Article

CHORUS: A Solution for Public Access to Scholarly Research

Howard Ratner

Background

As the first service of the nonprofit CHORUS Inc., the Clearinghouse for the Open Research of the United States (CHORUS, www.chorusaccess.org) offers an open-technology platform to meet the public-access needs of US federal funding agencies, researchers, institutions, and the public. CHORUS is focused on five principle sets of functions: identification, access, discovery, preservation, and compliance.

CHORUS was established in response to a US Office for Science and Technology Policy (OSTP) memorandum in February 2013 (www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf). The memorandum stated that each federal agency with expenditures of more than $100 million in research and development had to develop a plan to support public access to research, including "results published in peer-reviewed scholarly publications". It called for the establishment of public–private partnerships to facilitate access, and it encouraged leveraging existing infrastructure rather than creating new, expensive ventures.

CHORUS represents a good-faith effort by the publishing industry to partner with funding agencies to comply with the terms of the OSTP memo. CHORUS facilitates public access to peer-reviewed publications after a determined embargo period (where applicable) for each discipline and agency. By leveraging existing tools—such as CrossRef, FundRef, and ORCID—CHORUS allows research funds to remain where they are most needed: in the hands of researchers, funding research. There is no substantial cost for agency use or participation in CHORUS.

By focusing on open standards and an open architecture, CHORUS is a scalable solution that offers maximum efficiency for all parties by automating as much of the process as possible. That saves researchers time and effort and minimizes the investment of funds and the continuing work necessary by funding agencies and publishers.

CHORUS identifies articles that are reporting on federally funded research and enables readers to access the “best available version” free of charge via the publisher. The best available version could be either the accepted author manuscript or the version of record. CHORUS launched in a pilot phase in September 2013, and the production phase will begin in early 2014.

Stakeholders

CHORUS has identified many key stakeholders and has been working closely with each group to meet its needs:

• Funding agencies want to meet OSTP guidelines, measure grantee and agency compliance with the OSTP memorandum, show how the agencies’ investments are having an effect (return on investment), and provide the widest possible access to articles that report on the research funded by the agencies.
• Researchers want to obtain funding for their research, comply with their funding agencies’ requirements, know the sources of funding in their fields, and have access to the best available version of content in their fields.
• Librarians want to have access to the best available version of content for their patrons, conduct text mining and data mining, have confidence that their articles will be readily available in perpetuity, help researchers to comply with funding-agency requirements, and build discovery tools for researchers.
• The public wants to have access to the best available version of content to research a problem or drive economic development, see what the government is funding, learn the effects of specific agency grants, understand the latest developments in science, and have content connected to learning tools.
• Publishers want to help their authors and institutions to comply with funder mandates and retain traffic on journal Web sites to demonstrate value to their customers.

How Does It Work?

CHORUS enables the identification of the appropriate article reporting on funded research, the discovery of the article on the publisher’s site, public access to the article, and preservation of the article in a long-term archive. CHORUS tracks each of those features and reports the level of compliance on its dashboards.

When researchers submit a paper to a journal through a typical electronic submission system, they will receive a pull-down menu to identify the funding agencies that supported their research and an opportunity to fill in specific grant IDs. That is all the effort required by the researchers; they have done their part and are now fully compliant with funder requirements. The paper then goes through the publisher’s regular peer-review process and production cycle and is published. The paper’s DOI and associated metadata—including the grant information—are then registered at CrossRef, which in turn feeds a variety of metadata services, including FundRef and CHORUS (Fig. 1).

The article is then made publicly accessible by the publisher’s host system after the funding-agency embargo period expires or immediately if an author has paid an article-processing charge (APC). Article reuse terms are posted by the CHORUS-compliant publisher and made transparent by the CHORUS services and application program interfaces (APIs). Users will have access to the best available version, either the accepted author manuscript or the version of record. Access is always through the publisher for either route to public access; this ensures that the reader is seeing the paper in the context of the journal in which notifications of updates, corrections, or retractions will be available (Fig. 2).
The publicly available articles will be indexed in the same way as any other journal article and can be discovered in a user’s favorite search engine, such as PubMed or Google. CHORUS provides an open API available to those discovery tools to allow them to enhance their search results with information on the funding source and availability of the articles. The open API also enables funding agencies to create their own discovery tools and build institutional portals. New text- and data-mining services are also enabled via the API (Fig. 3).

For our pilot program, CHORUS worked with the US Department of Energy (DOE). More than 4000 records have been harvested via the CHORUS system into the DOE’s “PAGES” agency portal, and the DOE’s discovery tool offers search results that point back to individual publisher Web sites via the DOIs of the papers reporting on funded research. This provides clear evidence that CHORUS is a workable solution for funding agencies.

When the paper is published, the publisher deposits a copy of the version of record into one or more of what are called dark archives for preservation. The dark archives include Portico, CLOCKSS, and any other archive required by funding agencies. The paper is permanently archived in these repositories but not made available unless compliance trigger events occur.

Trust is an important part of the CHORUS approach, so a great deal of work has been put into ensuring compliance. The CHORUS dashboard service allows stakeholders to monitor compliance constantly and provides automated mechanisms to make sure that what has been promised is happening. The dashboard keeps track of the numbers of articles identified, the numbers of articles preserved, the numbers of articles publicly accessible, and the numbers of articles for which there are agency-accepted reuse licenses.

The dashboard automatically checks to see whether papers that are supposed to be publicly available are indeed publicly available. If they are not, a trigger event occurs, and a notification is sent to the agency and the publisher. If the problem is not fixed promptly, the dark-archived version of the paper is brought to light and replaces the publisher’s version in the system until public access is restored.
The dashboard will evolve and can be customized for specific needs and reporting requirements. The open dashboard data can also be pulled into an agency’s, institution’s, or publisher’s own system via the CHORUS API.

International Relevance and Scale
CHORUS was designed to meet the specific needs of the US OSTP memorandum, but it was deliberately built on open standards and designed for interoperability. CHORUS does not prevent or damage the existence of any other repository or archiving strategy, and the open API can be used to enhance their efforts. CHORUS works equally well with gold and green open access, and the open standards used offer tremendous potential for further development.

Although CHORUS is in its launch phase, international scalability is an important future direction. The very name of the not-for-profit organization behind CHORUS, CHOR Inc, was deliberately chosen to reflect these plans and our hope that this framework can reach far beyond the United States.

Conclusion
CHORUS is a rapidly growing non-profit solution for public access to scholarly content. It is an open-technology solution focused on the identification, access, discovery, preservation, and compliance of scholarly content reporting on funded research and is internationally scalable. It keeps researcher funding in the hands of researchers and is as inclusive as possible. CHORUS addresses the public-access needs of funding agencies, researchers, institutions, publishers, and the public. To achieve maximum effectiveness, it needs maximum participation from the publishing community. Participating is easy. Find more information at www.chorusaccess.org.

Acknowledgments
Special thanks to David Crotty, Oxford University Press; Mark Doyle, American Physical Society; Fred Dylla, American Institute of Physics; Susan King, American Chemical Society; Alice Meadows, John Wiley & Sons; and Joe Serene, American Physical Society, for all their contributions.

Discussion
Volunteer editors undertake large responsibilities for overseeing the peer review of scholarly journals. Publishers of the journals are struggling to remain competitive while being mindful of the pressures put on volunteer editors. In our case, anecdotal evidence from the editors regarding the bottlenecks in processing time was inaccurate.

Manuscript-submission systems are increasingly sophisticated with respect to generating data reports. Many systems provide “canned” reports that show high-level statistics that are useful in some types of reporting; however, it is difficult to track specific concerns by reviewing only canned reports.

The Editor and Reviewer Performance Report has completely changed the conversation at editorial board meetings. The CEs and their AEs have data that show where the bottlenecks are in the process. Arming the editors with this information is critical for allowing them to do what they do best: solve problems.

From the publisher perspective, having these reports and more granular data has highlighted weaknesses in the process. More attention is being paid to late papers, and coordinators are more aggressive in contacting editors who appear to be struggling to keep up. Journal managers who perform a more detailed analysis will have data to either support or refute editor perceptions on their performance activity. Furthermore, by providing specific turn-around times for smaller pieces of the tasks involved with peer review, the editors have data for valuable discussions about how to manage the workload.

References