Open Public Response to Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research

1/12/2012

This is a public and open document intended to draft a collective response to the request of information posted by the Science and Technology Policy Office (OSTP), on whether peer-reviewed publications resulting from federally funded research should be required to be made publicly available.

Dear Office of Science and Technology Policy,

Kitware applauds the initiative of the OSTP on seeking public feedback on these matters of high relevance to the scientific community and to the American public. However, please note that this is not an official Kitware response.

In order to contribute to this process, we reached out to our many collaborators and invited them to join us in writing a collective and thoughtful response to the insightful questions of the RFI. The result is the document attached to this submission letter. The names of the contributors and those in favor of this response are found at the end of the document.

Please find below our response to the RFI on "Public Access to Peer-Reviewed Publications from FFSR". NOTE: In the responses below we use the following acronyms:

FFSR: Federally Funded Scientific Research

Question 1: Are there steps that agencies could take to grow existing and new markets related to the access and analysis of peer-reviewed publications that result from federally funded scientific research? How can policies for archiving publications and making them publicly accessible be used to grow the economy and improve the productivity of the scientific enterprise? What are the relative costs and benefits of such policies? What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?

Response:

Grow Existing Markets Related to Access:
A vibrant market of open access publishers has developed in the past ten years, with thousands of journals covering many different fields. The Directory of Open Access Journals [http://www.doaj.org/] currently lists 7,372 journals. Furthermore, several traditional publishers - such as Springer, Wiley, and Nature Publishing Group - have implemented Open Access options that enable authors to choose to publish their articles openly, either in Open Access journals or by way of hybrid Open Access schemes. In the latter, authors can chose to pay the publisher to allow their papers to be freely downloaded by readers, or use the traditional approach by which readers are expected to pay the publisher in order to get access to the papers, either via subscriptions or one-time payments for a given article.

Federal agencies should support the adoption of "Open Access" as the standard way of publishing the results of federally funded research. What has been termed the "author pays" model should be understood as a "funding agency pays" model for publishing fees in Open Access journals. The NIH, as part of its Public Access policy, has already stated that publication fees can be charged to grant funds. This a standard practice of funding organizations such as the Wellcome Trust and the Howard Hughes Medical Institute.

It is also fundamental to produce a clear definition of what "Open Access" means. We strongly propose that the definition of "Open Access" for articles resulting from federally funded research must be articles that are distributed by the copyright holder under the Creative Commons by Attribution License 3.0: http://creativecommons.org/licenses/by/3.0/. Under this license, anyone is permitted to copy, distribute, and create derivative works of the article, with the only requirement of providing attribution. We propose that "proper attribution" be defined as citing the uniform resource identifier (URI) or digital object identifier (DOI) of the original article. No further requirements for attribution should be demanded, and particularly, the attribution methods must not be left to be defined by authors on a case-by-case basis. Instead the federal agency must specify this standard method of attribution to ensure a reasonably low bar of effort that will lead to compliance.

Previous experience with the NIH Public Access policy has demonstrated that the open access policy is ineffective if it is not enforced. Federal agencies should therefore implement a system for verification of compliance, which should be reported as part of the "past performance" section of future funding applications. In this way, researchers' compliance with public access policies will benefit applications for new funding.

An interesting and detailed set of suggestions for paying Open Access publication fees can be read here: http://blogs.law.harvard.edu/pamphlet/2011/11/16/how-should-funding-agencies-pay-open-access-fees/

Grow the Economy:
As an economic model, the goal of the scientific enterprise is to gather knowledge and information and to disseminate it in a usable form. The public availability of content will be the most important way of increasing efficiency and productivity of the scientific enterprise. Removing the barriers to scientific publications creates opportunities for developing new collaborations and for investigating previously unforeseen avenues of scientific research.

The public availability of articles will facilitate their access by academic institutions, companies, and citizens, and will reduce the amount of time it takes for research to impact small businesses and start-up endeavors. On a global scale, developing countries will profit immensely from free access to information as this will give them a chance to develop their own economies. Indirectly, the US and other countries will profit from the opening up of new markets.

The traditional process by which publishers request unpaid copyright transfers from authors, and then use those same copyrights to put articles behind toll-gates that restrict access to information for the 95 years awarded by copyright laws is detrimental to the further development of the scientific enterprise. Worst of all, they are an impediment to the education and appreciation of scientific research by the general public, as well as to the participation of the public in furthering those research efforts.

Examples of successful efforts for engaging the public in the practice of scientific research include the Polymath project, the NASA Galaxy Zoo, the Moon Crater Zoo, and the regular involvement of amateur astronomers in comet discovery. Michael Nielsen, in his book "Reinventing Discovery," goes into more detail on how massive participation of regular citizens, endowed with online collaboration tools, are transforming the practice of science. The need for and benefits of a networked society with free and unrestricted access to knowledge are dealt with in depth in the recent book "Too Big to Know" by David Weinberger.

Costs:
The current publication process is inefficient and not cost effective. For example, articles submitted by authors are unnecessarily deconstructed, retyped, and re-edited by the publisher to recompose a final version with only minimal incremental refinements. This process would be far more efficient and cost effective if publishing software was widely available for researchers to write initial versions of their articles in a collaborative fashion with their colleagues. In turn, articles can be passed directly to publishers using open standard file formats.

Federal agencies can help increase the efficiency and reduce the cost of publishing by supporting the development of open and royalty-free standards for scientific publications, and encourage commercial applications to implement these standards.

Types of Access Necessary:
The status of "public availability" must be defined in terms of:

(a) Placing the articles in Public repositories, (without requiring registration or fees)
(b) Distributing the articles under Creative Commons by Attribution License

Other licenses that impose restrictions such as *"non-commercial"* or *"only for research"* are not useful in practice, given that U.S. courts consider most universities and non-profit organization

activities to be of a commercial nature. Authors of papers must be required by federal agencies to retain the copyright of their articles (or at least be precluded from transferring copyright away) and in this way remain empowered to make licensing decisions about the articles. Princeton University and Harvard University faculty members have stopped the common practice of transferring copyright of articles to journals. The practice of not transferring copyright must become a policy of federal agencies to be applied to recipients of federal funding.

[See for example: http://dx.doi.org/10.3897/zookeys.150.2189 "Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information"]

[See Princeton University report http://www.cs.princeton.edu/~appel/open-access-report.pdf]

Question 2: What specific steps can be taken to protect the intellectual property interests of publishers, scientists, federal agencies, and other stakeholders involved with the publication and dissemination of peer-reviewed scholarly publications resulting from federally funded scientific research? Conversely, are there policies that should not be adopted with respect to public access to peer-reviewed scholarly publications so as not to undermine any intellectual property rights of publishers, scientists, federal agencies, and other stakeholders?

Response:
In addition to the stakeholders listed in this question, it is critical to note that the general public is the primary stakeholder to be considered here. Given that in the context of FFSR it is the public's tax dollars that are paying for the scientific research, the public's interest should be considered foremost when considering trade-offs between available options.

In order to have a productive discussion on intellectual property, it is important to first deconstruct the term "intellectual property" and clarify its meaning in the context of current U.S. laws. We do this in **Appendix A** and conclude that **copyright** is the only concept of intellectual property that is relevant to this RFI.

Copyright is originated by the authors of articles, when they put the expression of ideas required to disseminate the outcome of their research in a tangible medium. In the context of federally funded research, authors are performing this work as part of their job duties. Therefore the articles are the outcome of *"work for hire"* and it is the employer of the authors who holds the copyright of the resulting articles. It is commonly the case that universities and other research institutions assign that copyright to the authors themselves, but this is a matter of policy choices by the institutions. It is a common commercial contractual practice that when one organization contracts another to develop creative works, the paying organization will retain some of the copyright rights (if not all) of the resulting creative work. In the context of federally funded research, it will be then consistent with common commercial practice of requiring awardee institutions to return the copyright of the articles resulting from federally funded scientific research (FFSR) to the federal agency. The U.S. government does not originate copyrights, but it can hold the copyrights of creative works when they are transferred to it.

For example, the Federal Acquisition Regulations: *FAR Subpart 27.4—Rights in Data and Copyright.* https://acquisition.gov/far/html/52_227.html

"(1) Data first produced in the performance of this contract.
(i) Unless provided otherwise in paragraph (d) of this clause, the Contractor may, without prior approval of the Contracting Officer, assert copyright in scientific and technical articles based on or containing data first produced in the performance of this contract and published in academic, technical or professional journals, symposia proceedings, or similar works. The prior, express written permission of the Contracting Officer is required to assert copyright in all other data first produced in the performance of this contract.
(ii) When authorized to assert copyright to the data, the Contractor shall affix the applicable copyright notices of 17 U.S.C. 401 or 402, and an acknowledgment of Government sponsorship (including contract number).
(iii) For data other than computer software, the Contractor grants to the Government, and others acting on its behalf, a paid-up, nonexclusive, irrevocable, worldwide license in such copyrighted data to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly by or on behalf of the Government. For computer software, the Contractor grants to the Government, and others acting on its behalf, a paid-up, nonexclusive, irrevocable, worldwide license in such copyrighted computer software to reproduce, prepare derivative works, and perform publicly and display publicly (but not to distribute copies to the public) by or on behalf of the Government."

It has also been clarified that, for the purpose of Federal Acquisition Regulations, computer software falls into the category of data.

Federal agencies could therefore introduce requirements by which the copyright of articles resulting from FFSR should be transferred back to the federal agency.

As described in **Appendix A**, the economic logic behind copyrights does not hold for the case of articles resulting from funded scientific research. This is because in the context of scientific publishing, copyright does not have an economic role in "encouraging the creation of works of authorship" (since authors do not get paid by publishers), nor does it play a role in protecting creative works as scientific articles are expected to have a minimal amounts of "creative, invented" content.

The weak level of copyrightable content that a serious scientific research article should have, combined with the rationale that the purpose of copyright is to benefit the public by making available the result of creative works, and that the public has provided monetary compensation to produce those creative works, leads to the conclusion that copyright protection is not really required in the context of FFSR scientific research.

It is also important to clarify that in the standard practice of scientific publishing, publishers are not originators of intellectual property. The real creators are the researchers who write articles intended for publication. Publishers acquire intellectual property on those articles through the practice of requiring authors to sign copyright transfer agreements as a condition of publication. In these transactions, publishers do not provide any monetary compensation to authors. In economic terms, the transfer of copyright from authors to publishers is essentially a donation.

As a result, there is no need to provide any protection for publishers, as they are already acquiring for free a product in which they have not invested any significant financial resources to produce. The contribution of the publisher is limited to coordinating the work of associate editors to compose the final collection of articles for publication, to host the digital documents online, and to provide the gates that regulate access to the publications to paid subscribers only.

In most cases, associate editors and reviewers who contribute the bulk of the peer-review process are volunteers who are not paid by the publisher. Therefore, it should be questioned why publishers benefit for free from the work that scientific researchers as authors create under the support of FFSR, rather than the taxpayers, who are paying for the bulk of the research enterprise and are the rightful copyright holder.

This doesn't mean that publishers do not need to be paid for providing the service of disseminating articles. As providers of commercial services, publishers certainly deliver a valuable contribution to the scientific enterprise and must be compensated for such services. However, that compensation does not have to be achieved at the price of restricting access to FFSR articles. A variety of business models that make Open Access a viable financial endeavor have been demonstrated in the past ten years.

Our point is that intellectual property, particularly copyright, is not needed in this economic transaction that compensates the publishers for their services; under modern business models of open access publishing, there is no need for publishers to hold the copyright of the articles. Instead, publishers simply need to be the recipients of a license given by the copyright holder, allowing the publisher to copy, distribute, create derivative works, and perform public displays of the articles. The prime examples of licenses suitable for this purpose are the Creative Commons by Attribution license, and the Creative Commons Share Alike license.

Note that other Creative Commons licenses, such as the CC Non-Commercial license and the CC No Derivatives license, will not be suitable for allowing publishers and other institutions to productively use the articles resulting from FFSR.

More can be found on these topics at:

- https://svpow.wordpress.com/2011/10/22/economics-of-open-source-publishing/
- http://blogs.lse.ac.uk/impactofsocialsciences/2011/11/09/functionality-academic-publishing/

Question 3: What are the pros and cons of centralized and decentralized approaches to managing public access to peer reviewed scholarly publications that result from federally funded research in terms of interoperability, search, development of analytic tools, and other scientific and commercial opportunities? Are there reasons why a federal agency (or agencies) should maintain custody of all published content, and are there ways that the government can ensure long-term stewardship if content is distributed across multiple private sources?

Response:

Centralized Option

Pros

- It may facilitate the creation of uniform methods of accessing and searching for articles.
- As with Apps platforms such as the Android and iPhone operating systems, a cohesive platform can lead to innovative development of different means of accessing, searching, deconstructing, and analyzing articles.

Cons

- It tends to result in bottlenecks, delays, congestion, and a lack of flexibility and agility.
- Centralization results in creating a single point of failure, where the entire system depends on a critical piece to be working all the time.
- It will take longer to be put in place.
- It makes it very difficult to innovate over time and to introduce new functionality that can transform the way that the data is used.

Decentralized Option

Pros

- It spreads and distributes the load of the system across multiple archives.
- When combined with smart redundancy, it provides protection against potential loss of information.
- It provides an open market for innovative methods to evolve, which enables researchers and the public to consult and data-mine the content of scientific publications.

Cons

- It requires a concerted and coordinated effort to define standard mechanism for
    - Replication of data
    - Federated search
    - Interoperability

Conclusion:

The system should be decentralized based on agreed standards and interoperability. Federal agencies should host archives of the published materials, but those archives should be commonplace and be replicated in different institutions (for example, university libraries).

Wide replication is the best way of ensuring continuous availability. This method is the essential mechanism used by the Internet itself, and has also been demonstrated by large scale source code repositories in Github (http://github.com), where some of the most popular code repositories have been replicated thousands of times (for example, see https://github.com/popular/forked).

Rich environments of replication, combined with SHA1 hashing that makes it possible to verify differences between multiple copies of a resource, guarantee the perpetual availability of a digital resource. To be more specific, a worldwide cataclysm would be the only way to wipe-out all copies of the “rails” repository, for which 2,569 copies have been made available worldwide: [https://github.com/rails/rails/network]. Distributed replication, versioning, searching, and indexing are standard features in peer-to-peer software applications, of which several open source implementations are available.

A decentralized storage solution, however, must be paired with a federated system of indexing and searching for content to ensure ease of search and access to the publications. Such systems are widely available and have been used to support many legal applications of peer-to-peer networks.

Question 4: Are there models or new ideas for public-private partnerships that take advantage of existing publisher archives and encourage innovation in accessibility and interoperability, while ensuring long-term stewardship of the results of federally funded research?

Response:
University libraries, archives, and public libraries already archive articles and provide long-term stewardship of the results of FFSR. It has been only due to the recent publishing practices of copyright control, such as DRM and online-only-licensing for access to articles, that libraries have been prevented from playing their natural role of long-term stewards of published content. This has been true for many years in the case of books.

Long term preservation of published materials is the job of libraries and archives, not the job of publishers. Libraries and archives have a much better guarantee of longevity than publishing businesses and societies that provide the services today.

Once federal agencies implement policies that preclude researchers from transferring copyrights to publishers, and that require researchers to make articles available in public repositories using appropriate licenses, libraries and archives will be able to regain their historical role as long-term stewards of these published materials.

No individual organization can be a sole, reliable provider of long-term, fail-safe storage for the large body of articles resulting from FFSR. As the Internet itself has demonstrated, only a distributed, decentralized system built upon light and open standards can provide reliable, long-term, and innovative support of the public dissemination of information.

The adoption of permissive practices on copyright and licensing for FFSR publications will be of fundamental importance ing enabling the unfettered replication of articles in any medium, including digital ones. It will therefore empower decentralized systems to host replicated archives of the articles, along with experimenting with innovative technologies for maximizing the dissemination and collective exploitation of the information contained in the articles.

There are open access journals and publishers that are examples of successful models of publishing innovation and stewardship, such as the Insight Journal, PLoS, and BiomedCentral. These journals foster accessibility to the results of scientific research and are creating a new paradigm for scientific publishing. More importantly, they are reviving the support for verification of reproducibility, which should be the hallmark of scientific research.

Question 5: What steps can be taken by federal agencies, publishers, and/or scholarly and professional societies to encourage interoperable search, discovery, and analysis capacity across disciplines and archives? What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities? How should federal agencies make certain that such minimum core metadata associated with peer-reviewed publications resulting from federally funded scientific research are publicly available to ensure that these publications can be easily found and linked to federal science funding?

Response:
Adopt standards of publishing technology that:

- Do not rely on proprietary formats and
- Are not subject to proprietary restrictions (patents or copyrights).

There is an abundant body of publishing technology that is openly available to the public. Examples include RTS, Latex, HTML, Wiki formats, ODT, ebooks.

All the adopted formats must be machine readable (digital) to facilitate indexing and large-scale data-mining of the literature. The Library of Congress, in collaboration with the National Library of Medicine should define a minimalistic schema of metadata, and it should be done in less than six months. Much of this work is already done by PubMed and Medline.

Standards of unique resource identifiers such as the ones provided by “handle.net” should be required.

Question 6: How can federal agencies that fund science maximize the benefit of public access policies to U.S. taxpayers, and their investment in the peer-reviewed literature, while minimizing burden and costs for stakeholders, including awardee institutions, scientists, publishers, federal agencies, and libraries?

Response:
1. Streamline the process of paying for publishing services. For example, a standard R01 grant should have a pre-specified budget for paying for publication charges in open access journals based on the “authors-pays” model (http://www.plos.org/publish/pricing-policy/publication-fees/). Publishers should negotiate their rates with the government in the same way that all other contractors (from service providers up to manufacturers of military equipment) do.

Negotiated rates are already a requirement for all universities and companies that receive grants from and do contracted business with the federal government. There is no reason why publishers shouldn’t be subject to the same conditions of rate negotiation when they are providing services to the federal government.

[For more details, see: https://www.acquisition.gov/far/html/Subpart%2042_7.html]

2. Implement a distributed system of repositories that provides redundancy of storage along with a distributed system for indexing and search that can be easily navigated without a single point-of-failure or bottlenecks.

This technology is already available in the form of peer-to-peer networks, for which multiple free and open source software implementations are available.

3. Define a specific set of copyright licenses that will be admissible for labeling articles as "publicly available," and then require that all articles resulting from FFSR be made available by distributing them under one of these accepted licenses.

In particular, these licenses must not have any restrictions on the commercial use of the content, and must allow for modification and redistribution of the copyrighted content. Ideally, this would be the Creative Commons by Attribution license 3.0, and the Creative Commons by Attribution Share Alike 3.0 licenses. The Creative Commons Non-Commercial license should be excluded. This is consistent with what open source communities did for the open source definition, which requires that licenses allow for modification, redistribution, and commercial use of content.

[For more details, see: http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.0030097]

Question 7: Besides scholarly journal articles, should other types of peer-reviewed publications resulting from federally funded research, such as book chapters and conference proceedings, be covered by these public access policies?

Response:
Yes, from both an economic point of view and that of promoting the progress of science and technology, all non-classified information that derives from FFSR must be made publicly available. This must include book chapters, conference presentations, articles in conference proceedings, audio recordings, podcasts, video recordings, and training materials whose content is **substantially** based on FFSR results.

If the public has paid for the development of any of these materials, then the public must have unrestricted access to them. Note that this is not "free access," given that the taxpayers have indeed *already* paid for those materials beforehand. This is simply returning to the public what the public has paid for.

More on this topic at:
[http://www.nytimes.com/2012/01/11/opinion/research-bought-then-paid-for.html?_r=2]

Aside from classified and export controlled materials, the only other exception that should be made is the protection of the privacy of human subjects participating in medical research. Note however, that once medical datasets have been anonymized properly, they should fall in the category of public dissemination. Federal agencies, particularly the NIH, should create an easy option for patients to consent to share their medical information if they wish to do so once properly informed of the implications, both in the sense of risk and in the sense of the potential benefit for the advancement of scientific research.

Question 8: What is the appropriate embargo period after publication before the public is granted free access to the full content of peer reviewed scholarly publications resulting from federally funded research? Please describe the empirical basis for the recommended embargo period. Analyses that weigh public and private benefits and account for external market factors, such as competition, price changes, library budgets, and other factors, will be particularly useful. Are there evidence-based arguments that can be made that the delay period should be different for specific disciplines or types of publications?

Response:
No embargo period should be required or allowed. Articles should be made publicly available immediately after being published. In a typical FFSR project, the American taxpayer has paid for the research with one or more years in advance by dutifully paying their contributions to the federal budget in the form of taxes. There should not be any further delay in making the results of the research available to taxpayers, as they have already paid for it.

Publishers' business models must be restructured in such a way that they are no longer an obstacle to the public dissemination of scientific information. The viability of such models has already been demonstrated by open access publishers, and with the options offered by hybrid publishers; that is, publishers that offer authors the choice of processing their articles as open access articles, or as traditional closed access articles.

The cost of publishing is about 1% to 2% of the cost of performing research. This cost is already paid by the federal agencies through the indirect channels of overhead (indirect costs) that goes to finance the operation of research institutions, including their libraries, and particularly the subscriptions that the libraries pay to publishers. It would be a lot more efficient to clearly incorporate the cost of publication upfront into the preparation of research proposals and utilize such a fraction of the budget to pay for the publication fees of open access publishers. Notice that this doesn't at all diminish the peer-review process that is required to ensure the high quality of content, given that this activity can continue to be performed on a volunteer basis, as it is done today. Publishers today do not pay authors, reviewers, or associate editors for the work they contribute to the endeavor of preparing and reviewing articles for publication.

Publishers should be paid up front from grant funds, so they will not need to engage in the practice of using copyright to implement toll-gates that restrict the public's access to FFSR results. Instead, publishers just need to receive a license to publish the FFSR articles from the authors or their institutions. In this way, articles can be made immediately available to taxpayers and the general public, the rightful copyright holders of the articles content. Having been compensated for their services, publishers will not need to further restrict access to readers.

For more suggestions on how to pay for open access, please see:
http://blogs.law.harvard.edu/pamphlet/2011/11/16/how-should-funding-agencies-pay-open-access-fees/.

This reference above discusses the topic of balancing library budgets with an open access payment fees taken from grants. It also includes discussion about funding libraries to play the role of archives, and as nodes in a decentralized system that facilitate access to FFSR results. This is after all, what Web technology was invented for at CERN.
[http://public.web.cern.ch/public/en/about/web-en.html]

Signatures

Name	Title	Institution
Luis Ibanez	Technical Leader	Kitware Inc.
Katie Osterdahl	Communications Specialist	Kitware Inc.
Katie Sharkey	Communications Specialist	Kitware Inc.
Arno Klein	Asst. Prof. Clinical Neurobiology	Columbia University
Tom Olijhoek	Scientific Consultant	Dutch Malaria Foundation
Pascal Christoph	Developer	hbz-nrw.de
Cameron Smith	Computational Scientist	Rensselaer Polytechnic Institute
Wesley Turner	Technical Leader	Kitware Inc.
Raphael Ritz	Scientific Officer	International Neuroinformatics Coordinating Facility (INCF)

Appendix A - Intellectual Property in Scientific Data.

The term of "intellectual property" is commonly used as an aggregate of the concepts of

- Copyright
- Patents
- Trademarks
- Trade secrets

In order to understand how these concepts apply to the challenge of maximizing access to the results of scientific research funded by the federal government, it is important to analyze the concepts independently.

Copyright is a government-awarded monopoly given to the creators of works of art. This monopoly awards creators the exclusive right to (1) reproduce the work, (2) prepare derivative works of it, (3) distribute copies of it, (4) perform it publicly and (5) display it publicly. The duration of copyright is: (a) the lifetime of the authors plus 70 year, (b) 95 years for works created by a corporation, or (c) 120 years for unpublished works created by a corporation. The goal of copyright is to provide an incentive to the creators of works of art by giving them exclusive rights on the exploitation of the works for a limited time

In the context of dissemination of scientific data, the economic bargain of copyright bears very low or no relevance, given that researchers (those who acquire and process the data) do not get paid when publishing that data. Instead, they get funded proactively for performing the research that leads to gathering information that is later published. Therefore, a very concrete economic incentive has already been provided and delivered to the researcher in the form of funding that American taxpayers have invested in the acquisition of the data.

As opposed to a novelist, whose income if purely based on the sale of copies of her/his book, the salary of a researcher is based on their performing the duties of scientific research. Granted, publishing datasets is part of such duties, but it is not equivalent to the creative activity of writing works of art (such as novels, music, or poems). Given that, in the context of FFSR, researchers are already paid by the public beforehand and so there is no need for the economic incentive of copyrights to address any "market failure" on the production of public goods (in the economic sense of non-rival and non-excludable goods), as is the case for novels, poems, and music. On the contrary, once the FFSR data has been acquired, every day that passes without this data being publicly shared is a day in which economic waste takes place and the economy at large performs less efficiently. It is also a day in which American taxpayers do not get anything back from the funds that they provided to the research enterprise.

Additionally, the nature of scientific research requires that the content of scientific datasets must be measurements of facts and should be devoid of any "creative elaborations". In other words, the more "scientific" a dataset is, the less "creative artistic content" it should have in it; therefore, the less it deserves the protection that copyright is intended to provide to creative works of

authorship. The creativity of the researchers lies in the definition of the acquisition protocols, the experimental design, and in the specific apparatus or software used during the data acquisition, which sometimes are made especially for a specific dataset. The dataset itself, on the other hand, shall not include any creative content. A high quality scientific dataset must be a concise collection of facts, measurements, and computations on those measurements. Datasets with high levels of "creative content" are by definition not scientific datasets, and should not be produced as the outcome of federally funded research, or any other process that aspires to be called "scientific".

Patents are government-awarded monopolies on the commercial exploitation of an invention. This 20-year long monopoly is awarded to the inventors in exchange for the public disclosure of the invention, and its eventual delivery (at the expiration of the patent term) to the Public Domain. Given that public disclosure is a requirement of the patent economic bargain, for awarded patents there is no concern about including information in articles intended for publication. The full information about the invention should already be publicly available at the U.S. Patent Office at the time that the patent is awarded to the inventors. Data is not "patentable subject matter" given that it is not the result of a creative process and is not useful, non-obvious, or novel. Datasets collected in the course of scientific endeavors are expected to be a collection of factual data, and therefore, they are as far as they can get from the type of "creative" work that patents are intended to protect.

Trademarks are symbols, designs, and terms that identify a product, service or company in the public marketplace. They are intended to prevent confusion in the marketplace, to protect the reputation of the producers of goods and providers of services, and to reduce the transaction cost that consumers have to invest in finding good and services that satisfy their needs. In the context of dissemination of scientific data, trademarks play a minimal role given that datasets are not supposed to be mechanisms of marketing goods and services. It is actually contrary to ethical standards in the scientific research field to use dataset publication as a venue for promoting goods and services in the context of commerce.

Trade Secrets refer to information that organizations keep confidential. For a piece of information to be considered a trade secret, it must have some value and derive part of its value from the mere fact of being secret. Trade secrets are managed via contracts, typically established between organizations in the form of non-disclosure agreements and between organizations and their employees in the form of confidentiality clauses that are incorporated in employment contracts. It is the responsibility of the institution to take affirmative steps to prevent its confidential information from becoming public.

In the event that a piece of confidential information is leaked publicly, there is no legal protection that can prevent the further dissemination of such information, except from forbidding an intruder to make use of data that was acquired illegally (e.g. by trespassing into private property). Therefore, in the context of dissemination of scientific data, trade secrets are only relevant as a context in which institutions should establish policies and verification mechanisms that prevent confidential information from being included in any dataset that is submitted for public release. It is the responsibility of the institution and its employees to protect such confidential information. Once data is published, the institution has relinquished its claim for such data to be considered a trade secret.