E-Data Quality: How Publishers and Libraries are Working Together to Improve Data Quality

Carlen Ruschoff
University of Maryland, College Park, ruschoff@umd.edu

Suzanne Kemperman
OCLC, kempers@oclc.org

Elizabeth W. Brown
Project MUSE, EWB@press.jhu.edu

Rena D. Grossman
John Wiley & Sons, Inc., rgrossman@wiley.com

Noah Levin
Springer Nature, noah.levin@springer.com

See next page for additional authors

E-Data Quality: How Publishers and Libraries are Working Together to Improve Data Quality

Abstract

High quality data is essential for discovery and access of e-resources, but in many cases low quality, inaccurate information leads to low usage and a poor return on library investment dollars. In this article, publishers, aggregators, librarians, and knowledge base providers talk about how they are working together to improve access to e-resources.

Authors

Carlen Ruschoff, Suzanne Kemperman, Elizabeth W. Brown, Rena D. Grossman, Noah Levin, Alistair Morrison, Charlie Remy, and Jabin White

E-Data Quality: How Publishers and Libraries are Working Together to Improve Data Quality

Carlen Ruschoff (ruschoff@umd.edu)
Professor Emerita, University of Maryland

Suzanne Saskia Kemperman (kempers@oclc.org)
Director, Business Development and Publisher Relations at OCLC

Elizabeth W. Brown (EWB@press.jhu.edu)
Manager, Publisher Relations at Project MUSE

Rena D. Grossman (rgrossman@wiley.com)
Manager, Data Standards and Quality at John Wiley & Sons, Inc.

Noah Levin (noah.levin@springer.com)
Metadata Manager at Springer Nature

Alistair Morrison (a.morrison@jhu.edu)
Senior Product Manager for ScienceDirect and Digital Initiatives at Elsevier

Charlie Remy (charlie-remy@utc.edu)
Electronic Resources and Serials Librarian at the University of Tennessee at Chattanooga Library

Jabin White (Jabin.White@ithaka.org)
Vice President of Content Management at ITHAKA/JSTOR

Abstract

High quality data is essential for discovery and access of e-resources, but in many cases low quality, inaccurate information leads to low usage and a poor return on library investment dollars. In this article, publishers, aggregators, librarians, and knowledge base providers talk about how they are working together to improve access to e-resources.

Introduction

In late 2012, a small group of librarians began an informal discussion about what it would take to improve the metadata used for discovery of electronic resources and the timeliness of its distribution. We approached OCLC to suggest that a collaboration be formed to investigate the challenges in providing accurate, timely, and reliable access to e-resources. With OCLC's sponsorship, informal discussions were held with groups of publishers, aggregators, and knowledge base vendors to obtain a perspective on the challenges they faced in distributing metadata. We separately interviewed librarians to gain an understanding of the challenges they faced. In 2013, the group became more formalized, calling itself the *E-Data Quality Working Group[1]*. We expanded to include individuals representing libraries, publishers, data suppliers, and service providers, all of whom had a shared interest and responsibility to improve content discovery and access for library users. At the November *2013 Charleston Conference*, our group presented our research on the challenges we had identified and called for more feedback from our respective communities.

Using the responses obtained, in the fall of 2014 the group published *Success Strategies for Electronic Content Discovery and Access: A Cross-Industry White Paper*[2] which provided recommendations to publishers, data suppliers, and knowledge base service providers for improving the quality of metadata and its distribution. The paper generated a great deal of interest and afforded a number of opportunities to present our results to a variety of interest groups. Members of the working group made both in person and virtual presentations for the *National Federation of Advanced Information Services (NFAIS), the American Library Association, OCLC, and the 2014 Charleston Conference.*

In this article, members of the working group discuss the impact of the White Paper recommendations on both the supplier and consumer sides of the information industry; some of these findings were part of a presentation at the April 2016 *Electronic Resources and Libraries Conference*.[3] First, we review the main three problems that the Working Group isolated and the accompanying White Paper recommendations. Then, we discuss how different publishers, vendors, and aggregators have incorporated the recommendations in their operations to improve discovery and access to e-content.

Problems and Recommendations

The *E-Data Quality Working Group* identified three key issues that caused failures and prevented users from accessing digital content. Below is a review of those concerns along with the recommendations that were made.

Key Issue Number 1: Incomplete or inaccurate bibliographic metadata (needed for discovery) and holdings data (needed for access).

Recommendation: Improve bibliographic metadata and holdings data.

- Use e-identifiers instead of print identifiers in bibliographic metadata to describe e-resources.
- Provide consistent collection information to align data with the titles and collection names used in the sales and marketing materials.
- Verify data before sending to ensure that the data provided matches the library's actual holdings.

Key Issue Number 2: Bibliographic metadata and holdings data not distributed simultaneously.

Libraries and service providers have difficulty maintaining knowledge bases when they receive these two types of data for a single item or collection at different times.

Recommendation: Synchronize bibliographic metadata and holdings data.

- Follow a schedule to update data files at the same time as collections.

Key Issue Number 3: The distribution of data in multiple formats.

Library staff must spend time and resources reformatting and, in some cases correcting, erroneous data, which introduces the possibility of additional errors.

Recommendation: Use consistent data formats.

- Use Knowledge Bases and Related Tools (KBART) and Machine-Readable Cataloging (MARC) standards to exchange data throughout the supply chain.

- Provide change management records with scheduled data feeds to alert libraries to alterations in collection subscriptions.

- Provide direct holdings data to the knowledge base service providers so that libraries will no longer have to manage their holdings independently.

Progress with Recommendation 1 – Improve bibliographic metadata and holdings data

Each of the providers on the *E-Data Quality Working Group* has reported progress in improving the quality of data that is distributed and consumed, both through general efforts and in relation to the specific recommendations concerning identifiers, collection naming, and data verification against a library's holdings. OCLC adopted all the recommendations in the White Paper early on and has worked with its publisher partners on the implementation in their data strategies. Over the last two years, OCLC has continued to see the overall quality and consistency of incoming data improve – in discovery data, collection data and direct holdings feeds.

As an example of building and improving general data infrastructure, Wiley has restructured its Product Data Standards & Quality Team, and the company hired experienced librarians to guide best practices for data creation, maintenance, and cleanup. In recent months, senior members of this team have collaborated with developers and content producers to define data standards, business logic, thresholds for data quality, and error reporting.

These efforts have resulted in a well-defined enterprise product data hub which allows the team to review metadata from a central point of governance. New insight into data from disparate internals systems has helped Wiley's data team to confidently identify a single, trustworthy product record that can be made available for circulation across various channels. The team's data analysts are actively profiling, disambiguating, and guiding cleanup of journal and article data. Rena Grossman of Wiley anticipates that the product data hub will help the company perform root cause analysis of data early in the product lifecycle in order to resolve any data discrepancies before concerns are raised by service providers or librarians.

Identifiers

One of the biggest obstacles in the data quality area has traditionally been the use and misuse of identifiers in bibliographic metadata. Many records describing electronic resources hold identifiers for the print editions rather than the electronic. Publishers have acknowledged the problem, and many have set out to address it.

For example, JSTOR has put a routine in place to detect print ISBNs in ebook records and correct them before sending out the bibliographic feed. This seemingly minor step is a significant one in terms of impact on the cleanliness of the BOOKS@JSTOR metadata feed, according to Jabin White of Ithaka-JSTOR.

Springer and Wiley, on the other hand, are using both print and e-identifiers for every title (ISBN for ebooks and ISSN for journals).

Consistent naming of collections

Publisher and service provider naming practices for collections pose major challenges and complicate library workflows. Consistent collection information that aligns with the titles and collection names used in sales and marketing materials is sorely needed. Publishers continue to work to address these challenges, although some gaps remain in the information supply chain.

Project MUSE, Elsevier and JSTOR report that their companies are pulling sales files and KBART files from the same product databases. Consistency between the collection name used by the marketing and the KBART distribution arms are, therefore, reliable at the point of distribution.

Wiley has met with vendor partners to review their 2017 collections to be sure that products in the vendors' knowledge bases match Wiley's marketing materials.

Package names at Springer Nature are cleared with the sales teams to ensure that the collections named in the KBART files match what is being sold. These collection names also match the package names used on their platforms and MARC records.

Noah Levin notes that Springer Nature checks with the different knowledge base providers to ensure they too are naming the packages according to the file naming in the KBART packages. Unfortunately, some knowledge base providers still seem to use their own naming conventions, which has caused confusion with customers.

Alistair Morrison confirms that Elsevier has encountered the same problems regarding package names in the various knowledge bases; some knowledge base providers often change the names of collections or even combine collections into something that Elsevier does not actually sell, such as "ScienceDirect Books 2015." The Elsevier strategy is to append a unique identifier to collection names in their KBART files, however, these identifiers are often omitted by knowledge base service providers. Morrison suggests that the KBART recommendation be modified to allow a collection ID code assigned by the publisher as a solution to this problem.

Provide data that matches library holdings

Noah Levin states that, historically, Springer Nature has placed a priority on quality metadata for its digital collections. Springer Nature verifies its title data at multiple points in the workflow, with many checks on the data being entered into the system and also post-fact. Springer Nature staff also actively performs regular cleanups based on quality control reports, which is often why customers might see ebooks change packages.

Morrison of Elsevier reports that data distributed to service providers and libraries are coming straight from its entitlement systems and

therefore there is confidence that the data are accurate. Problems with data arrive when reports fail part-way through the process. Elsevier is working to improve monitoring of its system so that bad reports can be detected before they are posted on its website.

White reports that JSTOR performs quality control on the metadata in its database prior to distributing data.

Progress with Recommendation 2 – Synchronize bibliographic metadata, KBART and holdings feed

When data suppliers and service providers fail to provide bibliographic metadata, KBART, and holdings data simultaneously, users may follow links to resources no longer available, or they may miss out on important available resources altogether. The most significant recommendation in this area was to follow a schedule to keep bibliographic metadata and KBART files synchronized and up-to-date.

White reports that JSTOR creates a new KBART file every time a new collection is launched, and the ebooks file is updated weekly. JSTOR also does a massive update to its files when it is time to update the "moving wall" data to reflect journal content that has become available in its Archival Journal Collections. Once completed, the files are posted on the website where providers and libraries may pick them up at their discretion. Many libraries and knowledge base providers use scraper programs that grab files from the JSTOR website automatically while some providers do it manually. JSTOR targets getting all updated holdings information into its systems each Friday.

Project MUSE provides web pages where customers can download book MARC records, KBART files, and preliminary title lists for collections. The MARC records and KBART files are up-to-date as of the time that they are down-

loaded. Project MUSE also provides journal collection record sets, with monthly updates covering any new journals that have launched on the platform.

Wiley is working to align with vendor partners to ensure that their KBART reports are reaching libraries as efficiently and accurately as possible. Until recently, communication with vendors and service providers has been limited to monthly email alerts when KBART reports are available on its FTP site. Grossman acknowledges that, like other publishers, Wiley struggles with delays in knowledge base updates for data corrections. To improve and refine workflows, members of the Wiley Standards & Quality Team and Wiley's newly hired Library Technical Services Specialist held "meet and greet" sessions with vendors to develop an improved communication plan for 2017.

Elsevier has implemented automated data sharing with major providers of cataloging, discovery, and article linking services, including OCLC, ProQuest and Ex Libris, and EBSCO. Elsevier provides weekly updates to participating knowledge base service providers. (For more on Elsevier's automated data sharing, see the section below on Direct Holdings).

At the beginning of each month, Springer Nature posts KBART files to a public FTP site and publishes MARC files via a publisher downloader tool. Customers can request these MARC records at any delivery rate (daily, weekly, monthly, etc.). New titles are posted to both the KBART and MARC files as soon as they go live on the Springer platforms.

Levin notes that he encourages libraries to utilize Springer Nature's ONIX service, normally used by trade partners such as Amazon and BN.com, when interested in knowing about upcoming titles before publication. ONIX files are available in ONIX 2.1 and ONIX 3.0 at whatever schedule the customer requests.

Progress with Recommendation 3 – Use consistent data formats

To tackle the problem that libraries receive data in multiple formats from the supply chain, the Working Group recommended the use of consistent data formats (specifically MARC and KBART), change management records, and direct holdings feeds.

Consistent data formats

All of the publishers and aggregators participating in the Working Group provide both MARC records and KBART files. White highlights that the use of and advocacy for improvements in standards such as KBART and MARC are important for the supply chain.

It is worthwhile to note that all changes in data standards, such as moving from KBART Phase I to Phase II, require development work and multiple department efforts for content providers. For example, when Project MUSE was working to implement KBART, members of the KBART Working Group advised waiting for the release of KBART II, which was imminent. Waiting to focus on implementing KBART II was undoubtedly more efficient than implementing KBART only to start a new implementation process.

Change management records

If knowledge bases are not updated when both scheduled and unscheduled changes occur in collections, users will be unable to discover and access content even though it should be available to them. Providers should include details about changes in their regular data feeds.

For example, like many other providers, Project MUSE offers collections of forthcoming ebooks published throughout the year. MUSE makes sure that MARC records and KBART updates are available as soon as new books are released. MUSE also offers title lists that project collection

contents based on metadata received from publishers; the lists are generated from the central database and indicate which titles have become available on MUSE. When an occasional book must be removed from the MUSE site, Project MUSE issues MARC delete records, and the KBART and title lists reflect the removals.

From OCLC's perspective, content providers would establish a workflow for all data to provide library service providers with updated metadata, preferably within a designated timeframe, which would address issues related to pre-publication data not being updated immediately upon final publication. Publishers have taken different approaches with this recommendation.

Challenges with change management include the fact that different systems may track different aspects of the workflow, such as sales history and entitlements. Suzanne Kemperman of OCLC wonders if it would make more sense to put responsibility for tracking the add/update/delete transactions in the hands of the vendors who understand their knowledge base systems better than the publishers can. In the case of Elsevier's ScienceDirect MARC program, this is exactly what has happened: OCLC is comparing each new holding report it downloads from Elsevier against the MARC records OCLC has already shipped to it.

Direct holdings feeds

A significant improvement in the data supply chain is an increased adoption of direct holdings feeds. Traditionally, libraries have had to maintain holdings information manually. For libraries, direct holdings feeds from publishers support automated workflows and keep collections up-to-date with links for each provider and automatic updates when changes occur. Direct holdings make content and collections available more quickly, support all business models, and enable faster access and increased usage.

Kemperman notes that OCLC was the first organization to advocate for and implement direct holdings feeds and has made direct holdings a focus of its activities. OCLC is making a significant investment in the WorldCat knowledge base, focusing on capacity, speed, and quality. This investment supports increases in content coverage to meet the collection needs of libraries into the future, improvements in update frequency to ensure changes are reflected quickly to users, and an even greater focus on data quality to ensure high reliability in linking to electronic content/full text.

As noted, Elsevier has invested significant energy into implementing direct holdings feeds to major providers of cataloging, discovery, and article linking services. Morrison notes that the work with OCLC has been the most intensive work Elsevier has done with automated data sharing thus far. OCLC incorporated Elsevier's automated data sharing into its WorldShare system and began using the data to maintain the WorldCat holdings knowledge base for WorldCat Local customers and for libraries that receive ScienceDirect MARC records through OCLC. Libraries that use Elsevier's holdings service and the WorldCat knowledge base have their holdings symbol posted to WorldCat and receive holding updates that are unique to their institution. Individual customer holdings reports are generated on demand when requested via the API.

The University of Tennessee at Chattanooga is an example of an institution that has taken advantage of the work that OCLC and Elsevier have done with direct holdings feeds. The University uses OCLC's WorldShare Management System as an integrated library system, including WorldCat Discovery and the WorldCat knowledge base for its electronic resources holdings. In the summer of 2014, the University enabled automated holdings feeds for Elsevier ejournals and ebooks. The university's knowledge base holdings now mirror exactly

what they have access to on ScienceDirect, whether it be subscribed ejournals, perpetual access ebooks/ejournal backfiles, or open access content. Charlie Remy of the University of Tennessee Chattanooga reports that Elsevier automatic feeds have saved his library from having to manage holdings manually for over 2,000 ejournals in the *Freedom Collection* subscription, whose contents change throughout the year, as well as hundreds of ebooks. After performing several spot checks, staff at the University of Tennessee at Chattanooga determined that the feeds were accurate and up-to-date, two essential qualities that ensure optimal discoverability and accessibility for patrons. As a result, the usage of Elsevier content has steadily increased in the time that the University has enabled the feeds.

Since entitlement data for every customer cannot be sent to the entire world without the customer's prior approval, Springer Nature is also developing an online portal that customers will need to use to sign off on Springer Nature sending the customer's automated holdings feeds to knowledge base providers. This manual step prolongs the process but is necessary so that sensitive data is not being sent without consent.

Next Steps

Access versus entitlement

The major focus of the E-Quality Working Group has been to improve discovery and access to digital resources so that library patrons can seamlessly reach content. The group recognized early on that the lowest hanging fruit was to work with vendor files that contain metadata for content *access* rather than the sales files which show the titles to which libraries have purchased entitlement. Carlen Ruschoff, University of Maryland, observes that there seems to be a gap between the sale of some titles and the availability of either the online content or the linking metadata. Whichever of the two, the

problem is that the data for some titles is missing from the files currently distributed to service providers. Feedback from libraries, and the vendors themselves, indicate that the next stage of work should be to provide complete entitlement files.

Levin highlights that in theory access and entitlement systems should be the same, although in reality they are rarely in sync. Ideally the entitlement files should show what the customer has purchased, but reflecting this can take several years of system development. On the other hand, pulling from what the customer has access to on Springer Link is readily available to be used for automated entitlements lists. These system abilities for a publisher are often based on how a publisher's systems were set up many years ago, often predating any concept of KBART or discovery data. In this case, it is a question of which system can interact with the bibliographic data needed for the KBART files.

Elsevier has come down strongly on limiting the role of the knowledge base to entitlements. Elsevier's reporting system pulls data from the ScienceDirect entitlement system. Morrison explains that there is a close alignment between the goals and the data architecture of a knowledge base and an entitlement system. They are both designed to support access to resources, and they identify resources at the title level. Removing the complexity of sales history has made it much easier to design the Elsevier system and ensure its accuracy. Consider, for example, the challenge of collections that include forthcoming titles. Sales history shows that they have been purchased, but they are not entitled until they are actually published. Focusing on entitlement gives Elsevier a streamlined way of ensuring that the knowledge base matches what users can access.

More identifiers: customers and collections

As mentioned, Morrison has suggested that the KBART recommendation be modified to allow a collection ID code assigned by the publisher in order to make sure that collections are correctly identified and represented.

Another question providers are grappling with related to identifiers is how one exchanges a customer ID with a knowledge base without industry-standardized IDs. In order to send entitlement lists one needs to identify the customer via a customer ID that every knowledge base will recognize. If a knowledge base service agrees, publishers can send internal proprietary ID's to use with the knowledge base, but this practice opens the question of sustainability once there is large scale adoption amongst many publishers all using their own proprietary customer IDs.

Levin notes that until an industry standard for a customer identifier is decided upon, development by publishers of automated entitlements lists will be slow while publishers are trying to build a consensus on their own for a delivery method and inevitably not meeting the needs of every group that receives that data. As an example, Springer Nature is researching the idea from EBSCO's Oliver Pesch to use SUSHI-lite as a delivery method for the holdings feeds using the KBART delimited text file format. SUSHI might be an ideal candidate since it is used by many publishers, customers and companies and offers a method to have a standardized customer ID. In the meantime, a new offshoot of the NISO KBART Standing Committee will be working on creating this standard which is greatly needed by the publishing community.

New quality checks

Morrison posits that the greatest issue for automated data exchange has been the radical break it represents with past practice. The system does what it was designed to do very well. It pro-

vides a highly accurate title report of each library's ScienceDirect holdings to support discovery of and access to these resources. However, the system displaces procedures that libraries had in place and the quality checks that went with them such as comparing title counts in a MARC record delivery against a particular purchase. In fact, this sort of manual accounting does not work well for large online collections, and could often delay the addition of new titles to the catalog or discovery system. Morrison is calling for a new discussion among libraries, publishers, and vendors to develop quality checks suited to the automated data exchange process that is now emerging.

Continued improvement to data systems

Additional action items and next steps are sure to emerge as publishers and vendors continue to work on improving data quality and delivery as recommended in the White Paper. From OCLC's perspective, the recommendations continue to represent a real opportunity to remove friction from the entire metadata ecosystem, including libraries, content providers, and library service providers. More publishers and aggregators are now exploring adding direct holdings feeds, especially in cooperation with OCLC. Other publishers continue to work to improve their practices and infrastructure. For example, Wiley plans to work with its Library Technical Services Specialist to survey vendor contacts, members of Wiley's library advisory board, and the company's sales support to gain a better understanding of the changes libraries would value most. Grossman anticipates a survey will be distributed in early 2017. Possible work under consideration includes incorporating elements of the KBART II recommendation, improving data exchange workflows, and initiating an interdepartmental committee to begin the conversation about automated holdings.

In addition, publishers who have been working to implement the White Paper recommendations

have occasionally recognized unforeseen benefits. At the time the cross-industry E-Quality Working Group was formed, Elsevier was beginning to design a new system for reporting collection title lists and customer holdings based on the KBART II standard. The development of this system was guided by several of the recommendations in the White Paper, and enabled Elsevier to carry out their ambitious goal to provide direct holdings data to the service provider. Because the reports are all in KBART format, it has been easy to start importing them into other systems. Because the reports are based on Elsevier's entitlement system, they have provided a way to audit collection setup in the entitlement system. The "All Titles" collection report posted on their KBART site has become the authoritative title list for ScienceDirect and is used by several of its own systems to monitor when new titles become available to users. Finally, the customer holdings report has provided a new tool for customer support. Elsevier's own dependence on the same reports provided to libraries and customers creates a virtuous cycle that helps find entitlement or collection setup problems and ensures the accuracy of the reports. According to Morrison, in the two years since implementing this system Elsevier has learned, along with its library and service provider partners, a great deal about the promise and challenges of the vision laid out in the White Paper.

Conclusion

The guidelines published in *Success Strategies for Electronic Content Discovery and Access: A Cross-Industry White Paper* enable all partners in the supply chain to streamline their processes and thereby deliver purchased content to users within weeks, rather than months. While it is always challenging to actually put recommendations into practice, the industry has found these recommendations to be practical, common sense-based steps that moved the practice of metadata distribution in a direction that made sense for publishers and service providers.

More publishers seem to be realizing the importance of discoverability of their content and the need to invest resources to improve it. Developing these services requires a great deal of investment on the vendor side of the supply chain. Admittedly, vendors don't have bottomless pockets of money to invest in infrastructure and staffing on their end and therefore, without a customer push for automated holdings, publishers were not likely to spend resources on the development needed when other priorities are fighting for the same resources. Since the White Paper was published in October of 2014, libraries began to request automated feeds of bibliographic and customized holdings records and publishers fortunately began heeding the call. Some vendors have created full time positions and/or entire departments focused on providing libraries with better quality metadata and collaborating with discovery services/link resolver vendors. Others have already begun to develop new platforms or at least started to rethink their infrastructure. In addition, the White Paper gave a number of simple recommendations that publishers can put into place without infrastructure investments. Small changes like using standard file formats and different identifiers for print and electronic versions of the same product, providing complete, accurate identifiers and metadata, keeping titles and collections consistent, and following a schedule will all improve the workflow in the supply chain and support users in getting access. There seems to be a strong emphasis in finding new ways for publishers, service providers, and libraries to work together to prioritize ongoing development projects that have the greatest impact on their customers.

As the supply chain for bibliographic and holdings data improves, automated processes are likely to replace manual procedures that libraries have in place to check entitlement, data quality, and accuracy. It is clear that manual ac-

counting does not work well for large online collections, and it often delays the addition of new titles to the catalog or discovery system. What is needed is a new discussion among libraries, publishers, and vendors to explore and develop quality checks suited to the automated data exchange process that is now emerging. There is an opportunity for libraries to rethink their own workflows to create new efficiencies. Part of this process may be to explain to auditors and other officials how checks and balances are achieved using technology rather than manual comparisons.

Ultimately we want to make it as seamless as possible for libraries to receive and process publisher metadata, know that they have done so accurately, and provide resource access to their patrons. By working together to address cross-industry problems with data quality, parties involved in the content supply chain can improve the value of their content and their service to library users.

Endnotes

[1] The original Working Group included Bill Brembeck (Senior Product Analyst, Data Services and WorldCat Quality Management, OCLC), Elizabeth W. Brown (Manager, Publisher Relations, Project MUSE), Alexandra de Lange-van Oosten (Head of Third-Party Platform Relations, Elsevier), Theodore Fons (Executive Director, Data Services and WorldCat Quality Management, OCLC), Catherine Giffi (Director, Strategic Market Analysis, John Wiley & Sons), Suzanne Saskia Kemperman (Director, Business Development and Publisher Relations, OCLC), Noah Levin (Metadata Manager, Springer Nature), Alistair Morrison (Senior Product Manager for ScienceDirect and Digital Initiatives, Elsevier), Carlen Ruschoff (Director, Technical Services and Strategic Initiatives, University of Maryland), Gregg A. Silvis (Associate University Librarian, Information Technology, University of Delaware) and Jabin White (Vice President of Content Management, ITHAKA/JSTOR).

[2] Kemperman, Suzanne Saskia, Bill Brembeck, Elizabeth W. Brown, Alexandra de Lange-van Oosten, Theodore Fons, Catherine Giffi, Noah Levin, Alistair Morrison, Carlen Ruschoff, Gregg A. Silvis, and Jabin White. 2014. *Success Strategies for Electronic Content Discovery and Access: A Cross-Industry White Paper*. Dublin, OH: OCLC. http://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf.

[3] Ruschoff, C., Bill Carney, Noah Levin, Kim Maxwell, Alistair Morrison, & Jabin White. (2016 April) *E-Data Quality: How Publishers and Libraries Are Working Together to Improve Data Quality.* Panel presented at the ER&L Conference in Austin, Texas.

Appendix: About the Authors

Editors:

Elizabeth W. Brown is the Manager, Publisher Relations at Project MUSE. (EWB@press.jhu.edu)

Suzanne Saskia Kemperman is the Director, Business Development and Publisher Relations at OCLC. She moderated community input sessions at the 2013 and 2014 Charleston Conferences and presented findings at NFAIS workshops. Ms. Kemperman is a founding member of the *E-Data Quality Working Group*. (kempers@oclc.org)

Carlen Ruschoff is Professor Emerita, University of Maryland. Prior to retirement, she served as Director, Technical Services and Strategic Initiatives at the University. She participated in the panel presented at the 2016 ER&L Conference: *E-Data Quality: How Publishers and Libraries Are Working Together to Improve Data Quality*. She provided updates on the "Success Strategies" White Paper in a number of other venues including the Charleston Conferences, National Federation of Abstracting and Information Services (NFAIS) workshops and ALA conferences. Ms. Ruschoff is a founding member of the *E-Data Quality Working Group*. (ruschoff@umd.edu)

The Contributors:

Rena D. Grossman is the Manager, Data Standards and Quality at John Wiley & Sons, Inc. Ms. Grossman joined the group in 2016. (rgrossman@wiley.com)

Noah Levin is the Metadata Manager at Springer Nature. He participated in the panel presented at the 2016 ER&L Conference: *E-Data Quality: How Publishers and Libraries Are Working Together to Improve Data Quality.* Mr. Levin is currently a member of the NISO KBART Standing Committee. He also served a member of the NISO KBART Working Group that published *Knowledge Bases and Related Tools (KBART) Recommended Practice*. (noah.levin@springer.com)

Alistair Morrison was the Senior Product Manager for ScienceDirect and Digital Initiatives at Elsevier. He is now Manager of Library Applications at Johns Hopkins University. Mr. Morrison participated in the panel presented at the 2016 ER&L Conference: *E-Data Quality: How Publishers and Libraries Are Working Together to Improve Data Quality.* He serves on the NISO Vocabularies Working Group and is a member of the NISO Book Metadata Working Group. (a.morrison@jhu.edu)

Charlie Remy is Electronic Resources and Serials Librarian at the University of Tennessee at Chattanooga Library. Mr. Remy presented at the 2016 ER&L Conference: *Achieving the Holy Grail of Electronic Resource Management: Automated Holdings Feeds in a Knowledge Base.* (charlie-remy@utc.edu)

Jabin White is a Vice President of Content Management at ITHAKA/JSTOR. Mr. White participated in the panel presented at the 2016 ER&L Conference: *E-Data Quality: How Publishers and Libraries Are Working Together to Improve Data Quality and* presented *Achieving the Holy Grail of Electronic Resource Management: Automated Holdings Feeds in a Knowledge Base.* He also serves on the NISO Board. (Jabin.White@ithaka.org)