Free Science Publication Standard

Discussion Version 0.1

FSPS workgroup

2020/06/02

Abstract. This is a proposition for a scientific publishing framework that removes the role of a publisher in the traditional sense. This is possible because the publishers' role has been largely rendered dispensable by the Internet and online publishing. It is useful because it removes a layer of bureaucracy from the process, thereby allowing faster publishing as well as more flexibility for specialty formats such as data repositories, technical manuals, or negative results, while maintaining rigorous peer review procedures. It is necessary because extensive corporatization and commercialization of the scientific journals in recent decades has led to increasing concerns over detrimental influence of monetary interests on the scientific process. Whether or not these concerns are warranted in any specific case, they need to be addressed in order to buttress public trust in science, which is quite possibly more urgent today than ever before.

The proposed publishing framework is based on an open standard, which defines the editorial and review process, and the responsibilities of all involved parties. It does not define specifics such as formatting rules or number of reviewers; these decisions are left to the editor. It is intended as an additional path of scientific publishing for topics of formats that are difficult to publish otherwise and for authors who wish to avoid working for certain publishers, which may hold an effective monopoly on journals in these authors particular field of research. It is explicitly not intended as competition to classical society journals but rather as a way to facilitate the creation of new journals from within the scientific community when the need for new topics or formats arises. Scientific societies, universities, or similar organizations may act as hosts or platforms for editors and publications and define additional requirements on top of those detailed in this document, thus acting as virtual journals.

1 Rationale

Scientific research is ultimately a service to the public and is paid for largely by the taxpaying public. Yet, the results of this research are generally not made freely available to the public but must be bought from publishing companies by anyone who wants to read them. This includes publicly funded research institutions, which means that the public, who have *already* paid for the research, must pay for it again. Consolidation in scientific publishing has created de facto monopolies in this market over the past few decades, increasing the cost of this system for the taxpayer and the revenue of the publishing companies.

At the same time, the advent of the internet and desktop publishing has made the role of the publishers –typesetting, layout, printing, distribution– almost completely obsolete. Articles are now largely distributed electronically, which is effectively free of cost; hosting them on a website is very cheap and can just as well or better be done by libraries, which are already publicly funded; and software that makes proper typesetting and layout of scientific articles so easy that it is (or at least should be) expected of any undergraduate student, has been freely available for half a century. The lack of transparency of the results of scientific research for the public and the transfer of large sums of public funds to private companies for services that are no longer required, makes justifying this expensive system to the taxpayer difficult, to say the least.

A substantial part of the big influence of publishers (as owners of journals) over the scientific publishing process stems from what we might call the "reputation lock-in effect". Scientists publish their best research in journals that they consider reputable and prestigious because these journals published important discoveries in the past. Having published a lot in these prestigious journals is consequently critical when applying for a position or a research grant. The reputation of individual researchers thus depends on the reputation of the journals they publish in, which causes them to publish as much as possible in these journals, lending further credence to the journals. This feedback loop provides life support for a publishing system that is no longer the most efficient one possible. However, breaking such a loop requires a conscious effort and good alternatives.

While trust in scientific publications is currently based to a disconcerting degree on the reputation of the journal they are published in, rather than the actual contents of the individual paper, there is no reason why it would have to be. In fact, the reputation of the journals itself rests on the good work of scientists who have contributed to it as authors, editors and reviewers in the past. Taking the publisher out of this equation merely removes an unnecessary but costly middle man, and allows to assign proper credit and responsibility to all parties involved, including editors and reviewers.

Removing the role of the publisher in the scientific publishing process requires first and foremost to provide an *alternative authority* to supervise and lend credence to the review process. However, this need not be a personal entity. The Free and Open Source Software community has demonstrated that these functions can be exercised by open standards, to which all involved parties agree to adhere. Such a standard needs to be transparent to the general public and assign all involved parties their rights and responsibilities. It also needs to ensure that the result of the review process is transparent to the public and to give proper credit to all involved.

When using such a standard, the prestige of a publication no longer rests on the reputation of the journal but on the reputation of the editor and reviewers, and the institutions where they are employed. In return, the editor's and reviewers' reputations also profit from their involvement in high quality publications or may suffer from green-lighting papers of insufficient quality. This removes commercial interests from the decision of what is being published, while incentivizing the best possible review quality. Ultimately, such a system requires scientists to take back responsibilities they relinquished to for-profit companies to the detriment of the quality of and public access to science.

1.1 Outline of the review and publishing process

This section gives a brief overview of how the proposed Free Science Publication Standard (FSPS) ensures a proper review process and publishing of scientific manuscripts. The responsibilities of authors, editors, and reviewers will be introduced here and can be found in detail in the Guide for Authors (GfA), Guide for Editors (GfE) and Guide for Reviewers (GfR), respectively. Numbers in brackets reference the flow chart in Fig. 1.
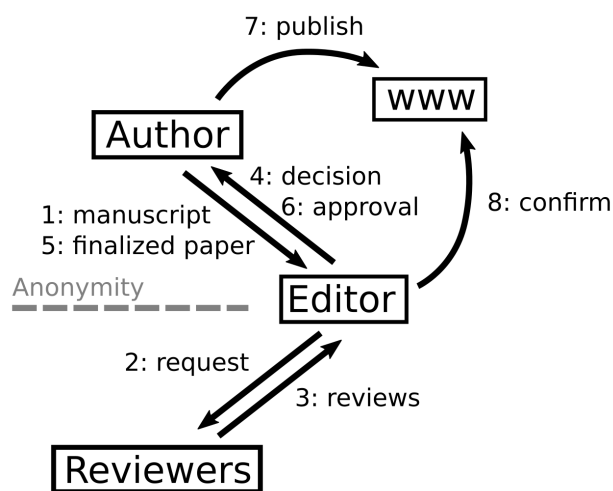

Fig. 1: Flow chart of the submission, review and publication process. For detailed explanation see text.

After the author or authors have completed a manuscript that they wish to publish, they format it in

a way that is suitable for review (see GfA). Any information about the authors shall not be included in the manuscript itself but in a separate form (GfA) to allow double-blind reviewing. The corresponding author then contacts a potential editor with the abstract and an invitation to preside over the review process for the manuscript based on the FSPS. How potential reviewers, who are willing to participate in the proposed system can make themselves known to prospective authors, is detailed in the GfE. If the editor accepts the invitation, editor and authors agree on how the final result shall be formatted and made public and the corresponding author sends the entire manuscript to the editor [1].

The editor then contacts an appropriate number of reviewers (depending on the scientific field and scope of the manuscript) with the abstract and an invitation for review and sends those who accept the entire manuscript [2]. The identity of the authors (including their affiliations) and reviewers shall be known only to the editor during the entire review process to counteract conscious and unconscious bias one way or the other (GfA, GfE and GfR). However, once the paper gets published all involved parties, including all reviewers, will be made public on the title page, to discourage nepotism and dispel any suspicions thereof.

Once the reviews are completed and sent back to the editor [3], the latter decides whether or not the reviews are overall in favor of or opposed to publication of the manuscript in it's current or revised form and informs the corresponding author about this decision [4]. The decision may or may not include the necessity for additional editing and/or review rounds [1-4].

When and if the editor is satisfied that the manuscript's content is ready for publication, the authors format it in a way that is suitable for publication and send the finalized version to the editor [5]. The authors may choose to outsource the typesetting and layout at their own discretion. The editor checks the manuscript one last time, adds the names and other relevant information of the reviewers to the front page and sends it back to the corresponding author with a letter of approval [6].

The exact version that has thus been approved is published online in a way previously agreed upon between the authors and the editor [7], including all information mandated by the FSPS. This can for example be the homepage of the institution of one of the authors or of a scientific society, or a library or archive. Finally the editor publishes a confirmation that the information, which the authors have published, is correct, along with a duplicate of the publication or the hash value (see following chapter) of the file, on their own personal or professional homepage [8] to counter attempts at misuse of the system. The reviewers are free to also publish confirmations of their involvement, which gives them well deserved credit for their work, increases the reach of the paper and buttresses the credibility of the FSPS.

1.2 Benefits and opportunities

The obvious benefit of the proposed publishing framework is that it untangles monetary interests from scientific publishing. Why that is a good idea has been so thoroughly debated in recent years that it seems redundant to repeat all arguments here. The same is true for the benefits of publishing only open access. However, there are two main problems with most current open access journals, where the author(s) pay a fee for processing and publication of their manuscript: 1) The high fees of "prestigious" journals shut out colleagues from lower income countries, which filters out many high quality studies and important findings before the review stage; and 2) the fees create an incentive for journals to rubber-stamp any and all submissions, which has led to a flood of predatory journals publishing rubbish papers and does considerable damage to the credibility of science as a whole in the eyes of the public. Both of these problems are addressed by removing pecuniary motives from the publishing process.

The second benefit is that the proposed scheme in principle allows peer reviewed publication in arbitrary formats. While written accounts are a perfectly good way of presenting many or even most types of scientific findings, technical advances in recent decades have provided us with more and more data sets that are challenging to represent adequately due to their sheer size. As a result, large sets of raw data disappear in desk drawers and only condensed and filtered data and their interpretation gets published. Raw data in forms that cannot be printed on paper may be relegated to electronic appendices but these are often not properly reviewed. There is an argument to be made that this is a good thing and that we should not clutter the world with raw data that may only contain noise. However, modern big data methods are becoming better and better at pattern recognition on a scale that humans cannot oversee. Good raw data, published alongside peer reviewed methodology in a form that makes many such data sets accessible to automatized analysis will therefore become an increasingly valuable research tool in the future. The individual data sets may seem meaningless by themselves but when combined, they may reveal important new information.

The FSPS is format agnostic, and thereby allows to publish anything that authors may find worthy of publication, provided they find an editor, who agrees with this position, and that the reviewers approve the quality of the research. This means that things like time resolved raw data in the form of videos can be treated like any other data that one may normally present in a table. It also means that the FSPS can be used for publishing anything of scientific interest with proper peer review, be it regular papers, textbooks or textbook series, educational videos, classroom materials, technical manuals, negative results, etc.

Finally, the FSPS allows quick and easy setting up of "impromptu magazines" by providing all necessary guidelines and processes. This can be used for topical

issues or conference proceedings and should also allow smaller scientific societies or small conferences to create their own journals without the need for the time and resource consuming development of the technical details. Ideally it should make it easier for the scientific community to react quickly to new developments that may require new forms of publishing.

1.3 Preventing fraud with hash values

The *hash value* of a file is a unique identifier created from the file by applying one or several *hash functions* to the file. A variety of *hash algorithms* (combinations of hash functions) are available as free software. For example, the commonly used SHA256 algorithm produces a hash value of 256 bits, expressed in 64 hexadecimal digits.

The authors and the editor shall independently produce such a hash value from the file that they agreed on as the final version of the manuscript and publish it independently. The used algorithm must be the same and must also be reported. This allows readers to confirm that the paper they downloaded is indeed the exact same that the editor green-lighted for publication. To do so, the reader only has to apply the same hash algorithm to the downloaded paper and compare the result to the values published by authors and the editor. This method is extremely well tested and ubiquitous in software publishing due to its simplicity and security.

1.4 Double-blind reviewing

Reviewers and authors of scientific papers are humans, at least for now. As such, they are never entirely free of bias, which is a problem when evaluating the scientific merits of a manuscript. Knowing who wrote something influences our perception of the content of the text in ways that may not be justified. Likewise, being aware that we do not know the author of a manuscript on a topic we fancy ourselves experts in, also affects our evaluation of that manuscript. This effect makes it harder for young scientists to get their works published via a non-blind review process than for established experts.

What is worse, is that it creates echo chambers, which stifle innovation and help cement established concepts irrespective of their actual merit but based on *assumed expertise*. While it may *often* be true that the seasoned expert knows their topic better than a PhD student trying to publish their first paper, it is unquestionably *always* true that a good manuscript must be able to convincingly stand on its own without support of name recognition. Therefore, the actual quality of the research and of the resulting manuscript should be evaluated purely on the basis of the actual manuscript, independent of implicit or explicit assumptions about the authors qualifications. The same holds true for the quality of reviews, and authors should not be able to dismiss a reviewer's criticism on personal grounds either.

1.5 The importance of transparency

Science lives and dies by the trust in the scientific method, which is monitored and verified by the review process. It is therefore pivotal that the fairness and independence of the review process be beyond any doubt. This is only attainable by making the process and all involved parties transparent to the public. Conflicts of interest should be avoided where possible but must be openly addressed where they are unavoidable.

Supposed and sometimes imaginary conflicts of interest have become one of the chief criticisms of the current anti-scientific animus in parts of society, not least as a result of lack of transparency and communication about quality control procedures on the part of the scientific community toward the public. To ignore this tendency is dangerous and it can only be counteracted effectively by full transparency. The FSPS therefore mandates all information that may be pertinent to evaluating potential or alleged conflicts of interest to be published on the front page of the paper (see The Front Page).

1.6 Platforms

Although not strictly necessary for the proposed publishing framework, having some kind of organizational platform facilitates its implementation. This could be an existing organization such as a scientific society, a university (library), a science funding agency, or a (scientific) social media website. It could also be a new kind of platform that would be founded for the specific purpose of publishing science openly and not for profit, using the FSPS. It is up to these platforms to decide what services they want to provide and what they want to use the FSPS for. The most obviously useful service would be to provide a point of contact between prospective authors and reviewers, e.g. in form of a list of colleagues who have agreed to participate as editors, along with their respective fields of expertise and contact data. The platforms may also choose to publish and archive the publications that have been handled by the editors associated with them. They may wish to define certain things that are left open in the FSPS, such as how many reviewers are required, what topics or formats they want to publish, or what hash algorithms are to be used, as well as imposing specific formatting restrictions for things like citations or figures. If they aim for a coherent look of publications, they may want to provide formatting templates or a place where such templates can be shared. The platforms may also opt to take on the role as a supervisory body with the power to comment on or retract publications associated with them, publish responses or issue rebuttals. Finally, the platforms should either assign a DOI number to all publications managed through them, or provide guidance on how to obtain a DOI number.