



Standards and the Scholarly Publishing Ecosystem

By Richard Wynne (<http://orcid.org/0000-002-9217-0407>) and Alison O'Connell (<http://orcid.org/0000-0002-9032-3983>)
Aries Systems

Well, that's probably the world's most boring title for an article—maybe I should have entitled it “Five things you need to know about sex and food that can keep you healthy while watching sports on TV”! But, if you're still reading, I'm impressed, and I hope that you're rewarded with some useful insights, even if I stick to the subject of standards in publishing.

The Scholarly Publishing Ecosystem

A couple of months ago, I was invited to give a presentation at a workshop attended by librarians. One of my slides illustrated the main steps in manuscript workflows, including the use of vendors to create structured XML from author manuscript files and compose first page proofs. One of the librarians, in all sincerity, asked me, “What is that, and why do publishers want to do that?” Naturally, I explained, and the librarian was enlightened and grateful.

I tell this story because it illustrates how ignorant most journal users and buyers are about the supply chain that delivers the manuscript to their eyeballs. It takes a global ecosystem of organizations and services to create the final product. Kent Anderson of the Scholarly Kitchen publishes and updates a useful list of functions publishers perform.¹ Many of these steps are complex and require coordination between multiple independent organizations and systems, each with a particular specialty.

But here's the problem: while the scholarly publishing ecosystem thrives with pockets of rich innovation, it is hugely inefficient when it comes to the transfer of data and content. Frequently,

data has to be rekeyed, time is lost by incompatible transfers, and users (including authors) become frustrated because they are bounced between different systems. There is way too much friction.

One solution to the problem might be to build one big publishing system that everyone uses. Believe it or not, there are several organizations that actually plan to do this. To them, I say: good luck.

Personally, I prefer a vision for the future where multiple independent organizations compete and innovate in specialized areas, but where they are nevertheless able to efficiently cooperate to serve authors and readers. Agreeing upon data standards makes this integration possible and removes the friction.

Fortunately for us, we are working at a time when viable data standards have emerged and are about to transform the scholarly publishing industry. As publishing professionals, we should understand and ride this wave.

To help illustrate the impact and potential of standards, I discuss three use cases that show how standards are being used to solve real-life pain points.

Use Case 1: Account Maintenance and Single Sign-On

Ever since the first publishing systems were launched, users have complained about the need to register for and log into different systems. Solving this problem has been a structural challenge because organizations naturally do not wish to cede user credentialing to competitors. While some companies have tried to build in-house solutions, such parochial approaches miss the key user need—the ability for users to sign in to multiple independent publisher systems.

The ORCID identifier (ORCID) is principally promoted as a solution for contributor disambiguation. However, the ORCID API (application programming interface) also allows this

1 K. Anderson. 82 Things Publishers Do (2014 Edition). Available at <http://scholarlykitchen.sspnet.org/2014/10/21/updated-80-things-publishers-do-2014-edition/>. Accessed November 10, 2014.

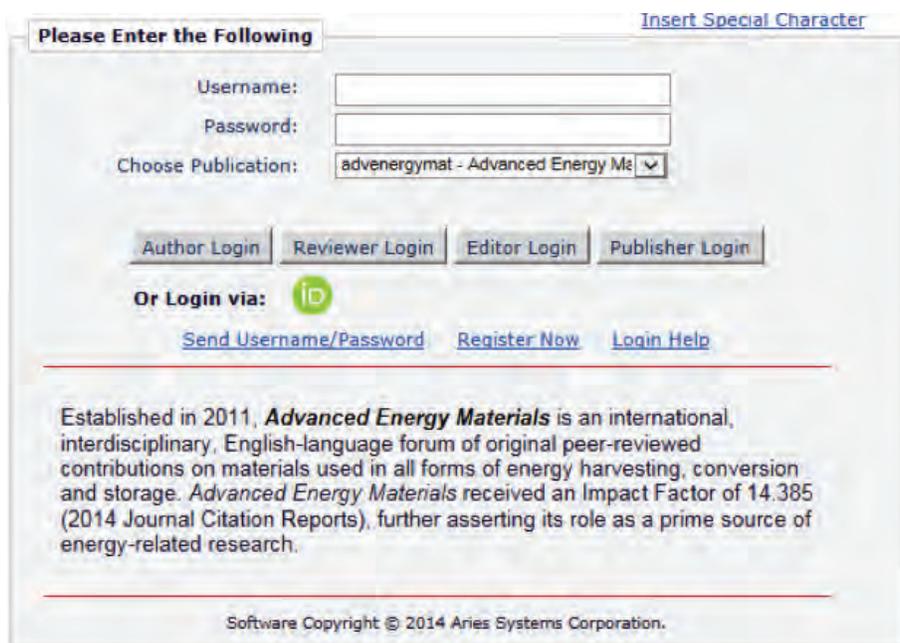


Figure 1: ORCID login option for EM

identifier to be used by systems like Editorial Manager (EM) as an independent authentication system. For example, several publishers, including Maney Publishing, Cambridge University Press, Wiley VCH, DeGruyter, and the American Psychological Association, have all activated ORCID single sign-on for EM. This means that authors can sign into EM deployments for all these publishers using the same credentials. I expect many additional publishers to activate this capability in EM, and other systems, in the near future.

As other vendors and other applications (*e.g.*, membership management systems) activate the ORCID API, this will further accelerate the benefits delivered to end users.

Use Case 2: Manuscript Transfer and Editing Services

Authors complain about having to resubmit rejected manuscripts. Reviewers are also frustrated about re-reviewing manuscripts that were rejected elsewhere. The obvious solution to these problems is to implement seamless manuscript transfer so that authors don't have to resubmit, and prior reviews can be passed along to recipient journals. In theory, this sounds easy; in practice, implementing seamless manuscript transfer involves many details. For example, did you ask the reviewers for permission to transfer their review? Does

the editor recommend one or more journals for the author? How do you match disparate submission data elements?

However, the emergence of standards has begun to simplify this problem. JATS (the Journal Article



Figure 2: Ingest services for EM

Tag Suite from the National Library of Medicine) provides a standardized XML format that can be shared between journals. The use of ORCiDs ensures that authors and reviewers can be identified with greater fidelity in the transfer process.

A similar problem exists for manuscript language editing services (*e.g.*, Cactus, Editage) or collaborative authoring platforms (*e.g.*, Overleaf). These services already store manuscript data and author information that could potentially be used to streamline submission. The emergence of standards such as JATS and ORCID now makes this possible. For this reason companies like Aries have launched “ingest” services so that third parties can directly submit manuscript data and files in standardized formats.

Use Case 3: Open Access Payment Processing

The societal and political demand for Open Access (OA) has resulted in the emergence of new business models to support the cost of article publication. Initially it was thought that a single APC (article publication charge) paid by the author would be the dominant model. However, it has become increasingly apparent that processing single APCs is inefficient from the point of view of the payer, payee, and indeed the entire system.

The reality is that most APCs are actually paid by institutions or research funding organizations. Using the author as an administrative intermediary in this commercial transaction introduces unnecessary friction and is a waste of the author’s precious time. Some publishers have identified this as an opportunity and implemented new models. For example, the Royal Society of Chemistry’s “Gold-for-Gold” initiative offers free APCs to institutions with subscriptions; and BioMed Central offers APC discounts to member libraries.

As OA business models proliferate, the back office “plumbing” needed to administer fees needs to be flexible and efficient. That’s where data standards become so important. With standardized metadata it becomes possible to administer and change business models with minimum disruption to workflow.

For example, by integrating the Ringgold institutional identifier into manuscript submission, the APC terms can be modified based on the institution(s) of the author(s). Thus, if the principal author is from MIT and the OA publisher has a membership arrangement with MIT, then the APC can reflect the terms of that membership.

Similarly, by capturing the FundRef identifiers during manuscript submission, the payment processing system can notify authors of any license limitations imposed by research funders such as Wellcome Trust.

EM is the first peer-review system to have collaborated with the Copyright Clearance Center’s RightsLink to leverage peer-review metadata in this way. Several publishers now process APCs through RightsLink with standardized data via FundRef, Ringgold, and ORCID passed out of EM through a standard API.

Conclusion

The scholarly publishing ecosystem is large and teeming with innovation, but the impact of innovation is frequently muted because of an inability to connect the dots. Emerging data standards (along with routine software mechanisms such as APIs) enable greater collaboration and integration to better serve authors and readers.

I hope that I’ve communicated the immediacy of the transition that is taking place right now. Both understanding and leveraging this transition offer significant career opportunities to ISMTE members.