

It has been observed that reviewers who provide negative feedback are less willing to publish their identity. Since negative feedback may potentially lead to personal conflicts, the opportunity to hide reviewer identity might hence be important to secure the quality and honesty of the reviewer feedback. Publishing the reviewer identity might, on the other hand, lead to improved quality of the reviews. At the moment the journals experimenting with peer review are leaving the decision on publishing their identities to the reviewers. Open access to the review history provides a full record of academic feedback and subsequent improvements to the manuscript during the editorial process. This can be useful for other researchers and the public to identify potential shortcomings of the work, and how these have been taken into account. This facilitates transparency of the scientific process and can provide valuable extra information that supplements the actual manuscript.

When intermediate versions of the article are published during the review process, also the wider research community and the general public have the opportunity to provide feedback, for instance via personal communication, email, commenting options provided at the online publishing platform, or even through media. This can further support a key goal of peer review, which is to help the authors to improve their original work. Another advantage of publishing intermediate versions is that the new scientific information can be made immediately available upon submission, whereas the full quality control provided by peer review is at the same time taking place as a post-publication process. This can further speed up distribution and expansion of scientific knowledge. The public preprint archives, such as the arXiv, have already facilitated such communication processes for a longer time in certain fields, such as particle physics, but now the new online publishing platforms are catching up and providing improved technical tools to support such process.

4.3 Support for open data and open source

Finally, data and source code have an increasing role in scientific publishing. Open availability of these key research resources is also an element of peer review (see the Open Science Peer Review Oath). [6] When the community of researchers and other parties have the opportunity to review and test the code, supported by appropriate technical platforms such as Github (for source code) or Data Dryad (for data), the potential bugs or inaccuracies can be spotted more efficiently. This emphasizes the role of peer review as a continuous process of quality control and improvements, ideas that are now spreading from the domain of open source code to the domain of open peer review as the technical solutions are becoming available and provided by an increasing number of online publishing platforms.

5. OPEN PRACTICES ADVANCE RESEARCH INTEGRITY

5.1 What is responsible conduct of research?

A person is said to possess the virtue of integrity if his or her actions are based upon an internally consistent framework of principles. Responsible conduct of research (RCR) is a set of principles and values that all fields of research are expected to follow, and thus a precondition for ethically acceptable, reliable and credible research.

The flip side of RCR is research misconduct and research fraud, usually categorized into fabrication, falsification and plagiarism (FFP). [7] There are a number of national and international guidelines and statements defining RCR. The issues covered in them include usually at least research methods, data management, reporting of results, giving credit to others and dealing with liabilities. They do not use the term open science, but are in line with its core values (i.e. they generally promote responsible practices, transparent and reproducible research processes and free access to research results). [8]

5.2 Widespread bad practices

Breaches against responsible conduct of research are a serious problem whether they are common or not. Research misconduct scandals jeopardize the public trust on science, and there have been quite a few during recent years. There is reason to believe that they are not individual instances, but a symptom of a wider-spread culture of cutting corners. A meta-analysis of surveys on research practices found out that up to 2% of respondents had resorted to scientific fraud and 33,7% to other questionable means, often referred to as the “grey area” between good and bad research conduct.[9]

There is very little direct evidence on the amount of misconduct taking place. One indicator is the number of retracted articles. A review of all of the some two thousand retracted articles indexed by PubMed until May 2012 revealed that in only 21,7% of the cases the retraction was due to error, while misconduct was to blame in 67,4% of retractions.[10] It is hard to give a definite number on the yearly amount of retracted articles. A feature article published in Nature gives an estimate of 400 for the year 2011 and the number has likely risen since. [11]

5.3 Open data management key in misconduct prevention

It can be argued that the current model of publishing research results both rewards cheating and makes recognizing fraudulent science difficult. Top journals have been blamed for choosing articles based on their newsworthiness rather than academic

excellence. Nobel laureate Randy Schekmann called in 2013 for a boycott against Science, Nature and Cell for this reason.[12] The accusations are backed by a research finding from 2011, showing a strong correlation between a high impact factor and a high amount of retractions.[13] This could indicate that for the most ambitious, "publish or perish" has become to mean "publish in an acclaimed journal or perish", an attitude that can make misconduct more appealing.

Famous fraudster Diederik Stapel, number four on the leaderboard held by the blog Retraction Watch, has compared himself to a junkie searching for a bigger and better high.[14] He continued fabricating results because he was rewarded for it time and time again by getting into top journals and advancing his career. Stapels fraud wasn't particularly cunning and once a more systematic investigation begun the house of cards came down. He was able to get away for so long because he was allowed to collect and manage his data all alone. With an open data policy in place either at institutional or publisher level his frauds would have been easily discovered.

The Royal Dutch Academy reacted to the scandal caused by Stapel (his last position was at the Dutch Tilburg University) by conducting a study on data management practices among Dutch researchers and came to the following conclusion: "Maximum access to data supports pre-eminently scientific methods in which researchers check one another's findings and build critically on one another's work. In recent years, advances in information and communication technology (ICT) have been a major contributing factor in the free movement of data and results." [15]

6. ACKNOWLEDGMENTS

We would like to thank the Academic Mindtrek organizers for providing us the platform, Open Knowledge Foundation Finland and the Open Science and Research Initiative by the Finnish Ministry and Education for financial support and Infocrea and CloudStreet for making live streaming and recording of the workshop possible.

REFERENCES

[1] Bernius, S., Hanauske, M., König, W., Dugall, B. 2009. Open Access Models and their Implications for the Players on the Scientific Publishing Market Reasoning about naming systems. *Economic Analysis & Policy*, Vol. 39 No. 1, (March 2009), 103-115.

[2] Bornman, L. 2015. Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics. *Scientometrics, 103*(3), 1123-1144.

[3] Costas, R. 2015. Do 'altmetrics' correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science & Technology, 66*(10), (2003-2019).

[4] Wouters, P., & Costas, R. 2012. *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*. Utrecht: SURFfoundation (Feb. 2012).

[5] Strasser, C. 2013. Universities can improve academic services through wider recognition of altmetrics and alt-products. *Impact of Social Sciences*. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2013/10/15/universities-can-improve-academic-services-through-altmetrics/

[6] Aleksic J, Alexa A, Attwood TK *et al.* An Open Science Peer Review Oath [version 2; referees: 4 approved, 1 approved with reservations] *F1000Research* 2015, **3**:271. DOI= 10.12688/f1000research.5686.2

[7] *Responsible conduct of research and procedures for handling allegations of misconduct in Finland 2012*. Finnish Advisory Board on Research Integrity (2013), 32-33.

[8] Laine, H. 2015. What is research misconduct and why it matters for open science. The Honest Broker Blog. Retrieved from http://thehonestbrokerblog.org/2015/07/13/what-is-research-misconduct-and-why-it-matters-for-open-science/

[9] Fanelli, D. 2009. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE* 4(5), e5738.

[10] Fang, F. C., Steen, R. G., Casadevall, A. 2012. Misconduct accounts for the majority of retracted scientific publications. *PNAS*, 109, 17028–17033.

[11] Van Noorden, R. 2011. Science publishing: The trouble with retractions. *Nature* 478, 26-28 (2011), DOI=10.1038/478026a

[12] Sheckman, R. 2013. How journals like Nature, Cell and Science are damaging science. *The Guardian* (9 December 2013). Retrieved from http://www.theguardian.com/commentisfree/2013/dec/09/how-journals-nature-science-cell-damage-science

[13] Ferric, C. F., Casadevall, A., Morrison, R. P. 2011. Retracted Science and the Retraction Index. *Infection and Immunity*. vol. 79 no. 10 (Oct. 2011). DOI=10.1128/IAI.05661-11

[14] Bhattacharjee, Y. 2013. The Mind of a Con Man. The New York Times (28 April 2013). Retrieved from http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all&_r=0

[15] *Responsible Research Data Management and the Prevention of Scientific Misconduct*. The Royal Netherlands Academy of Arts and Sciences 2013.