The 2nd SPARC Japan Seminar 2018
Quality Control in the Age of Open Science

Moving Beyond the Journal:
A Changing Role for Publishers, Funders and Institutions

Rebecca Lawrence
(F1000)

Abstract

The traditional approach of publishing new findings in research journals is becoming outdated, costly and unsustainable, and can delay and damage scientific progress. With a growing worldwide shift towards more open access and open science policies, together with significant technological advances, it has now become possible to address many of these problems. In 2013, F1000 launched the world's first open research publishing platform, F1000Research, combining the ability to publish rapidly with functionality to ensure greater transparency, robustness and reproducibility of research. Our approach uses near-immediate publication together with FAIR data sharing, followed by transparent invited peer review and article versioning. It brings control back to the author(s) and aims to remedy many of the problems associated with traditional and increasingly outdated modes of publishing research, to facilitate the transition to more transparent, collaborative and efficient ways of doing research and delivering impact.

Testimony to the rising interest in, and demonstrable benefits of, open research publishing, we are now providing publishing platforms to a large number of high-profile research-funding agencies and research-performing institutions across the world (e.g. Bill & Melinda Gates Foundation, Wellcome Trust, African Academy of Sciences). This approach changes the respective role of publishers, funders and institutions in the ecosystem, and has the chance to finally address many of the well-known problems with the current research and researcher evaluation system.

Rebecca Lawrence

Rebecca Lawrence is Managing Director of F1000, providers of a series of tools and services to support the research community in writing, publishing, discovering and evaluating new scientific findings. She was responsible for the launch of the novel open science publishing platform F1000Research. She has subsequently led the initiative behind the recent launches of Wellcome Open Research, Gates Open Research, and many other funder- and institution-based publishing platforms that aim to start a new trajectory in the way scientific findings and data are communicated and ultimately research and researchers are evaluated.

She is a member of the High-Level Advisory Group for the EC's Open Science Policy Platform, chairing their work on next-generation indicators and their integrated advice: OSPP-REC. She has been a co-Chair of a number working groups focussing on data and peer review, for organisations including the RDA, CASRAI and ORCID. She is also an Advisory Board member for the data policy and standards initiative, FAIRsharing, and for DORA (the San Francisco Declaration on Research Assessment). She has worked in STM publishing for almost 20 years for several publishers including Elsevier where she built and ran the Drug Discovery Group. She originally trained and qualified as a pharmacist, and holds a PhD in Cardiovascular Pharmacology.

Challenges with Existing Publishing System

There are number of different challenges with the existing publication system. The first one is around the fact that a large proportion of new research discoveries are not openly accessible. They are still stuck behind a subscription wall.

Another issue is that there are long delays between when researchers have discovered something and are ready to share it with other researchers and the world and when others get to actually see that research. It can be months to often years, and in fact, there is no good reason and no benefit for authors and for readers and users in such a delay.

Most journals operate an anonymous peer review process which means that you do not know what has happened, who peer-reviewed the article and what their decisions were, and why the editors made the decision they have made. Therefore, it has inherent conflicts and biases built into that which we think are not necessary.

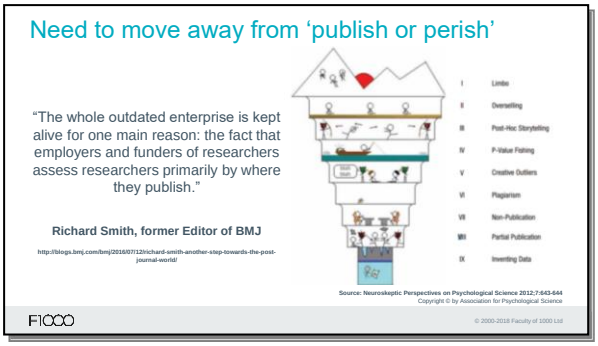
Next, most research that is published is published without underlying data being made available. It is very hard to see how you can review and assess a research discovery without the data. It also makes it very hard to try and reproduce and reuse new findings without that data.

There is also a lot of good research that is never published at all. Some examples are that of negative findings or incremental findings. They are not published because journals do not want to publish them as such research does not bring many citations, and citations affect a journal's impact factor which affects submissions. Therefore, journals are not interested. Estimates suggest almost half of all good quality research is never published.

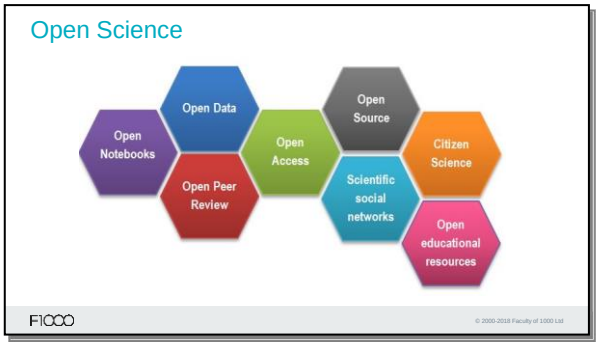
This not only skews our understanding of science because we only see the positive findings and not the negative ones, but also leads to significant research waste due to research being conducted repeatedly and being funded needlessly.

Hence, we need to move away from a "publish or perish" system which causes many issues such as researchers overselling their findings, post-hoc storytelling, p-hacking, and in some cases even plagiarism or fraud in terms of inventing data (Figure 1). As Richard Smith, former Editor of the *British Medical Journal* said, "The whole outdated enterprise is kept alive for one main reason: the fact that employers and funders of researchers assess researchers primarily by where they publish."

Open Science aims to address many of these issues. It comprises of many aspects not just open access, which often gets confused with open science, but also open data, open peer review, open notebooks, and citizen science. All sorts of issues are part of open science (Figure 2).



(Figure 1)



(Figure 2)

Global Shift towards Open Science

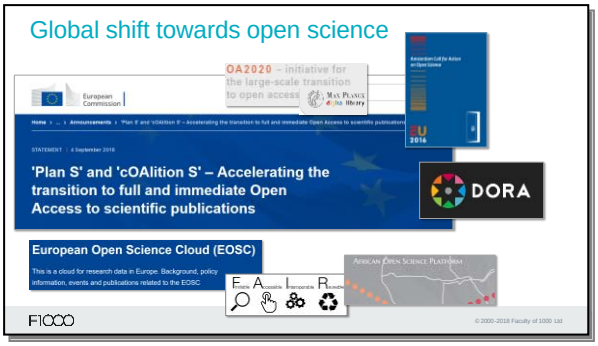
As you are aware, there is a global shift towards open science. There is a real shift in policymaking around the world, but particularly in Europe. The European Commission put out their Amsterdam Call for Action on Open Science in 2016 (Figure 3). This has led to a number of initiatives to implement policy and changes towards an open science system. There are a number of initiatives to move towards open access specifically. Max Planck Digital Library launched an initiative called OA2020. This is trying to shift and flip subscription journals to become open access journals and to encourage institutional subscriptions to combine open access fees in that subscription bundle.

One of the most significant initiatives recently is Plan S, which is the European Commission with the European Research Council together with another 11 funders (and the cOAlition S is growing), where those funders are stating that they will require immediate open access from 2020. There is still quite a bit to work out, but it is an important shift forward. DORA is also an important initiative. It is the San Francisco Declaration of Research Assessment and was launched in 2012. Many institutions and individuals have signed this Declaration to emphasize that they plan to change the way they assess research and researchers away from using impact factors, but there is still quite a bit of work to do that.

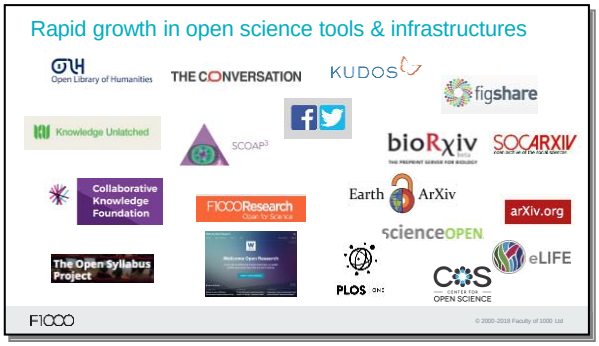
With respect to data, there are many major initiatives around world, in the US and Japan and elsewhere. Key initiatives include the FAIR initiative to make data Findable, Accessible, Interoperable and Reusable, as well as platforms to enable sharing of data and data services such as the European Open Science Cloud (EOSC) and the African Open Science Platform (AOSP).

There has also been a rapid growth in open science tools and infrastructures (Figure 4). For example, there are number of preprint service that have launched recently, and of course, arXiv which has been around for a long time in physics to enable rapid sharing of new findings and articles immediately on the server. There are number of new publication models such as *PLOS ONE*, which was the first journal to move to a process of technical peer review only and eLife that uses a collaborative peer review approach.

There is also a growing change in the way that we discuss and comment on research. There is an increasing use of Twitter and Facebook or blogs and tools like The Conversation to not only flag up new research but also to discuss and debate it. There are many new infrastructure projects, for example, the Collaborative Knowledge Foundation (Coko) and the others. There are also a number of approaches such as SCOAP[3], Knowledge Unlatched,



(Figure 3)



(Figure 4)

or Open Library of Humanities where institutions are coming together to try to underwrite the costs of open access so that the author does not have to worry about finding the funds to publish. These are all great and the speed of change in new approaches is increasing, but the challenge is to shift fully to open science to enable researchers to take up some of these new approaches.

Main barriers to Uptake of Open Science

What are the main barriers to uptake of open science? One of the biggest is that researchers are still typically judged by the impact factor or the journal brand of the articles that they publish. Impact factors and journal brands are ingrained in the whole assessment system, and it is very hard to see how you will displace it when they are so easy to use. Any replacement that is as easy to use will be just as flawed as evaluation is a complex issue and the correct metrics need to be used depending on what it is you are trying to assess. There is also some misconception that open science is not quality science, and it is important that we understand quality science can be open science and open science can be quality science. It should be open quality science.

There is also a challenge of the reality on the ground. Some funders like the Wellcome Trust in the UK have for a long time said that they judge research not on the venue of publication but on the research itself. The challenge is to ensure adherence of that policy on the ground in the review groups where when you talk to those individuals, they often say, "Yes, but I can see they have publications in *Nature*, *Cell or Science* etcetera and it is too easy to slip into using them." It also requires a change at all levels in the system right up to the university league tables because if universities are being assessed on the impact factor of their researchers' publications, they will ask their researchers to ensure they publish in high impact factor journals, and this is going to filter all the way back down the system regardless of changes in assessment elsewhere.

The key: separate publication from evaluation

In our view, the key to changing this whole system is to separate the decision about publication from the evaluation of the content and to move away from journals. Now that we publish online, there is no real need for journals. Readers certainly do not need journals. Readers typically search Google Scholar or PubMed or use other tools and approaches to find articles. It is only the authors that need the journals at the moment for the reflected benefit they provide to their careers. If a researcher has discovered a new finding, they should be able to share it with the community without a gatekeeper or an editor saying you cannot or it is not interesting enough for us. Equally the research community should be able to view new discoveries without any delay. Readers and users would also really benefit from being able to see the expert views of their peers on new findings. In addition, peer reviewers should receive due credit for the very important work they do in reviewing those findings and providing their expert opinion. New discoveries should be judged purely on the quality of the finding itself and not on the venue of publication.

The first important step towards this is preprints, being able to share your research immediately on a preprint server. There has been a very significant growth particularly in biology. For ex-

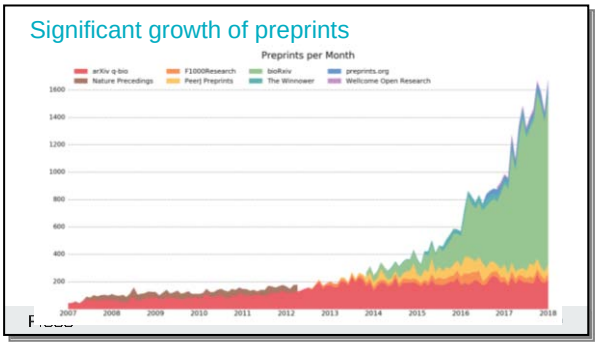
ample, in bioRxiv, we can see exponential growth in preprints being deposited (Figure 5). Preprints provide the benefits of being able to share rapidly and share all sorts of findings including null and negative findings and incremental findings, and the reader can access those immediately. Interestingly though, researchers even in physics who have been doing this for very long time still go from journal to journal to try and get the work published. The reason is so that they can get the impact factor of that journal to help their career.

F1000Research

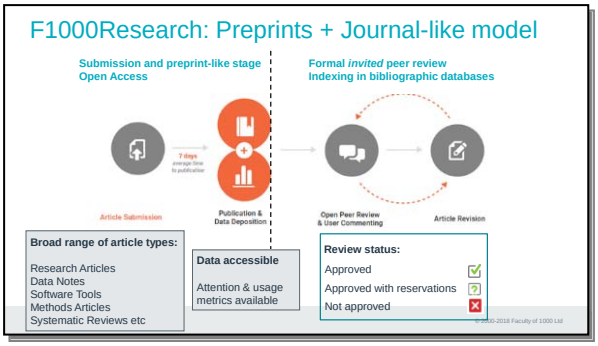
This makes no sense, and so we developed a publication called F1000Research (Figure 6). It is a platform which combines the benefits of preprints with the benefits that you get from a journal such as independent peer review, archiving, indexing, and XML. It starts off like a preprint. An author can submit a broad range of article types in the format that makes the most sense for that output. Our internal team then conducts a set of objective checks. We make sure that it is a piece of research, that the authors are from a recognized institution, that it is not plagiarized, that we have the data that underpin the findings, and if it meets those checks, we then publish the article. It is very much like a preprint, but the key difference is that this is now published, so you cannot take it to a journal to publish it. This is your journal publication, and at the point of publication, peer review starts.

When the author submits, they also choose a set of suggestions for reviewers from a list that we provide to them. They can suggest others as well, but our team check to make sure they are experts in the relevant field and that there are no conflicts of interest. The key to this process is that it is completely open and transparent. Who the reviewers are and what they say are published alongside the article. The other key point is that there are no editors. This is an author-driven process, and the author decides what they want to do throughout that process.

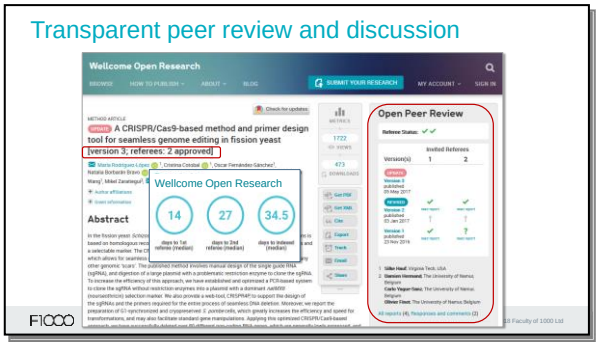
At the point of publication, we invite expert reviewers to review on behalf of the authors and the reviewers do two things. They provide a peer review report and they provide a peer review status. There are three options: 'approved' (which is shown



(Figure 5)



(Figure 6)



(Figure 7)

by a green tick); 'approved with reservations', which is a major revision (shown by a green question mark); and 'not approved', which of course is not reject as there are no rejects as it is published, it is just not approved (shown by a red cross). We have agreed with major bibliographic indexes like PubMed and Medline and Scopus and others that once an article achieves two 'approved' statuses, or one 'approved' and two 'approved with reservations', then it is indexed and all future and all past versions are then indexed.

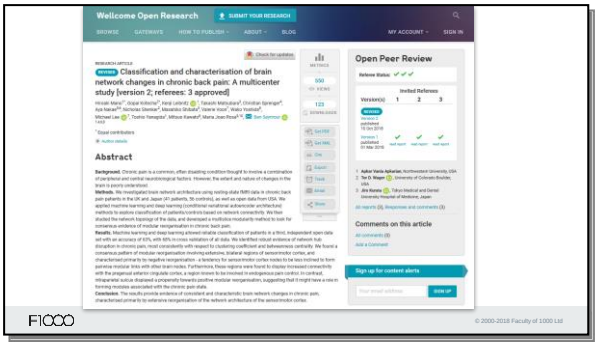
It is easier to understand the process with a real article (Figure 7). You can see all articles on the right-hand side have open peer review box which explains what is going on with the article. You can see who the reviewers are: you can see their names, and interestingly, reviewer two is three reviewers reviewing together. You can also see the versions. You can see this article has three versions and you can see the reviewer statuses of different versions. You can also see that versions can be used for revising an article, but they can also be used for updating it. It might be a review that you keep updating, or it might be a software article. New versions are independently citable, and the author is in charge of deciding if they want to revise or not. When they decide they are happy with the article, that is when that process stops.

The other important thing is that the article title also includes details of both the version number and the peer review status. It starts by saying 'awaiting peer review' and then it updates as review reports come in. This whole process speeds up publication significantly. The median time from submission to publication is 7 to 8 days. The median time to first referee is 14 days, to second referee is 27 days, and to indexed is 34.5 days, which is very fast.

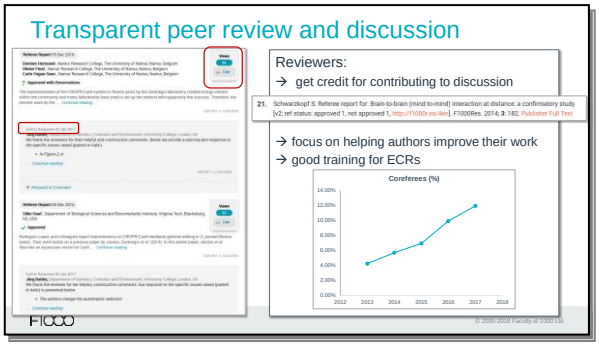
Figure 8 shows Ben Seymour's paper. He will be talking after me more about this article. If you click through from the links on the open peer review box, you get to see the peer review reports (Figure 9). If you click Continue reading, it expands out. You can see the author's response. Usually, they may say thank you or great ideas, but sometimes they refer to a point which was missed or disagree and have a discussion with the reviewers.

Peer review reports are also independently citable from the article, so the reviewers can get credit and we also track views. A lot of readers tell us that they look at the title, then the abstract, and then at the peer review reports before deciding whether to read the article. Reviewers can get credit for contributing to the discussions. This is an example of a citation of a peer review report cited not by the authors or by the reviewers.

It also changes the nature of the peer review.



(Figure 8)



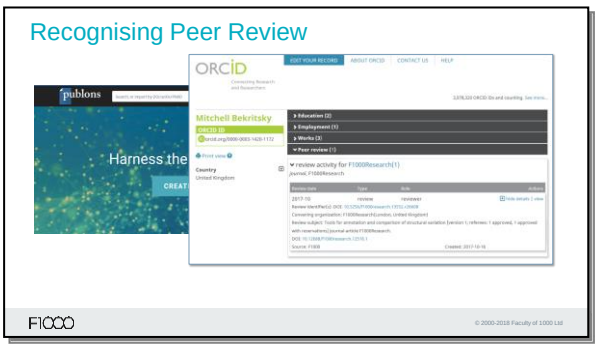
(Figure 9)

You are not helping an editor decide whether to accept or reject the article because it is already published. You are simply helping the authors improve the quality of the article, which is what peer review should be focusing on.

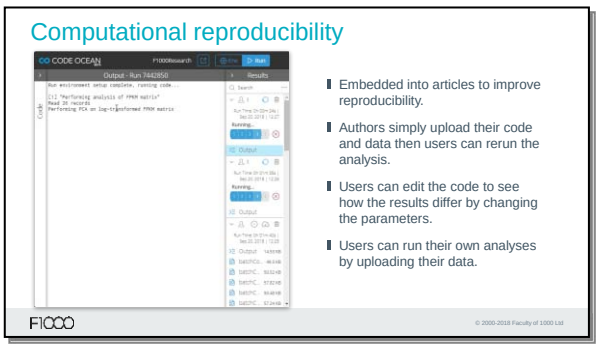
It is also often very good training for early career researchers. One interesting thing that we have seen is a significant growth in the number of peer review reports that are co-authored, so two or more people review together. Typically, this is a more junior scientist co-authoring the reports but also it is often collaborators on the other side of the world writing those reports together, which is interesting and important.

There are also other tools to help get credit for the peer review that you do. You may be familiar with Publons (Figure 10). It is a tool to enable researchers to capture all their peer review activity. If it is open peer review, you can see the review report, but equally you can track other review activity on closed review. You can also track it on your ORCID profile. ORCID is a unique identifier for researchers. I would encourage any researcher who does not have an ORCID ID to get one because it is simple and quick to do but has significant benefits in terms of pulling all your research activity together in one place.

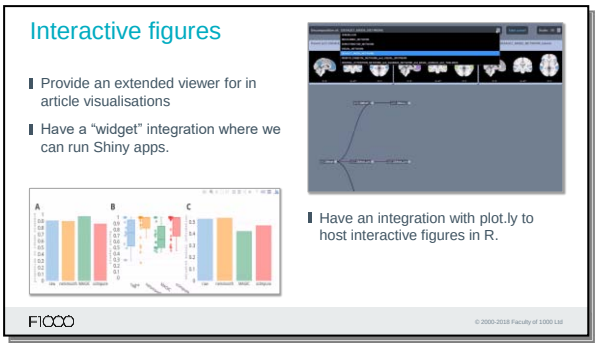
As mentioned, we also enable sharing of data, and one of the things that we have done is work with various providers to put widgets within the article to enable readers and peer reviewers to interrogate that data (Figure 11). We have integrations, for example, with the Shiny apps, with R code, and with other tools as well. We also have integration to support computational reproducibility. We work with an organization called Code Ocean, which enables authors to upload their code and data and then readers and peer reviewers can rerun the code (Figure 12). They can edit the code and see how that changes the analysis.



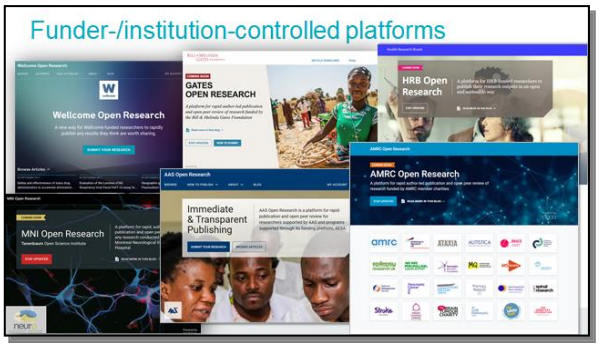
(Figure 10)



(Figure 12)



(Figure 11)



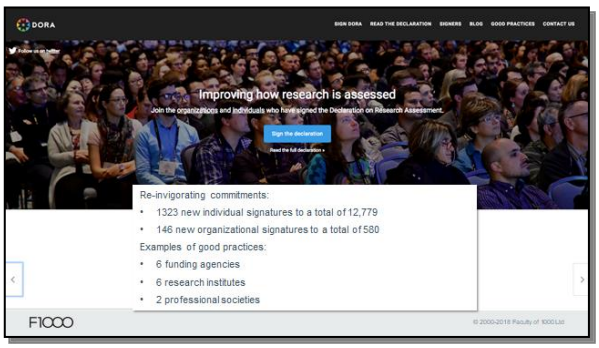
(Figure 13)

F1000Research has been running for almost six years, and we have worked during that time to refine the model. We are now working with funders and others to provide a similar publication model to their grantees to shift the whole system and provide researchers with reassurance that this is a safe way to publish even though we do not have an impact factor (and we intend to ensure we do not get one). We have also been working with funders and institutions to help support the system (Figure 13). The Wellcome Trust in the UK was the first to launch their platform called Wellcome Open Research where we are a service provider. It is an optional platform for their grantees and uses the same model of publication. The idea is that the Wellcome is saying: we have decided to fund you, the researcher, and now here is a platform to enable you to share anything and everything immediately on it. We have also launched similar platforms with the Gates Foundation, with the Irish Health Research Board, and many other groups. In fact, the European Commission put out a tender earlier this year for a similar platform for their framework program.

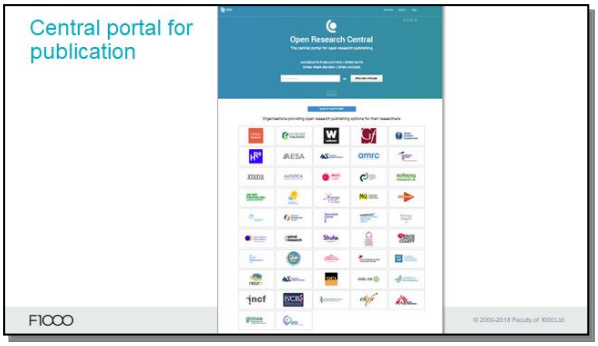
Figure 14 shows an article that was in the *Nikkei* a couple of weeks ago, which is about the Gates Open Research platform.

We do not want to end up in a situation where funder platforms compete with each other and you have funder impact factors. This is a transitory state. This is part of a trajectory towards where we think we should end up. Ultimately, we think researchers should be able to publish centrally, the citation should not be matter, that central platform should be maintained by the scholarly community, and publishers should simply compete for authors based on the quality of the service that they provide to enable publication on such a platform (Figure 15). If we cannot now use the citation as an indicator of quality, which of course we should not do, then we

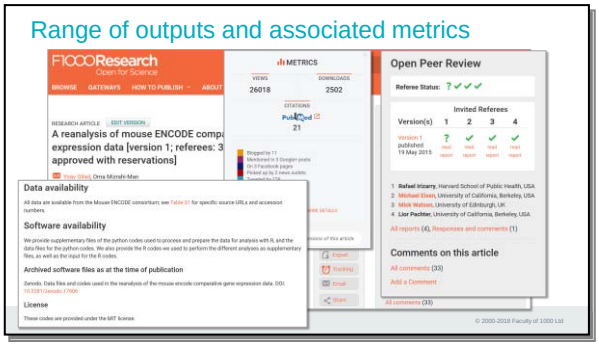
(Figure 14)



(Figure 16)



(Figure 15)



(Figure 17)

obviously still need tools to assess the quality level, the importance and potential impact of those outputs so that we can make decisions on funding and on careers.

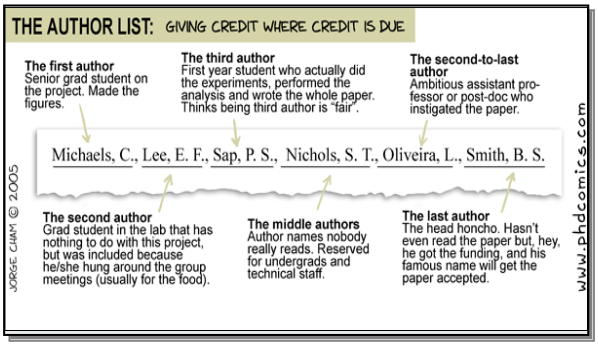
As mentioned earlier, DORA has been working for a number of years to encourage a shift in the way that researchers and research are evaluated and have been capturing a growing number of signatories and many more recently since its relaunch, but they are also capturing examples of good practice that we can then learn from and hopefully replicate (Figure 16).

There are many existing metrics that can be used and that we capture (Figure 17). For example, citation is still relevant and important and we capture those. Social media activity and altmetrics can be important in certain circumstances depending on what you are measuring. We think there is also something interesting and important that can be captured from the open peer review because you know who the reviewer is and you know what they said. Often, the reviewers may be the only people who have read the article from the beginning to the end, so it is an important opportunity to capture what they think of the article. We can also capture metrics around the elements within the article - around data, software, and other elements - and capture citations and usage of those.

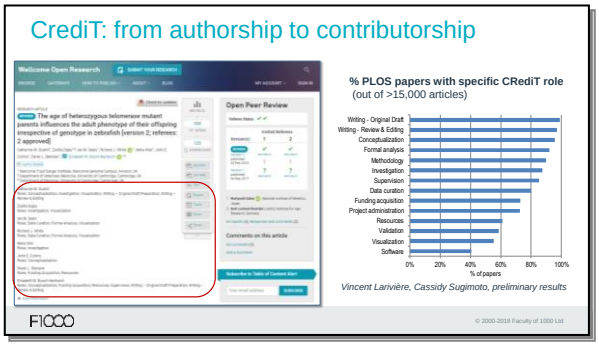
Figure 18 explains a situation that often arises around authorship and who contributed what and who should be first and who should be last and all those arguments.

A group of publishers, such as Nature and PLOS and others, have been working together on something called CRediT to put together an ontology of 14 different types of contributions that authors typically make to an article (Figure 19).

Many major publications require authors to provide this information. Figure 20 is an example.



(Figure 18)



(Figure 20)

Term	Definition
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims.
Methodology	Development or design of methodology; creation of models.
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
Validation	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
Formal Analysis	Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
Data Curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
Writing – Original Draft	Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
Writing – Review & Editing	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
Project Admin	Management and coordination responsibility for the research activity planning and execution.
Funding Acquisition	Acquisition of the financial support for the project leading to this publication.

CRediT

https://forum.casrai.org/c/standards

(Figure 19)

(Figure 21)

You can see exactly what the different authors contributed and therefore what can get individual credit for. There is an interesting analysis from PLOS where they have looked at over 15,000 articles. This shows the types of activities that authors contribute. You can see that about 80% of authors contribute to data curation, for example. Those authors often do not get proper credit, but if you are looking for somebody who is a good data curator, you need to know who the key authors are that contributed to that.

F1000Prime

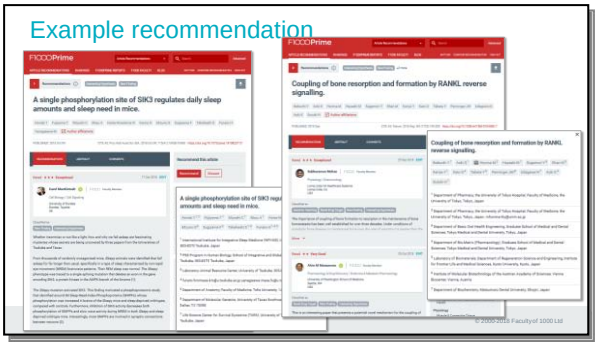
We also need tools to assess quality and importance. One of the reasons F1000 was set up at the beginning and named as such is because we have a service called F1000Prime, which comprises a large virtual expert faculty of over 8,000 members in biology and medicine (Figure 21). What they do is, as they read the literature as part of their research, they identify papers they think are particularly interesting or important, and write a short recommendation of the article. What is interesting is that the majority of the articles they recommend are not in the top impact factor journals. It shows there is a lot of high-quality research published all over the place. The faculty is global, and we have very prestigious experts, particularly in Japan, for example (Figure 22).

Figure 23 shows the recommendations, so you can see the article being recommended and the recommendations. In fact, articles can get many recommendations from different experts.

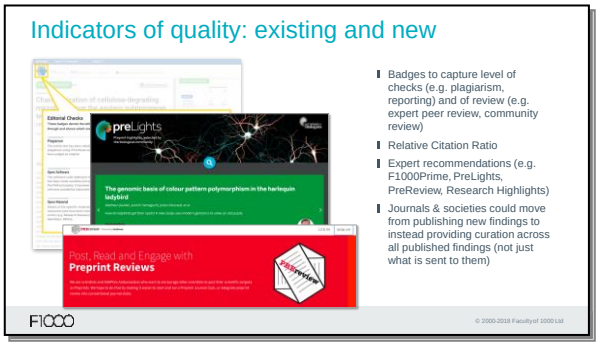
Indicators of Quality

There are a number of different indicators of quality (Figure 24). We are working with groups to look at badging, for example. It is often hard between various different preprints and journals to know what checks have been done. We are pro-

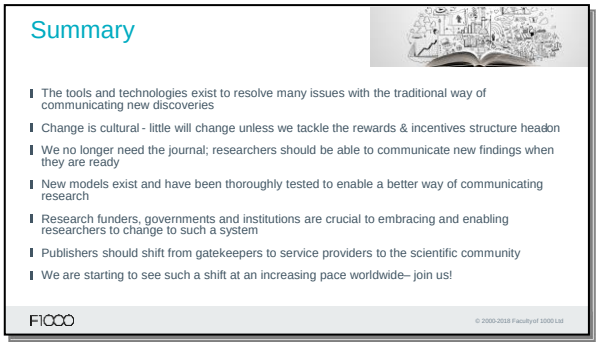
(Figure 22)



(Figure 23)



(Figure 24)



(Figure 25)

ducing a set of badges to capture the level of checks, for example, for plagiarism, for ethics, etcetera. There are also a number of other expert recommendation tools such as PreLights and PreReview, who look at preprints and other things. Indeed, the role of journals could change so that journals such as *Nature*, *Cell*, *Science*, etcetera, instead of publishing outputs could focus on the curation or on highlighting of articles that are published centrally and identify those that they think are particularly important and that reach their standard of novelty and importance.

Summary

To summarize, the challenge to change towards open science is not a technical issue; it is a cultural issue (Figure 25). Key to this is that we need to change the reward and incentive system. We no longer need the journal. There are new models already available, and they have been well tested and many funders and governments are now embracing those. We are now starting to see a shift worldwide towards some of these approaches. It would be great to work with colleagues and groups in Japan to explore some of these here too.