

Beth A Sullivan, PhD

beth.sullivan@duke.edu

Assistant Professor, Duke University School of Medicine

Associate Editor, PLoS Genetics, PLoS ONE

Durham, NC

Currently, only 8% of all scholarly papers are Open Access (OA). Thus, 92% of papers that are published are unavailable as a “market opportunity and size” for new innovative tools to start new business more jobs. Most importantly the majority of papers funded by tax dollars are not accessible to the public who has funded them. In my opinion, this is a serious breach of the trust between scientists and the American people who are the financial benefactors in scientific research endeavors.

Comment 1: NIH has implemented a public access policy that has proven to be extremely cost-effective at \$3.5-4.6 million per year. This means that an investment of approximately 1/100th of 1 percent of NIH’s overall budget results in access to 2.2 million scientific articles. The database for searching and accessing the articles is accessed daily by over 500,000 users. Most of these users are actually from OUTSIDE of education, emphasizing the demand for this information throughout the public sector.

Scientific research is an inter-dependent process in which each experiment is informed by the results of previous experiments. The scientists who perform research and the professional societies that represent them have a great interest in ensuring that research results are disseminated as immediately, broadly and effectively as possible. Electronic publication of research results offers the opportunity and the obligation to share research results, ideas and discoveries freely with the scientific community and the public. OA articles provide scientists with new tools (machine reading, computational tools, analysis programs) to obtain and read more information faster. Such new uses will encourage opportunities for commercial development through private investments to capitalize on a public resource. Importantly, OA expands the potential for interdisciplinary research that increase the breadth, depth and value of our public investment in science.

The type of OA is a crucial issue. In my opinion, there should be full OA, with free immediate access coupled with rights for full digital re-use. Mechanisms are necessary to enable full use, such as distribution, reuse, text and data mining, computations, and creation of derivative works. Appropriate licenses should be implemented, such as Creative Commons (CC-BY) licenses.

Comment 3: If a federal agency, such as NIH or NLM, maintains custody of all published content, it should make it easier to ensure accountability across the various institutes that funded the research. It would also assist grant reviewers in identifying publications during the peer review process to evaluate previous work and progress of grantees. In my opinion, distribution of information by a private source is not ideal, unless there is a federal mandate of how content of databases is structured and maintained, and curated.

Otherwise, distribution of information by multiple private sources is at risk of compromising standards and archiving of data. In other words, each source for archiving will set up the databases as they wish. It seems more logical that there be consistency in how the information is collected, tagged with identified, maintained and accessed.

Comment 4: Publishers are in an important area to participate in public/private partnerships by providing approved repositories that meet conditions for public accessibility, use rights, interoperability and long-term preservation of publicly funded articles.

Publishers need to commit to providing an open access option for any research article published in any of the journals they publish and declare a specific timetable for transition of journals to open access models. They should work with other publishers of open access works and interested parties to develop tools for authors and publishers to facilitate publication of manuscripts in standard electronic formats suitable for archival storage and efficient searching. Finally, they should ensure that open access models requiring author fees lower barriers to researchers at demonstrated financial disadvantage, particularly those from developing countries. This specific issue is going to be difficult to implement since many of the high profile biomedical journals still cling to impact factors as measures of standing and clout in the field. A broader issue as well is antiquated university tenure system that rates faculty progress and stature based on journal impact factors rather than article impact metrics. Article levels metrics measures how often an article is downloaded and cited in other scholarly works. Article level metrics (ALMs), such as those adopted by the PLoS journals, appear to be a more accurate method by which to evaluate the impact and reach of a specific journal article, rather than simply citing the impact of journal as a whole. ALMs could also be directly linked to grants and funding agencies to emphasize and track the direct impact of federal funding by a) individual investigator, b) research initiative, c) research topic, or d) technology (to name a few).

Comment 5: Metadata affords ways to enable specific actions and parameters, rather than basic item description. Parameters could include journal name, conference, publication title, researcher name, researcher grant code, institution publication data, and other identifiers such as subject/topic, technique/technology or method. Provenance metadata is a distinguishing feature of this effort as a means of facilitating specific desirable actions for use, reuse and analysis of publications. It should be a priority to develop machine-readable/machine-operable metadata for rights of downloading and reuse.

Comment 6: For OA models to work effectively, there must be consistency of requirements and mandates for depositing peer-reviewed literature across all funding agencies. Consistency in policy will maximize returns to taxpayers by ensuring complete results are widely available in a timely fashion. It will also minimize duplication of efforts in depositing literature that is coming from various publishing houses/publishers. Policies should integrate articles with grants managements systems, so that there is better agency accountability and transparency between the type of science that was funded, the level at

which the research was funded, and ultimate publication outcomes. The policies should create opportunities to create and maximize productivity management tools for federal and internal reporting.

Comment 7: In my opinion as basic scientist, if an investigator's laboratory research is federally funded, that means that most of his/her activities and those of the lab are supported by taxpayer dollars. Thus, honing an area of expertise over years/funding periods is afforded by federal funds. This expertise is called upon for writing book chapters and it make sense that encyclopedia entries or book chapters should also be publicly available. In fact, reviews and book chapters will be more intellectually accessible to the public since they are usually written broadly and with the intent to educate.

Another topic that might be considered (perhaps in future discussions of public access and scientific accountability) is archiving of all data from federally funded projects. Publications resulting from federally funded projects only account for a fraction of the actual data generated from the grant or fellowship. Many experiments result in unexpected or unexplainable findings or even failures that are not suitable for publication but could be useful to the broader research committee. In other words, these experiments, while considered "scientific dead ends" or failed experiments within a specific line of research, the data could aid another researcher who is approaching a similar question or using the same technology. The benefit derived by a subsequent researcher could be in the form of "don't make the same mistakes we did in this experiment" or "we attempted this protocol/technique multiple times by changing a, b and/or c". Since ALL of the research performed under a specific grant is federally funded, it might be useful to the broader scientific community to have a searchable repository of experimental findings that were too brief or incomplete for a proper journal publication. It will increase accountability and also illuminate to lawmakers and the general public the amount of trial-and-error that is involved in basic science research. Of course, this endeavor would need to be well managed, would require enormous amounts of storage, and benefit from proper curation of experimental data/results to prevent such a database from becoming a waste heap for useless information.

Comment 8: Immediate access to publications funded by federal resources is the ideal option and will optimize scientific and commercial utility of information contained in these articles. However, there should be an option for investigators to select a longer embargo period, particularly if the research is directly linked to a patent application.