



ORCID in Book Workflows: Report and Recommendations

*Bill Kasdorf, VP and Principal Consultant
Apex Content and Media Solutions*

April 30, 2017

Revision History

Rev Num	Release Date	Changes	Author
Rev. 1.0	April 30, 2017	Initial Release	Bill Kasdorf, Apex
Rev. 1.1	May 10, 2017	Two small clarifications requested by an interviewee, no change to meaning; deletion of interviewee emails in the appendix to enable public distribution of report	Bill Kasdorf, Apex

Contents

1.0	Executive Summary	1-4
1.1	The Survey	1-4
1.2	The Interviews	1-4
1.3	Key Issues and Recommendations	1-5
1.3.1	Book vs. Journal Workflows.....	1-5
1.3.2	Updating Systems	1-6
1.3.3	The Discipline Factor.....	1-6
1.3.4	Distinguishing Types of Books.....	1-6
1.3.5	The Backlist Problem.....	1-7
2.0	The Survey	2-8
2.1	The Respondents.....	2-8
2.2	Use of ORCID by Book and Journal Publishers	2-8
2.3	Use of ORCID Specifically in Book Workflows	2-9
3.0	The Interviews.....	3-11
3.1	Publishers	3-11
3.1.1	Factors Contributing to Lack of Implementation.....	3-11
3.1.2	Perceptions of Value	3-12
3.2	Technology and Workflow Systems Providers	3-13
3.2.1	Manuscript Submission and Peer Review Systems	3-13
3.2.2	Title Management Systems	3-14
3.2.3	Editorial Management Systems	3-15
3.2.4	Production Workflow Management Systems	3-16
3.3	Service Providers.....	3-16
3.3.1	Content-Related Services.....	3-16
3.3.2	Metadata-Related Services	3-18
3.4	Standards Organizations.....	3-20
3.4.1	EDItEUR.....	3-20
3.4.2	NISO	3-21
3.5	Libraries and Library-Related Organizations	3-21
3.5.1	The British Library	3-21
3.5.2	OCLC	3-22
3.5.3	KB, the National Library of the Netherlands.....	3-23
3.6	Funders	3-23
3.6.1	The Mellon Foundation.....	3-23
3.6.2	The NIH.....	3-24
3.6.3	The Wellcome Trust	3-25

4.0	Key Issues and Recommendations	4-26
4.1	Book vs. Journal Workflows	4-26
4.2	Updating Systems.....	4-27
4.3	The Discipline Factor	4-27
4.4	Distinguishing Types of Books	4-28
4.5	The Backlist Problem	4-29
5.0	Appendix: The Interviewees	5-30
5.1	Commercial Publishers	5-30
5.2	University Presses/Nonprofit Publishers	5-30
5.3	Technology and Systems Providers.....	5-30
5.4	Service Providers.....	5-31
5.5	Standards Organizations.....	5-31
5.6	Libraries and Library-Related Organizations	5-32
5.7	Funders	5-32

1.0 Executive Summary

Recognizing that the use of ORCID iDs and their associated metadata records—by authors, publishers, and others in the publishing ecosystem—has thus far been almost exclusively in the context of journals, ORCID contracted with Apex to explore the extent to which ORCID is used in book¹ workflows and to make recommendations to promote and facilitate this.

In the context of virtual meetings with the ORCID in Books Community Working Group (CWG), it was determined that this could best be accomplished by two vehicles:

- A survey that could be sent out to a wide variety of recipients and which would contain questions that would be designed for systematic tabulation and analysis.
- A set of 25-some interviews with the CWG members and a carefully selected group of publishers and related organizations that could provide more wide ranging and nuanced insights.

1.1 The Survey

Although the survey was too small to be statistically significant, and although responses were mainly from Europe and North America, it did provide useful insights into the use of ORCID in book publishing. The respondents were well distributed as to the size of their book publishing programs, and evenly distributed between STM and HSS publishers.

Unsurprisingly, the survey indicated very little use of ORCID in book publishing workflows. However, many of the respondents also publish journals, and for them, the use of ORCID is surprisingly evenly distributed throughout the workflow. Of those currently using ORCID, 72% use it in acquisition and peer review; 58% use it in metadata management (e.g., disambiguating names in metadata); 50% use it in identity management in editorial and production workflows; 47% use it in citations and linking; and 33% use it in researching the work of contributors or potential contributors.

When asked specifically about the value of ORCID in book workflows, the answers were quite different, with 86% seeing value in disambiguating names in metadata and 73% for citations and linking; 50% saw value in identity management in editorial and production workflows; and only 32% saw value in acquisitions and peer review and researching the work of contributors or potential contributors. This is reflective of the significant differences between book and journal workflows.

1.2 The Interviews

Twenty six interviews were conducted involving 39 interviewees. Eight of the interviews were with publishers; the others were from carefully selected representatives of significant aspects of the book publishing ecosystem where the use of ORCID is relevant, such as technology and

¹ In keeping with the mission of ORCID, the focus for this engagement was limited to academic books.

systems providers, service providers, standards organizations, libraries and library-related organizations, and funders.

Of the publishers interviewed,

- Only half of them are using ORCID in some way in their book workflows.
- Of those who are, all but one rely on manual insertion of the ORCID iD into metadata.
- Of those who are not, half are in the process of implementing ORCID in at least one system.

The lack of ORCID implementation in publishers' book workflows was not unexpected, but the lack of awareness of ORCID on the part of authors was striking. It was clear that most humanities scholars are largely unaware of ORCID, and it is in the humanities where books, rather than journals, are the key to professional advancement. Because ORCID is primarily designed to benefit authors, the lack of pressure from authors on book publishers retards its implementation.

Nevertheless, the publishers interviewed did recognize that ORCID has value for books, citing the following benefits:

- The value of matching authors across book and journal systems.
- Simplifying sign-on into systems.
- Improved discovery, citation, and marketing.
- Benefits in peer review.
- Empowering semantic systems.

Publishers also often expressed that the systems they use in book workflows are not currently capable of doing more than simply storing an ORCID iD in metadata. Ironically, the interviews with technology, systems, and service providers made it clear that they consider it straightforward to incorporate good ORCID functionality into their systems, but that they are not hearing sufficient demand for this from their clients. In fact, all of those technology, systems, and service providers expressed frustration about this, because they would very much like to make better use of ORCID.

1.3 Key Issues and Recommendations

The interviews identified a few key issues affecting the implementation of ORCID in book workflows, which resulted in concrete recommendations that ORCID could pursue to remedy the situation.

1.3.1 Book vs. Journal Workflows

Book workflows are very different from journal workflows. They are typically much less automated and much less integrated, using a collection of siloed systems or vendors with little or no interoperability, and often involve no direct contact by authors.

RECOMMENDATIONS:

- Focus on the points in the book workflow where there is the most direct author contact and the most perceived benefit: acquisitions or marketing.

- Focus initially on publishers of both books and journals, where some systems could be leveraged and where awareness of ORCID benefits is likely to be greater.

1.3.2 Updating Systems

Updating systems to properly interoperate with ORCID is an issue both for publishers and for the vendors or developers of the systems that they use. The obstacles to doing this are not primarily technical, but lack of demand, lack of urgency.

RECOMMENDATIONS:

- Focus on title management systems, the metadata management systems commonly used by book publishers. Work with them to help them see ORCID implementation as a competitive advantage and help them to educate their customers.
- Publicize successes to create demand. Don't rely on the ORCID website; scholars won't come and look. Get the word out to where they are by promoting publishers and organizations who successfully promote ORCID to their constituencies.

1.3.3 The Discipline Factor

The lack of awareness of ORCID in the humanities and social sciences is a significant problem. Strategic initiatives with organizations of humanities scholars and social science researchers will be necessary to remedy this situation.

RECOMMENDATIONS:

- Explore interoperability with MLA Commons and its peers so that ORCID records can be created as a byproduct of the identity management systems most critical to scholars and already used by them.
- Create a boilerplate white-label document making the case for ORCID that societies and humanities and social science publishers can use to promote and facilitate ORCID registration.

1.3.4 Distinguishing Types of Books

Most discussions of ORCID for academic books focus on monographs, including serials and contributed volumes. But ORCID implementation is much more likely in conference proceedings and encyclopedias and other large reference works, which have much more journal-like workflows.

RECOMMENDATIONS:

- Work with societies in fields where conference proceedings are of prime importance for professional advancement, such as engineering.
- Work with publishers that publish large volumes of conference proceedings or large reference works to help them implement ORCID functionality.

1.3.5 The Backlist Problem

Many book publishers have large backlists. The authors of those books do not have ORCID iDs and are unlikely to obtain them. This is particularly important in the humanities, where research may involve scholarly literature that goes back many decades. This is also an issue for libraries and other large repositories of academic book content, like JSTOR and Portico. These organizations often use other identifiers, such as ISNI, or name authority files, to manage identities. There are small collaborative initiatives underway, but a more comprehensive look at the place of ORCID in the broader context of author identity management could result in significant uptake in the use of ORCID for books.

RECOMMENDATION:

- Convene a meeting—virtual if necessary—with OCLC, ISNI, the BL, the KB, and JSTOR. The interviews indicated a strong interest on the part of all of these groups in exploring ways to partner with ORCID to better address the interoperability of their systems with ORCID and the need for more comprehensive identity management. These organizations have very significant technical resources and infrastructures to contribute to a solution that could result in significant benefit to ORCID in the book world.

2.0 The Survey²

ORCID sent the survey to its own members and promoted it on Twitter. In addition, three major scholarly publishing organizations publicized it to their members: the Society for Scholarly Publishing (SSP), the Association of Learned and Professional Society Publishers (ALPSP), and the International Association of Scientific, Technical, and Medical Publishers (STM).

2.1 The Respondents

There were 85 responses to the survey, 67 (79%) of which were from publishers of books or organizations serving the book publishing industry. Most questions were answered by 50-some respondents, almost all of whom appear to be publishers rather than publishing related organizations such as service providers. Although this is clearly too small a sample for the results to be statistically significant, the responses do provide useful insights into the perception and adoption of ORCID in academic publishing.

The majority of responses were from Europe and North America. Of the 56 respondents who answered the location question, 26 (46%) were from Europe and 20 (36%) were from the US and Canada; four (7%) were from Asia; two each (4%) were from Africa and Australasia; and single responses were received from Central/South America and from the Middle East.

The range of the annual book volume published by the respondents was more evenly distributed. While 27% of the respondents publish twenty or fewer books annually, 25% publish more than 500; 19% publish 20-50, 17% publish 50-100, and 12% publish 100-500.

Thirty-one percent were commercial publishers, closely followed by 29% from non-profit publishers; 16% were university presses, 4% were government publishers, and a surprising 20% classified themselves as “other.”

As it happened, there were an equal number of respondents who publish STM books and those who publish HSS (humanities and social sciences) books: 61% each, with some obviously publishing both. Thirty-two percent publish professional books (e.g., law and medicine); 21% publish in higher education; and 11% publish technical documentation.

Thus apart from the geographical imbalance (which is a reflection of the geographical distribution of members in the organizations surveyed), the sizes and types of publishers responding provide quite a good cross section of the academic book publishing industry.

2.2 Use of ORCID by Book and Journal Publishers

Interestingly, although fifty of the respondents indicated awareness of ORCID, only four answered the question about whether they thought that such an identifier could be useful in their acquisition, editorial, production, or marketing workflows. Nevertheless, 56% of 52 respondents said they currently use ORCID in their “book and/or journal workflows.” It can only be concluded that the

² The results of the survey are provided in the PowerPoint presentation “ORCID in Books Survey” that accompanies this report.

question about the usefulness of ORCID confused the respondents with regard to whether it was about only books or about both books and/or journals, and so most skipped the question. It is no surprise that most users of ORCID use it for journals and almost none do for books.

Twenty-seven respondents answered the question about where in their workflows ORCID is used (thus are likely those who publish journals). The responses to this question were quite encouraging: the use of ORCID is surprisingly evenly distributed throughout the workflow. Of those currently using ORCID, 72% use it in acquisition and peer review; 58% use it in metadata management (e.g., disambiguating names in metadata); 50% use it in identity management in editorial and production workflows; 47% use it in citations and linking; and 33% use it in researching the work of contributors or potential contributors.

Purposes for which respondents are planning to use ORCID but are not yet doing so were, logically, complementary to the purposes for which it is already used: 75% plan to use it in researching the work of contributors or potential contributors, 60% for citations and linking, 55% for identity management in the workflows, 47% for disambiguating names in metadata, and 33% for acquisition and peer review.

These responses thus indicate that those respondents not only see value in ORCID, they are aware of its value in many aspects of their work: all of the aspects addressed by those questions were clearly valued.

2.3 Use of ORCID Specifically in Book Workflows

Keeping in mind that the respondents to the previous question were likely the users of ORCID for journals, the next question specifically asked where ORCID was *perceived* as having value specifically in book workflows. The responses there were quite different, with 86% seeing value in disambiguating names in metadata and 73% for citations and linking; 50% saw value in identity management in editorial and production workflows; and only 32% saw value in acquisitions and peer review and researching the work of contributors or potential contributors.

This reflects the significant difference between the workflows in books vs. journals, discussed in more depth in section 4.1 below and borne out in the response to the question regarding whether the systems used in book workflows are capable of incorporating ORCID iDs: 47% said their editorial and production management systems are capable of this; 40% said their manuscript management and peer review systems are; only 19% said their marketing systems (e.g., systems generating their ONIX feeds) are; and 37% answered “don’t know” or “other.”

Even fewer were aware of any capabilities their systems have for interfacing with ORCID via APIs. Thirty-one percent said their manuscript management and peer review systems are capable of this; 21% said their editorial and production management systems are; only 2% said their marketing systems are; the majority, 57%, said “don’t know.”

A strong majority, 83% of respondents, said that they see a benefit to contributors to their books for associating their ORCID iDs with those books. This would indicate that while book publishers do see benefits to their own operations in the use of ORCID, they recognize that the primary benefit is to the contributors. Nevertheless, a strong 58% did recognize that it would be beneficial to be able to identify or update other information on their contributors’ ORCID records.

Finally, it is no surprise that most respondents (78%) identified manuscript submission and peer review as the touchpoint where they could best obtain ORCID iDs; 49% cited production and proofing, another phase where there is direct interaction with authors; and 43% cited editorial; only 14% cited marketing.

Again, the small sample size of this survey should be kept in mind when considering these results.

3.0 The Interviews

Apex conducted 30-to-60-minute interviews with 26 organizations across the publishing ecosystem: commercial publishers, university presses, technology companies, standards organizations, service providers, libraries and library-related organizations, and funders. Nine of those organizations involved more than one person in the interview, resulting in 39 individuals being involved in the discussions. (The organizations and individuals interviewed are documented in the Appendix, section 5.0 below.)

3.1 Publishers

Eight publishers were interviewed: three large commercial academic publishers and five large, mid-sized, and small university presses.

- Only half of them³ are using ORCID in some way in their book workflows.
- Of those who are, all but one rely on manual insertion of the ORCID iD into metadata.
- Of those who are not, half are in the process of implementing ORCID in at least one system.

3.1.1 Factors Contributing to Lack of Implementation

All the publishers interviewed observed that there is very low awareness of ORCID in book publishing. Even those who have substantial journal programs that very actively use ORCID find that awareness of ORCID is low on the book side of their business.

The most significant reason cited for this was that professional advancement is dependent on journal publication in STM fields but on books in the humanities and social sciences. This divergence is responsible for a number of factors that result in not just lower *awareness* of ORCID for books but a perception of lower *value*.

- Whereas a researcher may publish many journal articles in a given year, a humanities scholar typically takes years to produce a single book.
- Because journal publishing is article based, the volume of transactions needing to be managed demands automated systems, most notably manuscript submission and peer review management systems. Most book publishing is still largely managed by manual systems in which the need for the ORCID iD is much less compelling.
- This is compounded by the fact that far more manuscripts are submitted than accepted in journal publishing, whereas much scholarly book content is commissioned.
- Book metadata tends to be at the title level; there is often little or no metadata at the chapter level, even for contributed volumes.⁴

³ Because of the small sample sizes involved in the interviews, proportions are indicated in general terms rather than with specific numbers to avoid creating a false impression of precision.

⁴ It should be noted that even in journals, it is most often only the corresponding author's ORCID iD that is collected and maintained, not those of all of the contributors to an article.

- Most publishers saw little value in displaying an author’s ORCID iD in books themselves. This is due to the fact that unlike journal content, most book content—although discovered and obtained online—is not actually consumed online.
- The automated systems that book publishers do use—typically “title management systems” that manage metadata, budgets, schedules, royalties, etc.—are not yet capable of anything more than simply storing the ORCID iD in a field. They are not yet typically capable of authentication or accessing a contributor’s ORCID record. The book publisher is typically dependent on an external commercial system vendor to provide such capabilities.
- Unlike journal systems, the systems used in book publishing typically do not involve direct access by authors.
- Whereas the focus for journals—especially STM journals—tends to be on current or recent research, humanities scholarship often involves research across a deep historical corpus of work by scholars who do not have ORCID iDs or did not have them at time of publication.

These issues will be discussed in more depth in section 4.0, “Key Issues and Recommendations,” below.

3.1.2 Perceptions of Value

Despite the low implementation of ORCID in book workflows, the publishers interviewed generally acknowledged that it has value. Most stressed that the value for book publishing would be greater for the authors than for the publishers. However, the following benefits to the publishers were acknowledged by some of the interviewees.

- *The value of matching authors across book and journal systems.* This benefit tended to be cited by the larger publishers and those who have both journal and book programs. Several lamented that currently they lack a systematic way to discover journal articles they have published that a book author contributed to. This capability would facilitate suggestion features on their websites as well as aiding acquisition, editorial, and marketing processes.
- *Simplifying sign-on into systems.* The ORCID iD is seen as having obvious value for enabling authors to simply provide their ORCID iD and then for the publishing system to retrieve complete and trustworthy metadata required from that author. While this is seen as primarily a convenience to the author, this would also reduce the burden on editorial staff, who typically manually enter such metadata about authors, or copy and paste it from a questionnaire sent to authors (which also has the drawback of lacking the authentication provided by ORCID).
- *Discovery, citation, and marketing.* Several publishers recognized that an author’s ORCID iD would have value in facilitating discovery, citation, and metrics—especially altmetrics—that would increase the visibility of book content as it currently does for journals.
- *Benefits in peer review.* Several publishers said that it would be beneficial for them to be able to search across the ORCID database to identify peer reviewers qualified to review specific manuscripts (although other interviewees pointed out that there are better, more discipline-focused resources for this). One publisher cited the benefit of linking their system to Publons via the ORCID iD to provide credit to peer reviewers.

- *Empowering semantic systems.* As more publishers develop knowledge graphs, such as Springer Nature’s recently announced Scigraph, that enable the mapping of complex relationships not only within a publisher’s own corpus but with the entire publishing ecosystem, unambiguous identifiers for authors are as important as DOIs are for content.

It was striking, in conducting the interviews, how often the discussion stimulated awareness of potential benefits from ORCID that the interviewees had not considered. One book publisher interviewed commented that access to a potential author’s ORCID metadata would be “a goldmine for acquisitions,” that knowing what areas a potential author is expert in would be of “enormous interest,” and that this capability would be “absolutely useful” in finding qualified peer reviews: “That functionality would be extraordinarily useful to us.” Another interviewee in that same interview commented that ORCID would be “incredibly valuable in marketing” and that their marketing department is likely unaware of this.

That publisher’s journal division does an exemplary job of explaining the value of ORCID, facilitating its use by contributors, and making it easy for a contributor without an ORCID iD to get one. This is reflective of the gulf that still exists between the journal and book operations of most scholarly publishers who publish both. These silos are beginning to be broken down; the implementation of ORCID can contribute to that improved integration.

3.2 Technology and Workflow Systems Providers⁵

The interviews with technology and systems providers confirmed the lack of implementation of ORCID within book workflows. However, whereas the publishers cited that one factor impeding this was that the system vendors have not updated their systems to facilitate the proper use of ORCID, ironically, the system vendors cited the lack of demand from the publishers, despite the fact that all of those interviewed felt that implementing ORCID in their systems would be quite straightforward.

3.2.1 Manuscript Submission and Peer Review Systems

The interview with Aries, whose Editorial Manager is the leading manuscript submission and peer review management system, and who also offers a complementary production management system, was illuminating. Although Editorial Manager is focused on and primarily used in journal publishing, Aries can “skin the system” to make it more book-like, as it has done for a leading medical textbook publisher. The interviewee observed that although much of the needed functionality is the same across journals and books, the terminology is different; they offer a “books language pack” to address this.

They are also in the process of developing book-specific functionality: they are developing a full-featured book production system, which is where automated functionality is more needed in book publishing than manuscript submission and peer review automation. This system will enable a publisher to require an ORCID iD to be fetched from the ORCID systems to discourage cut-and-paste; they feel that authenticated ORCID iDs are extremely important. In fact, that

⁵ The portions of this report that identify specific organizations or individuals were shared in draft form with those interviewees and updated where appropriate, based on their responses, for accuracy.

interviewee pointed out the importance of repeated authentication: it is otherwise too easy to assert an invalid ORCID iD. “We can bulletproof our system but we can’t guarantee that other systems it interacts with are bulletproof.”

In their systems, wherever an ORCID iD appears, it is linked to that contributor’s ORCID record. Their philosophy is not to retrieve and store that information, which could become outdated and could be misused, but to provide direct access to it as needed.

Because it was necessary to limit the number of interviews in the scope of this engagement, Aries was chosen to represent providers of this type of system. But the conclusion that any of the competing systems could be similarly updated—and may in fact be in the process of being updated—to facilitate proper use of ORCID iDs and ORCID metadata is likely warranted. For publishers who use these manuscript submission and peer review systems and similar production management systems, there should ultimately be little or no barrier to implementing ORCID in their book workflows—as long as they communicate the need for this to those vendors.

3.2.2 Title Management Systems

The same message was clear in the interview with the representative title management system. Unlike manuscript submission and peer review systems, title management systems are common in book publishing. Fundamentally, they are metadata management systems, used to capture bibliographic, administrative, and technical metadata associated with books such as information about contributors, vendors, and other business partners; budget and schedule information and tracking; rights and royalties information; and bibliographic and supply chain metadata, specifically the ONIX feeds that are sent to retailers, aggregators, and other recipients. Klopotek is the largest of these, with a particularly strong presence in large STM publishers. Several of the organizations interviewed in this engagement use the Klopotek system.

The Klopotek system has long supported ORCID—it can support any identifier, capturing both the value and type. A very modular system, Klopotek stores this information in its “business partners” module where a host of other information about a given business partner is stored, thus making it available to all other modules and subsystems, for example the production management module or the module that manages contracts, rights, and royalties. This enables an ORCID iD to be associated at the work level or the product level or both (since a given book is often managed as a number of different print and digital products).

The interviewee, Klopotek’s Head of Product and Solution Management, is very knowledgeable about ORCID and sees great value in integrating Klopotek systems with ORCID, despite the fact that they are not working with ORCID at the present time. He said that he was very interested in using the ORCID API and providing the ability to integrate the publisher’s data with ORCID data: “This is how I would like to go forward in product development. This is highly important in science. It is not a big thing to build this. It would be a huge value to a publisher.”

When asked whether this is actively being planned, he said that it is not on their current development roadmap. The reason, he said, is lack of customer demand: customers have not asked for this. Although publishers can simply capture the ORCID in the system, Klopotek hasn’t been asked for better integration, despite the fact that it would not be hard to do. This could change if they got customer demand.

Although Klopotek is only one of the title management systems used by academic book publishers, it is this writer’s informed opinion that a similar story would be heard from the others.

3.2.3 Editorial Management Systems

The leading editorial system in scholarly publishing is Inera’s eXtyle. Their CEO, Bruce Rosenblum, a thought leader who has made major contributions to the scholarly publishing ecosystem, was interviewed for this project.

Although eXtyle occupies only one segment of the full publishing workflow—fundamentally, the copyediting and XML conversion stage—it handles ORCID well. Instead of requiring ORCID iDs to be typed or copied-and-pasted, it retrieves ORCID iDs from the upstream manuscript submission system, synchronizes the submission system authors with authors in the Word manuscripts, and then automatically integrates iDs both in the Word file for the copyeditor and the underlying XML that is ultimately used for publication. Multiple layers of integrity checks in eXtyle ensure that the right iD is always associated with the right author. This system also allows for seamless and automatic merging of additional author-specific submission system metadata as CRediT and funding data.

However, he lamented that ORCID is not used well in general, even in the journal sphere. He observed that while publishers “struggle with whether or how to display the ORCID iD,” this issue is a red herring: “[ORCID is] fundamentally about the metadata, not about the display.”

“The vast majority of journal publishers are not *using* ORCID, they’re just *collecting* it,” he contends. Worse, when it is used, it is usually copied and pasted from a submission system to the final XML file, without any automated integration, which risks “mangling it or assigning it to the wrong author.” He also observed that ORCID is not well integrated into workflows: it often never gets beyond the submission system, thus “never making it to the final XML and to Crossref.” He noted that automated workflow integration of ORCID is challenging for small- and medium-size publishers, and often their vendors resort to inexpensive human labor (i.e. copy/paste) which technically invalidates ORCID authentication.

He is very skeptical that books are where ORCID should be putting its energy because the funding and promotion models for books are so different than those for journals. For example, “CHORUS doesn’t work without ORCID and FundRef,” which creates a huge financial incentive. His motto is “follow the money.” The perceived benefits for books—disambiguating and evaluating authors and reviewers—are social drivers, not financial drivers. Adoption would be much more likely for conference proceedings, on which some disciplines base professional advancement, or publications (such as government reports) based on funded research, where the funders will require it. He also noted that conference proceedings have a much greater representation of DOIs in Crossref than books, which indicates that publishers of conference proceedings may be more likely to adopt new identifiers than book publishers.

3.2.4 Production Workflow Management Systems

Justin Gonder, Product Manager of the Access & Publishing Group of the California Digital Library (CDL), was interviewed about their development, in partnership with the University of California Press and the Collaborative Knowledge Foundation, of Editoria, a robust open source content and workflow management system designed for scholarly monographs. Funded by a Mellon grant, it is expected, upon its planned full public release in 2018, to provide a free, open source⁶ system to the scholarly community, including university presses and library publishing programs, that can be installed on a publisher's own systems or hosted on their behalf by others.

Although ORCID integration is not in fact in the initial release of Editoria—again reflecting the low profile of ORCID in the book sphere⁷—it is of high interest to the CDL. The metadata about contributors available via ORCID would be of “enormous value,” especially the ability to “pull in an author's publication history.” Tools for citation building were seen as a good use case for ORCID, as was the use of the ORCID iD for login, Single Sign-On, etc. It was seen as especially useful with regard to authors outside of a given publisher's organization.

Mr. Gonder stated that it is likely that ORCID will be implemented in future releases of Editoria. Its value for authorization and authentication is obvious. Editoria is also designed to interoperate with title management systems, so its ability to populate those systems with authenticated ORCID iDs and related metadata would also be valuable. He observed that whereas title management systems are maintained by publishing staff, the Editoria system would be more likely to involve authors directly.

He also observed that library publishers, in contrast to commercial publishers, are very well positioned to work with ORCID because of their work with name authorities and related systems.

3.3 Service Providers

3.3.1 Content-Related Services

Three content hosting and delivery platforms were interviewed, providing contrasting and complementary perspectives:

- *ITHAKA*, which includes JSTOR, a digital library of academic journals, books, and primary sources with a focus on the humanities and social sciences, and Portico, a digital preservation service archiving over 77 million publications, including over 500,000 books.
- *OpenEdition*, a European provider of four related Open Access platforms for books, journals, conference proceedings, and events in the humanities and social sciences from over 70 publishers.
- *Atypon*, one of the largest hosting platform providers, currently used by over 200 publishers to host 11,400 journals, 24 million articles, 187,800 ebooks, and 998 websites.

⁶ Note that the fact that Editoria is free and open source does not mean it is limited to Open Access (OA) publishing: it will be available for the development and maintenance of paid content as well.

⁷ Mr. Gonder commented that he has never seen an ORCID iD associated with any books in CDL's eScholarship repository.

ITHAKA

At one end of the spectrum, JSTOR is seeing very little use of ORCID: “not even much in journals, just a trickle, and next to nothing—probably nothing—in books.” This is due to the preponderance of humanities content on their platform: “In STM, when you publish, as part of a group, 20-25 papers a year, it is very different than in HSS, where an author may publish a paper every year or two. And they don’t have funders wanting to track their work because there aren’t any funders!”⁸

The fact that JSTOR is a deep archive creates another issue: “We have a lot of dead authors in JSTOR,” for which ORCIDs do not exist. When ORCID iDs are provided in metadata supplied to them they will not discard them. Although this is minimal today, they stated that they “can actually see a day (when we start seeing some more traction from ORCIDs coming in from publishers via their digital submissions) that we go the rest of the way and start offering ORCIDs as a service *for* our publishers, much like we do with DOIs today for some of our publishers (some others register themselves and submit them).”

They definitely see value in associating ORCIDs or other reliable identifiers with content in both JSTOR and Portico for research purposes: “Researchers want to track people across scholarship, especially in the humanities.” Being able to disambiguate authors “would enable some very interesting science.” This is of high interest in Portico. There is “a massive amount of content in the Portico archive—400 terabytes,” including both HSS and STM content. The complication with Portico is that publishers “signed up for a dark archive.” It would probably be necessary to anonymize content; using reliable IDs under the hood could enable some very significant research.

OpenEdition

OpenEdition, on the other hand—although much smaller and newer than JSTOR and Portico—sees an urgent need for ORCID. In the OpenEdition Books platform, author names are currently simply a plain text field in metadata. They recognize that the ORCID iD is “critically important in the platform.” Because they host content across 70 publishers, it is very important for them to be able to reliably identify content in the platform by a given author that may have been published by different publishers.

OpenEdition is also a key part of the [HIRMEOS](#) project, “High Integration of Research Monographs in the European Open Science infrastructure,” which is devoted to integrating monographs into the European Open Science Cloud. More than simply a repository of Open Access content, HIRMEOS aims to bridge across disciplinary silos and build connections across the entire corpus. HIRMEOS is not another platform; it is a services layer across several platforms, presently including OpenEdition, OAPEN, University Press Göttingen, EKT, and Ubiquity. ORCID iD is the contributor identifier that will bridge across all five platforms, as will DOI for documents and FundRef for funders.

Currently, OpenEdition has DOIs but not ORCID iDs, which is why ORCID is such a priority. They also plan to use ORCID iDs to identify annotators and commenters.

OpenEdition also provides a content management system, Lodel, that is designed to facilitate the creation of long, complex texts in digital form for HSS content. ORCID integration is in development for Lodel, to enable the ORCID iD to be used at both the content and metadata levels.

⁸ See the observation in section 3.6.1: In the humanities, most of the projects funded are organizations, not people.

Atypon

A striking aspect of the interview with Atypon was the difference between the *capabilities* of their platform, Literatum, and the *use* of those capabilities by publishers. As Marty Picco, VP-Product Management, put it, “We give them the tools, but how they use the tools is up to them.”

Atypon’s goal is to “wrap the site around the user” to make the site a great experience. So they incorporate a host of features to enable that. Specifically with regard to ORCID in books, “there’s no reason we can’t do everything for books that we do for journals.” If the ORCID iD is in the XML submitted by the publisher, it’s used, both in metadata and in published content—depending on the publisher’s design and specifications. Publishers can also submit a data file that associates ORCID iDs with contributors and articles.

When an ORCID iD is displayed, it’s always an actionable link. Some publishers use this to enable linking to the contributor’s ORCID profile; others do not. Authors could log in with their ORCID iD and be recognized as a contributor without needing a password. Publishers systems could provide enhanced author services based on the ORCID, e.g., targeted calls for reviewers. The ORCID record is “a good signal of what an author is interested in”; publishers could use it to build a “fingerprint” of an author.

While the key use of the ORCID iD is disambiguation, Atypon also uses it in search results; the more ORCID iDs are used, the better search gets. Determining the intent of search could be enhanced by the ORCID iD. Mr. Picco’s ultimate vision is for “deep connectedness”: “When we get away from dead PDFs, we can really create a knowledge graph; ORCID can be key to that.”

Mr. Picco is a strong advocate of making the ORCID iD mandatory: “Publishers should require or facilitate it. I can’t think of a good reason not to require it.”

3.3.2 Metadata-Related Services

In addition to service providers who are primarily focused on content (which of course involves metadata as well), there are also key service providers that do not work with the content, focusing only on metadata. The three interviewed for this project were:

- *Crossref*, a nonprofit membership organization originally founded to enable cross-publisher linking, and which now provides a host of related services on a robust infrastructure that is based on bibliographic metadata—and particularly the Crossref DOI, which is now considered indispensable across the scholarly publishing world.
- *The Copyright Clearance Center (CCC)*, a leading Reproduction Rights Organization (RRO) that facilitates rights transactions and provides rights-related services globally. CCC is also considered indispensable across scholarly publishing (as well as publishing in general).
- *Altmetric*, the leading organization providing altmetrics, monitoring the conversation about scholarly research online and in social media to gauge the level of interest and impact published research is receiving.

Crossref

Crossref reports that there has been little use of ORCID for books; at the time of the interview, there were approximately 600 books or book chapters with an ORCID iD out of over 825,000

books (only 0.7%) and nearly 12,000,000 book and book chapter DOIs in their system.⁹ They speculate that these may mostly be for serials, which are journal-like.

They see the benefits of ORCID as being the same for books and journals. The main benefit, from the Crossref point of view, is auto-update. When an ORCID iD is submitted, Crossref updates the ORCID record of that contributor, providing they have permission to do so. Thus far, approximately 150,000 authors have given that permission, which they consider disappointing: fewer than half of the eligible authors have given this permission.

Many of the comments made by Crossref were also made by others, and have been reported above. One very interesting one that no other interviewee mentioned would be the value of relating ORCID iDs to organization IDs. It would be useful to be able to keep them in synch in order to keep affiliation information current. They have joined ORCID and DataCite to form an informal working group to look at the potential development of a community-wide organization identifier.

Chuck Koscher, Crossref Director of Technology, called attention to the difficulty of capturing ORCID iDs in the production process and into the metadata downstream: “It is slow going and hard to accomplish.” He stressed the importance of getting the ORCID into the metadata pipeline. Ultimately, the goal is to be able to map from a work with a DOI to a person with an ORCID iD affiliated with an organization with an ID. “Only ORCID can disambiguate the person.” He observed that metadata systems are more and more engaged with each other, but using only a very primitive, text-based method; identifiers are essential. “If books don’t get on the bandwagon, they’re going to become even farther behind in the ecosystem.”

Copyright Clearance Center

The interview with CCC was very illuminating, especially in light of Mr. Rosenblum’s previous dictum to “follow the money” (see section 3.2.3 above). With the proliferation of Open Access, CCC is actively involved in processing Article Processing Charges (APCs). This process significantly benefits from the precision provided by the combination of ORCID iDs for authors, FundRef for funders, and Ringgold IDs for organizations. “APCs are very specific to specific authors, specific publishers, and specific funders; lots of things can be tied back to an author’s ORCID iD.”

This is “of incredible value to OA.” Not only does this enable proper crediting of authors, it also enables properly acknowledging authorship to funders who ultimately pay the APC. This also ties to pricing and discounting.

When publishers integrate their submission system or production system with CCC, authors “land in the CCC workflow,” seeing the “persistent metadata bar” which prominently features the ORCID iD along with the funder ID and the institution ID.

Jennifer Goodrich, CCC’s Director of Product Management, made another important point not heard in the other interviews: how valuable it would be if ORCID records contained *society IDs*. “That would be tremendously useful to access via APIs.” Society affiliation is important because APCs can be discounted differently based on society. Big publishers represent many societies. “I would love it if they would include this in the ORCID record! This information is not centrally available anywhere.”

⁹ A useful report on Crossref statistics, including ORCID iDs and ORCID Auto-Update permissions and data, can be found at <https://data.crossref.org/reports/statusReport.html>.

She also pointed out that there are many other products at CCC where recognizing authors is useful; for example, waiving of licensing fees. They get lots of reuse requests and “being able to validate authorship claims would be fantastic.”

Another example: the ORCID iD as the basis for Single Sign-On (SSO). Authors work across many systems and may have many logins—to Aries, ScholarOne, the society’s membership site, RightsLink, etc. “The ORCID iD could solve so much frustration in the industry.”

It should not go unnoticed that this project has shown virtually unanimous support for the ORCID iD, including for books and book chapters, by such organizations and systems providers—along with unanimous lamentations about the lack of ORCID use, particularly in books. This opportunity will be discussed in section 4.0, “Key Issues and Recommendations,” below.

Altmetric

Altmetric provides what they characterized as “light touch integration of the ORCID iD.” Their Explorer platform gives a user everything Altmetric has access to. Entering an ORCID iD triggers an API call and retrieves all the DOIs associated with that ORCID iD. Currently, ORCID iDs need to be entered individually; they would like to enable a publisher to be able to submit a list of ORCID iDs.

The opportunity for Altmetric to provide ORCID access to the Altmetric data directly from the ORCID platform was offered, but this was not pursued by ORCID.

Currently, Altmetric doesn’t collect metadata at the person level; its focus is on published content. However, there is no technical barrier to integrating ORCID. For example, they could integrate with Crossref for ORCID functionality. They haven’t done that yet because there aren’t enough ORCID iDs. It wouldn’t be hard to do, but it is not on their roadmap.

3.4 Standards Organizations

The comments from the two standards organizations interviewed, EDItEUR and NISO, largely echoed the perceptions reported above from other interviews. This makes sense, because these organizations are not themselves involved in its use. But they both have good insight into related practices because of the prominence of their standards in the publishing ecosystem.

3.4.1 EDItEUR

Graham Bell, Executive Director of EDItEUR, confirmed that the ORCID iD can and should be included in ONIX feeds from book publishers: it is ID type 21 in ONIX, having been available since 2011. Thus there is not only a place to put the ORCID iD in the standard format for book supply chain metadata, there is a way to identify it in a machine-readable way *as* an ORCID iD.

Tim Devenport¹⁰ pointed out that the ORCID iD is especially valuable in serials and in contributed, multi-author books.

¹⁰ Mr. Devenport, Publishing Consultant and Product Manager at EDItEUR, was also recently named Executive Director of the ISNI International Agency.

Mr. Bell remarked that the need for disambiguation via the ORCID iD downstream is obvious in scholarly and academic publishing—giving the right credit to the right person is essential—but that it actually has little or no value in production. Thus there is a problem with the ORCID iD, which may be available on manuscript submission, flowing downstream to where it is of most value. He also observed how fragmented book publishing workflows are, consisting of various single-purpose systems. “Crossover is often badly handled—processes are too siloed, and systems frequently don’t communicate with each other.” Too often, plain text is what is easy to capture and reuse (if it is reused), but metadata and markup are lost.

3.4.2 NISO

Todd Carpenter, the Executive Director of NISO, said that the use of ORCID in book workflows is “modest in the extreme.” He gave the example of two recent books he contributed to: he provided the publisher with his ORCID iD but nothing was done with it.

He pointed out that although the production of a book isn’t that radically different from the production of a journal, historically there is “an aversion to looking at them in the same way.” Journals use standard formats, strict process management, and other systematic methods that could be very useful for book production; but these are seldom used for books.

He said that “Using the ORCID iD to streamline communications, rights management, etc. could be a huge benefit.” He predicted that shrinking margins in academic publishing will help drive book publishers to more standardization.

Another factor: “information sharing is key today.” There is great value to a contributor’s metadata in the ORCID record, especially because the contributor curates the information and keeps it up to date.

He pointed out that there is a serious lack of recognition of chapter authors for books: they are not listed as authors in the book’s metadata. The use of ORCID would help remedy this. This is also very important in reference/citation processing, which is increasingly done by machine: “an ID is *much* more important than a text string.”

He ended the interview on an optimistic note: “If you only look at monographs as they are today, you might say ‘why bother?’ But if you look at where they’re going, it’s a different story.”

3.5 Libraries and Library-Related Organizations

The three library organizations interviewed provided an especially interesting perspective and revealed an intriguing opportunity, building on work ORCID is already doing, that could have a significant impact on the ORCID-in-books issue.

3.5.1 The British Library

The British Library has been the recipient of legal deposits of electronic publications of all types since 2013, and has partnered with organizations throughout the supply chain to develop workflows to streamline ingestion of this content and its available metadata. Neil Wilson, the Head of Metadata Services at the BL, remarked that although they have not seen much use of ORCID in books, “if the ORCID iD were there, we would use it.”

The “dead author problem” is a significant issue, though. While it would be ideal to have ORCID iDs for all authors, this is not only impractical, it is impossible, given the design and purpose of ORCID. The BL has become an ISNI registration agency; they could assign ISNIs not only to authors lacking an ORCID iD (the vast majority), but also to those who do have them. It would be very beneficial for the publisher to supply ORCID iDs, which could facilitate this being done unambiguously. Linking the ORCID iD and the ISNI to “cluster” information would be highly desirable; the BL would even consider resupplying such enhanced metadata back to publishers.

“Bridging” is a key value of the BL. They maintain Authority Control databases and are embedding ISNIs into their workflows, including Cataloging in Publication, “so ISNIs are there right up front.” By then putting the ISNIs in the British National Bibliography, they will be able to provide an automated look-up service. This is a very big project, involving millions of catalog records. “Once we establish a trusted connection between the ISNI and the ORCID iD we can do a lot.”

They are working on the EThOS database of e-theses, beginning to link to the ORCID iDs and batch assigning ISNIs to enable future bridging: “We are trying to catch them [authors] early, wherever they are.”

He also remarked that rights organizations are another key player, echoing the comments from the CCC, and that there is interest outside of the library world: for example, Apple, in addressing the issue of micropayments, needs identities unambiguously. So establishing identities is critical. “Libraries have been working with controlled lists of authors for a very long time. Getting at the ORCID metadata would be very valuable in the establishment of authority control.” They sometimes wind up writing to the author to ask “are you this person?” From the library point of view, what ORCID could provide “would be *very* valuable.” The BL has also provided databases for research funders in projects, where identity management is key.

Finally, he pointed out that there are now multiple versions of books over time, and the way the author names are represented is not consistent. The ORCID iD could be invaluable in assisting the BL to associate all of them properly.

3.5.2 OCLC

OCLC was very eager to talk because they are working on a project to assess the use of name and person identifiers. They maintain VIAF, the Virtual International Authority File, which combines multiple name authority files into a single OCLC-hosted name authority service. They see VIAF as complementary to ORCID and ISNI because of its “deep historical representation” due to the participation of National Libraries. This represents the biggest collection of metadata about authors in existence.

One of the fee-based services OCLC offers for both libraries and publishers is the creation of MARC cataloging records. To date, none of their customers have requested them to put ORCID iDs in MARC records. OCLC would be willing to provide ORCID iDs in MARC records if they received such a request.

If OCLC gets an ID, they can map it to other IDs. They are not offering services along these lines currently, but they are very interested in providing services around names. OCLC would be quite interested in pursuing a conversation around a partnership with ORCID in this area. This

could go a long way to addressing the “dead author problem” cited by Brill, JSTOR, and others in this investigation.

What’s more, OCLC is responsible for maintaining ISNIs. This is done in their Leiden office, where they maintain the infrastructure for ISNI.

They commented that “the library community is becoming active in assigning identifiers to individuals.” OCLC provides a lot of infrastructure to facilitate this. They observed that “impact factors for monographs are almost impossible to determine.”¹¹ OCLC would like to “create a bridge between publishers and libraries.”

3.5.3 KB, the National Library of the Netherlands

The KB started a year ago to update their workflows related to everything discussed in this section. As the legal depository for the Netherlands, their first priority was to give all authors in the depository an ISNI. They chose ISNI because they needed a persistent ID; ORCID iDs, being self-registered, can get out of date; and VIAF and other authority information is being exported to ISNI. So ISNI appears to be the most useful “hub” identifier.

Mapping from the Netherlands Thesaurus of Authors (NTA) to ISNI is complex. In the past few months, they have been concentrating on publications deposited in 2016-17, which are done in real time on deposit from publishers. They capture all name variants in NTA and ISNI. Other Dutch university libraries will probably make their own connection to the ISNI database to request ISNIs.

There are 13 university libraries in the Netherlands, and the original intention was to issue ISNIs for all faculty. This exposed an important point: that the KB is a library, not a university. Libraries have a bibliographic reason for ISNI, whereas universities have a professional and research reason for the ORCID iD. It therefore makes more sense for the KB to concentrate on ISNI in the context of legal deposit and the universities to concentrate on ORCID for faculty.

Martin van Muyen at the organization responsible for ISNI in the Netherlands, OCLC (Leiden), is currently working on mapping ISNI and the ORCID iD. He’s exploring a possible workflow whereby when a scientist is requesting an ORCID, an ISNI can also be requested. The development of this ORCID-ISNI workflow is subsidized by the EU’s THOR project.

3.6 Funders

3.6.1 The Mellon Foundation

Don Waters, Senior Program Officer for Scholarly Communications in the Mellon Foundation, was interviewed because he is instrumental in the funding of a great number of important initiatives in the scholarly ecosystem. Two of those are in fact discussed above: JSTOR and the Editoria project from the California Digital Library and the University of California Press.

Mr. Waters stressed that although his comments are drawn from his experience at the Mellon Foundation, they should be attributed to him and not to the Foundation. Moreover, he

¹¹ Note the relationship of this comment to the work of Altmetric described in section 3.3.2 above.

emphasized that his comments focus mainly on the humanities. This focus in fact was a reason why his perspective is so important: because these interviews revealed little awareness of ORCID or use of the ORCID iD in the humanities, and because books are such a key factor in professional advancement for humanities scholars, getting insight into that sector is an important priority for this investigation.

Mr. Waters questioned an assumption fundamental to ORCID: that contributors will keep their ORCID records up to date. He cited what he characterized as “profile fatigue.” There are many places where a scholar’s profile currently needs to be maintained. He suggested that especially for humanities scholars, the most important profile is discipline-specific, for example MLA Commons. For humanities scholars, particularly those who are members of the Modern Language Association, this is where they will keep their profile up to date.

“There are disciplinary cultures. It’s better to go where the culture is rather than expecting ORCID to serve as the collection point.” Mr. Waters’ recommendation: “Figure out where humanities people keep their identities and go there.”

Developing interoperability between ORCID and MLA Commons, and compatibility with resources like LinkedIn, ResearchGate, etc., may be the best way to jump start getting humanities scholars into ORCID, and thus ORCID iDs associated with books.

3.6.2 The NIH

The perspective from Richard Ikeda, Director of the Office of Research Information at the US National Institutes of Health (NIH), is very different. The NIH is an extremely important funder of biomedical research in the US. They maintain PubMed, an open access database of 27 million citations for biomedical literature from MEDLINE, life science journals, and online books.

Although ORCID is critically important to the NIH and used extensively for journal contributors who have been funded by the NIH, the ORCID iD is “very rare in books,” and Dr. Ikeda has not seen it used for book chapters. (Only .1% of the citations in PubMed are to books.)

Dr. Ikeda places a very high value on the ORCID iD and provided another perspective not heard in the other interviews. The issue is the ability to track the impact of NIH funding *over time*.

He pointed out that by the time a book is published, its subject is typically in a mature state, having been the subject of many journal articles in the meantime, and is thus much less prominent to the NIH in the evaluation of a grant’s productivity.

Nevertheless, having the ORCID iD for books would be valuable—and especially for book chapters. Book metadata typically provides no metadata about chapter authors.

Because of public access requirements, the NIH gets good linkage to the journal publications reporting on the research they funded. But what they have no visibility into is what *other* work a researcher is doing—in the past, recently, and over time. “The value is in tracing people’s work over decades” in order to really understand the impact NIH funding had or has.

He provided a fascinating and high profile example in the area of gene editing. One of the two papers that is credited with providing the basis for CRISPR, the technology that is now so essential for gene editing, was done by a bacteriology researcher whose work the NIH funded. In fact, she was supported on an NIH training grant in 1986-1987 as a graduate student and has had

continuous NIH funding as a Principal Investigator since 1991. Thirty years later, it is clear that the impact of NIH funding has been enormous.

This is a particularly prominent example, but there are literally thousands of less prominent examples for which the ability to trace the impact of their funding would be extremely useful to them. ORCID could be a key to the solution. He also pointed out that public access to published literature by the NIH is solely for peer reviewed articles—not conference papers. Conference proceedings would be of particular interest to them. The ORCID record can give the NIH visibility into a researcher’s entire output, including journal articles, conference papers, books, and book chapters, and to see that output grow throughout that researcher’s career.

3.6.3 The Wellcome Trust

Robert Kiley, Head of Open Research at The Wellcome Trust, shares the biomedical perspective of Dr. Ikeda.

He made the ideal statement summarizing the purpose of this project: “It should be a no-brainer to encourage book publishers to use ORCID so that ORCID records can contain a complete record of all scholarly outputs.” He pointed out that, in addition to working to add ORCID IDs to monographs, work should also be done at the chapter level so that the authors of individual chapters can be uniquely identified.

The Wellcome Trust insists on ORCID iDs at the application process. “If you’re a Wellcome funded researcher, you get prompted at the start of the application process to supply the ORCID iD, and are provided a link to get one if you don’t have one. You can’t submit a grant application without an ORCID iD.” This is prompting other funding organizations to require the ORCID iD in grant applications.

He contends that the lack of ORCID in books is a timing issue: “First adopters were STM publishers but now it’s moving into other disciplines. Pretty soon it will be universal.” He expects that the pressure will come from researchers. It’s so easy for researchers to provide information via their ORCID record that they will be frustrated when the information in their ORCID record is not complete.

Wellcome uses the ORCID API to validate ORCID iDs and is working to both pull in relevant metadata from an ORCID record (e.g., publication data) and push grant data to successful grant applicants.

On the question of whether the ORCID database could be useful in finding grant reviewers, he said that this would be something they could consider downstream, but that systems are already in place to address this need.

4.0 Key Issues and Recommendations

The following are the key issues affecting the implementation of ORCID in book workflows that emerged from the survey and interviews. In one sense, these could be considered obstacles or barriers. Instead, these should be viewed as opportunities, revealing strategies by which current obstacles and barriers could be overcome.

4.1 Book vs. Journal Workflows

The differences between book and journal workflows were mentioned in virtually every interview.

Journals typically use standardized and often automated workflows, including the manuscript submission and peer review systems by which journal articles are first touched by the publisher. Increasingly often, these systems are integrated with downstream editorial and production systems. Plus, the metadata associated with journal articles is typically part of the article itself—in the header of the JATS XML file which has come to dominate journal publishing, not just for dissemination but now often integral to print and digital production.

Most book publishing workflows are far less standardized and systematic and far less automated. Books are often commissioned; the review process for manuscripts is typically manual (often managed by assistants working with email and spreadsheets); and the components of the editorial and production workflow are typically done by separate systems, and even separate vendors, with little or no interoperability. Authors have little or no interaction with these systems. The book metadata is even further divorced: it is typically a marketing function, using ONIX metadata that is completely separate from the digital book files themselves. The one end-to-end system used by some book publishers is the title management system. These don't work with the content, but just manage information about the content, parties, processes, budgets, and schedules.

RECOMMENDATIONS:

- *Focus on the point in the book workflow where there is the most direct author contact and the most perceived benefit: acquisitions or marketing.* The party to obtain the ORCID iD—and to provide a means to get one for authors who don't have one—is either the acquisitions or commissioning editor, who is working with the author to make the book successful, or the marketing department, in the collection of information that will be used to promote the book. In both cases, questionnaires are often used. Because of the much lower volume of transactions for books than for journals, obtaining ORCID iDs at these stages can even be done manually if necessary. Having those parties provide the author who lacks an ORCID iD with the appropriate link to ORCID would be easy and would benefit both the author and the publisher.
- *Focus initially on publishers of both books and journals.* Many interviewees remarked that the benefits of ORCID should be the same in books and journals. It will be much easier to get publishers who publish both to see the benefits of ORCID for their books. In some cases they will be able to leverage systems already in place on the journal side for the book side of the press. But be mindful that the workflows are quite different; assuming journal systems can simply add books would be a mistake. The interview with Aries (section 3.2.1) indicated this clearly. (See section 4.2 below regarding updating systems.)

4.2 Updating Systems

Updating systems to properly interoperate with ORCID is an issue both for publishers and for the vendors or developers of the systems that they use.

For vendors and developers, the interviews made it abundantly clear that there is little or no technical obstacle to doing this. Instead, the obstacle is lack of demand, lack of urgency. The interviews with all of the technology and system providers (see section 3.2 above) clearly indicated this. They are well aware of the value of the ORCID iD and they consider incorporating better ORCID-related functionality in their systems to be straightforward.

The same is true for publishers, whether they use these commercial systems, in-house development, or manual processes. One large university press that this writer has worked with extensively and knows to be highly technically capable acknowledged in the interview that the issue was simply priority: “There are just so many things on the roadmap for systems that ORCID does not rise to the top. Nobody would dispute the value of ORCID, but it doesn’t get priority over more urgent priorities.”

The engineering is not the obstacle; the perception of value on the part of the publisher is the obstacle.

RECOMMENDATIONS:

- *Focus on title management systems.* Although it is tempting to think that an effective marketing campaign focusing on the scholars and researchers for whom ORCID is the greatest benefit, or on the publishers who compete for those authors, would be the best strategy to create the necessary demand—and both of these strategies would be worth doing—from a practical point of view it may make the most sense to work more directly with the leading title management system vendors. There are only a few of them; finding out what would be necessary to get good ORCID integration onto their roadmaps, and helping them to see that doing so would give them, and their customers, a competitive advantage, might be the most effective and doable. If even a few of these make it virtually frictionless to incorporate ORCID well into book workflows, which shouldn’t be difficult for them, this could create a perception and eventually an expectation that this is necessary and a standard part of the book publishing process. These systems typically generate the ONIX metadata that goes out into the book supply chain; in any of these discussions, don’t neglect the advantage to marketing and the benefit of raising the profile of authors.
- *Publicize successes.* Look for every opportunity—joint press releases, vendor webinars, conference presentations, etc.—to tout those systems or publishers that do a good job of incorporating ORCID into their book workflows. Simply putting these on your website is not sufficient. You need to push this information out to get people to pay attention to it.

4.3 The Discipline Factor

Another common theme in the interviews was the lack of awareness of ORCID in the humanities and social sciences. This is particularly true in the humanities, where monograph publication is as critical to a scholar’s professional advancement as journal publication is to an STM researcher.

Don Waters from Mellon made a very astute observation (see section 3.6.1): that especially for humanities scholars, the most important profile is discipline-specific, for example MLA Commons. “There are disciplinary cultures. It’s better to go where the culture is rather than expecting ORCID to serve as the collection point.” Mr. Waters’ recommendation: “Figure out where humanities people keep their identities and go there.”

RECOMMENDATIONS:

- *Explore interoperability with MLA Commons and its peers.* It could actually be quite straightforward to simply issue ORCID iDs for scholars in the MLA Commons database; however, that would be contrary to the fundamental design of ORCID and could be seen as too aggressive by scholars. But discussion of interoperability would very likely bear fruit. At a minimum, the MLA could see it as a benefit to members to provide a link for them to register for ORCID from MLA Commons. With not much more effort, it’s likely that they could initially populate an ORCID record to some extent. If they were to tout that as a benefit of membership, this could result in a significant number of humanities scholars obtaining ORCID iDs. It could also provide a model for similar arrangements with other discipline specific organizations. This could ultimately lead to a beneficial critical mass of ORCID iDs for HSS scholars, especially if, as Mr. Waters suggested, the scholar would need to keep the profile up to date only in the discipline-specific context and the ORCID record would be automatically updated.
- *Create a boilerplate white-label document making the case for ORCID.* While this information is likely already available in many places and forms, including on the ORCID website, it could be beneficial to create a document that a society or publisher could adapt and incorporate in their own systems. Taylor & Francis, one of the CWG members, has an excellent online document entitled “ORCID: how to include it in your online submission (and why you should),” that is a good example of how to not only make a case for ORCID with authors but to make it easy for them to get an ORCID iD.

4.4 Distinguishing Types of Books

In most of the interviews, discussion of using ORCID in book workflows assumed that the books were scholarly monographs. Its value in serials was emphasized, and the issue of contributed volumes came up because of the regrettable lack of metadata about chapter authors in them.

But when prompted by the interviewer, many interviewees immediately grasped the importance of large reference works, including encyclopedias, and of conference proceedings. The workflows for these types of books are more like journal workflows than book workflows. They involve managing a large number of authors and of manuscripts that often arrive in random order. The authors of conference proceedings and reference work entries typically lack the recognition given to journal or monograph authors. And in some disciplines—notably engineering—it is conference papers rather than journal articles that are the most important factor in professional advancement.

RECOMMENDATIONS:

- *Work with societies in fields where conference proceedings are of prime importance.* Many of these—IEEE is an example—have sophisticated infrastructures and very likely already

completely appreciate the importance of ORCID. It would be worthwhile to reach out to such societies to determine if ORCID is being used for conference proceedings. If it is not, it will be a relatively easy argument to remedy the situation.

- *Work with publishers that publish large volumes of conference proceedings or large reference works.* Taylor & Francis, for example, has an operation in Leiden solely devoted to conference proceedings, and they also publish large STM reference works in the US and encyclopedias and other reference works in the UK. Since there are a small number of such publishers, it should be more practical for ORCID to address this segment than something that would require a more extensive campaign.

4.5 The Backlist Problem

Many interviewees were skeptical of the value of ORCID to their book publishing programs because their content is not mainly recently authored. Publishers like Brill, Cambridge, and Springer Nature have extensive book backlists reaching back decades or even centuries. This was also the case for organizations that aggregate large amounts of book content, such as JSTOR or national libraries like the BL and KB. For these organizations, other identifiers or name authority records like ISNI and VIAF are much more likely to be able to apply across the entire corpus they manage. Nevertheless, all of these organizations acknowledged the importance of ORCID and the desirability of obtaining ORCID iDs for books.

This may appear to be a problem requiring a complex and long-term solution, but there are factors that would argue for beginning to address this issue now.

RECOMMENDATION:

- *Convene a meeting—virtual if necessary—with OCLC, ISNI, the BL, the KB, and JSTOR.* Four of those organizations were interviewed, as well as the new Executive Director of the fifth, ISNI; it is this writer's clear impression that they would all be receptive to discussing this issue as a group. ORCID is already working with some of them on projects that could ultimately contribute to a more comprehensive solution. Developing a strategy for more complete interoperability between the systems of the six organizations would provide enormous value to scholarship. It would be a shame to miss an opportunity to get a conversation started. It is not inconceivable that a pilot project could be initiated by JSTOR Labs, and that a funder like Mellon would be willing to support it. We have the attention of all of these parties as a result of this engagement. They all expressed interest in following through in some way.

5.0 Appendix: The Interviewees

5.1 Commercial Publishers

- *Brill*

Marjon Jekel, Applications & Data Manager
Imre Zevenhuizen, Data Processing Manager
Tom Weterings, Production Editor Journals

- *Springer Nature*

Martijn Roelandse, Head of Publishing Innovation
Christina Hoppermann, Bibliographic Metadata Manager, Product & Metadata Management

- *Taylor and Francis*

Ben Denne, Editorial Director of Humanities/Media, Routledge
David Cox, Director of Innovation (books editorial team)
Nicola Parkin, Editor, International & Global Politics, Routledge
Fiona Macdonald, Senior Publisher, Chemical & Life Sciences, CRC Press

5.2 University Presses/Nonprofit Publishers

- *Cambridge University Press—Academic*

Emily Marchant, Author Services Manager
Kate Allen, Academic Product Data Manager

- *Göttingen University Press and Library*

Margo Bargheer, Head of Electronic Publishing

- *MIT University Press*

Amy Brand, Director
Bill Trippe, Director of Technology

- *University of Minnesota Press*

Terence Smyre, Digital Projects Editor

- *Wits University Press*

Andrew Joseph, Digital Publisher

5.3 Technology and Systems Providers

- *Aries Systems*

Tony Alves, Director of Product Management

- *California Digital Library*

Justin Gonder, Product Manager, Access & Publishing Group

- *Inera*

Bruce Rosenblum, CEO

- *Klopotek*

Michael Castner, Head of Product & Solution Management

5.4 Service Providers

- *Altmetric*

Kathy Christian, COO

- *Atypon*

Marty Picco, VP-Product Management

- *Copyright Clearance Center*

Jennifer Goodrich, Director of Product Management

- *Crossref*

Ed Pentz, Executive Director

Chuck Koscher, Director of Technology

- *Ithaka*

Jabin White, VP, Content Management

Alex Humphreys, Director, JSTOR Labs

- *OpenEdition*

Pierre Mounier, Associate Director for International Development

5.5 Standards Organizations

- *EDItEUR*

Graham Bell, Executive Director

Tim Devenport, Publishing Consultant & Project Manager;
also Executive Director of the ISNI International Agency

- *NISO*

Todd Carpenter, Executive Director

5.6 Libraries and Library-Related Organizations

- *British Library*

Neil Wilson, Head of Metadata Services

- *Koninklijke Bibliotheek (KB), the National Library of the Netherlands*

George.Siemensma, Team Lead for Digitization

Arjan Dekker, Information Collection Specialist

- *OCLC*

Lorcan Dempsey, Vice President, Membership and Research, Chief Strategist

Mary Sauer-Games, Vice President, Product Management and Product Marketing

5.7 Funders

- *The Andrew W. Mellon Foundation*

Donald Waters, Senior Program Officer for Scholarly Communications

- *National Institutes of Health (NIH)*

Rick Ikeda, Director—Office of Research Information

- *Wellcome Trust*

Robert Kiley, Head of Digital Services