Implementing Standards: Data and Data Exchange in Scholarly Publishing

MODERATOR:
Tony Alves (iD)
Director of Product Management
Aries Systems
North Andover, Massachusetts

SPEAKERS:
Michael Di Natale (iD)
Business Systems Analyst
Aries Systems
North Andover, Massachusetts

Gabriel Harp (iD)
Senior Product Manager
Cell Press
Cambridge, Massachusetts

Susan King (iD)
Executive Director, Rockefeller
University Press
Chair, CHOR Inc
New York, New York

REPORTER:
Colleen M Sauber (iD)
Editor, Instructor, Biomedical
Communications
Mayo Clinic
Rochester, Minnesota

Acronyms in scholarly publishing continue to proliferate. The initiatives they stand for represent the increasing interchange of scientific data and publishing support, transparency, and availability. Information collected during manuscript submission is translated into identifiers in standardized formats readable to digital systems. Information about funders of an author's research, for example, could be shared and the author's work tracked and recognized.

Speakers in this session assured attendees that the arrival of these and other systems is helping establish standards that benefit authors, journals, and the science community. Such systems can remove the "need to reinvent the wheel for each separate process," one speaker noted.

Susan King, Executive Director, Rockefeller University Press, and Chair of CHOR Inc., reported that nonprofit CHORUS, born in 2013, was built from existing infrastructure and standards, including CrossRef Open Funder Registry and ORCiD. The open-access policy-agnostic service currently comprises 46 publishers who work with 6 agencies, including the National Science Foundation and U.S. Department of Energy.

CHORUS, a suite of services and best practices, assists agencies and publishers to provide public access to published articles about funded research in the United States. Publishers collect funding details from authors during submission, and Rockefeller University Press journals mine acknowledgments to capture this information in a standardized format. CHORUS presents metadata and links through DOIs to the full text on publishers' websites. Its search application allows users to find the latest research articles through agency portals and common search engines. CHORUS takes in funder identification with CrossRef and verifies the metadata, content, and funder, as well as importing ORCiD data. CHORUS also sends data for archiving.

King stressed that active engagement with each federal funding agency is important. For the agencies, CHORUS supplies email alerts and establishes a dashboard application that assists the monitoring and reporting of funders and publishers. It enables ingestion of bibliographic funder and license information to enrich agency portals.

Next, Gabriel Harp, Senior Product Manager at Cell Press, spoke about CRediT, a cross-organization initiative that involves academic institutions and funders, among others. Harp echoed what various presenters cited during the conference: The number of authors per manuscript has increased markedly. "The traditional author list does a poor job in conveying what each of these people has done," Harp said. CRediT was created to address the fact that author lists cannot fully convey the complex and varied contributions that go into a scholarly work.

Harp explained the CRediT taxonomy of contributor roles (used primarily for biomedical and life sciences). After an initial workshop in 2012, the taxonomy was piloted on recently published papers in 2013. In 2015 Cell Press added CRediT to its author guidelines, encouraging but not mandating that authors use the taxonomy. Cell Press now mandates that authors provide a contributions statement near the Acknowledgment section. This statement is "a simple flat paragraph of text," Harp said. "In the long run, we want to go way beyond that." In fall 2015, Cell Press surveyed the first 100 papers that used CRediT. The 38 respondents recognized CRediT's potential value for standardization, and 76% found CRediT useful in accurately reflecting authors' contributions.

Today, ORCiD is building CRediT into its registry, Harp reported. Next steps are "sharing our learning with others; surveying authors who have chosen not to use CRediT; and to begin tagging the CRediT roles in the article XML."

Finally, Michael Di Natale, Business Systems Analyst at Aires Systems, described his company's endeavors with JATS (Journal Article Tag Suite). JATS is run in Aries's Editorial

CONTINUED

Manager (EM) for web-based manuscript submission and peer review and in ProduXion Manager, its production tracking tool. It sends and receives data during submission, facilitates manuscript imports from sister publications, and delivers editorial files to a third-party production system and to other services, including vendors. In this way, JATS can respond to multiple needs for data portability during publishing and beyond, and can be a standard for data exchange.

JATS supports ORCiD, Ringgold Institution IDs, and CrossRef, soon to be joined by CRediT, in a machine-readable arrangement in XML format for article metadata. JATS can import submission files from journals not in the EM system, possibly moving the submission to an open-access journal. Thus EM can communicate with any system that supports JATS. Di Natale explained the details from JATS XML include the corresponding author's name, address, and affiliation; ORCiD and Ringgold ID data; coauthors' information; funding information through CrossRef; and manuscript metadata, including title, sections, and supplementary pieces.