KENNEALLY: Publishers and authors who want to sell their books in the digital age need to learn a new language. Metadata. It sounds like a computer language, and in a way it is, but metadata is much more. It’s as old as the ancient library of Alexandria, Egypt, and as new as the World Wide Web.

Welcome, everyone, to Copyright Clearance Center’s podcast series. My name is Christopher Kenneally, host of CCC’s Beyond the Book. Even though metadata predates the world of e-books and tablets, it’s essential for publishing in 2013, because only with accurate and complete metadata can digital products be located and purchased online. Professional publishers large and small need to master the development and distribution of comprehensive metadata for the books they publish.

The new book The Metadata Handbook is a one-stop guide for success. And joining me from Columbus, Ohio is co-author Renée Register. Renée, welcome to Beyond the Book.

REGISTER: Hi, Chris. Thanks for having me on.

KENNEALLY: Well, we’re delighted to have you join us, and we’ll tell people that in addition to being co-author with Thad McIlroy of The Metadata Handbook, Renée is founder of DataCurate, a company supporting publishers and libraries in the development of data policies, practices, and systems designed to connect readers to content. Renée has over 20 years’ experience in building and growing innovative metadata systems, products, and services, including 10 years with Ingram Book Group and six years with OCLC, a global library cooperative.

On Tuesday, February 12th at the Tools of Change for Publishing conference in New York City, Renée Register will present Creating Powerful Metadata, a 90-minute workshop where participants will employ an online tool for metadata entry and transformation to Onix. And you know, Renée, we’re not going to get into it that deeply here on our little program, which is probably a good thing for me, because I’m not sure I could handle Onix as well as others would. But we should help our audience understand the critical point here, which is just how important metadata is to the future success of e-book publishing for authors and publishers.
So let’s just review a few basic facts, then. What is metadata intended for? Perhaps think of it this way – define metadata’s goal.

REGISTER: Well, in the context of the book business, and indeed, most e-commerce, whether it’s shoes or refrigerators or books, I like to talk about metadata, and the context is information needed for two things. First is the information needed for product description, what the potential reader or consumer might need to know, either to find or discover a book, and, importantly, to decide if that’s the book that they want.

And what we often don’t talk about is the information needed for effective commerce. And this is what publishers and their trading partners need to know and track and analyze for business reasons, and a lot of this metadata the consumer doesn’t even see. But it’s driving behind-the-scenes transactions and business analysis and strategizing, and all of that important stuff.

KENNEALLY: Well, as we know, in the old days, if you will, it was a brick-and-mortar distribution system in place for the selling of books, and now we’ve got kind of a combination of clicks and mortar, I suppose, with the clicks becoming more and more important. So all of that metadata is going to be crucial if the book business is going to continue to thrive.

REGISTER: That’s right. Because I mean, in years before, when we just had bricks and mortar, metadata was really important, and a lot of the tools that the retailers used to order books from publishers or wholesalers was out there, and there were interfaces that they could see. But really, the consumer didn’t see it, and a lot of that stuff was exposed when all metadata went digital.

So people began to understand, it can’t just be brief information for inventory purposes and transaction. It has to be really rich metadata that works for a lot of different publishing functions.

KENNEALLY: Right. And I’ve had a chance to look at your new book, The Metadata Handbook, and I understand that metadata grows and evolves over the life of a book. And in a print version, I suppose when you published it, the metadata was there, was fixed forever in that copy printed at that time. But metadata can grow in the digital space. Explain that for us.

REGISTER: That’s really true, and we all know that the information about a book is often released into the marketplace long before the book actually exists. In pre-Internet selling days, maybe that was released to potential resalers or wholesalers. But now the consumer can see it, and in the online world we can begin to take preorders way before a book is available. And it can even become a bestseller before it even kind of exists to the consumer.
So it’s really important that publishers get that information out early. But that also means that that information has to continue to grow and evolve during the pre-publication process. At some point you have a cover, at some point you have a table of contents – and then some metadata doesn’t even happen, important metadata to my mind, anyway, until after publication, or maybe even long after publication. Things like awards – a book receives a National Book Award, it may be a year later. And then, of course, reviews continue to come in, and those are very helpful.

So it’s important not to just send it and then it’s done, but to consider it an evolving description.

KENNEALLY: Right. And more than that, if I understand you correctly, integrating the creation of metadata throughout publisher workflow is a kind of holy grail for success.

REGISTER: It is, it is. And that’s where we’re trying to go, and that involves some rethinking of publisher processes, and thinking about who has the best knowledge of different aspects. That may mean that metadata doesn’t just go at the end to marketing or copywriters or technical people, but that through the editorial process – I mean, who knows a book and an author better than an editor? They may be working to get a good biography, a good description.

What is the subject? You don’t really want to leave that subject applications all toward the end, maybe an intern just slapped some subjects on there. You kind of want to integrate it into the process, and ideally – this is another holy grail – you have a central database where you’re storing your metadata, and the same metadata could be used for production tracking and all those kinds of functions, and be good enough to still then be exported out to your trading partners.

KENNEALLY: Well, with all those different players contributing to the creation of metadata, it sounds to me like there’s the potential for confusion, for misinformation, and all sorts of things. Where should the ownership of metadata lie within a publisher?

REGISTER: I don’t really have a good answer to that question right now, because it’s still evolving. I think now it’s more maybe in the marketing realm, in terms of the content of the metadata, and it’s more of the technical realm in terms of how the metadata is carried. And later, we can get into this a little bit more, but I think really, ultimately it may be more of a marketing function, because it’s how you get your information out to the world for buying and selling. But everybody has a part.
KENNEALLY: Indeed, I think that probably is the point, to say everybody has a part in this creation of metadata, and I think your identifying marketing as perhaps the primary source is a really interesting one, because of course, what we care about in the metadata is how to get the book discovered and eventually purchased online.

Well, we are talking today on Beyond the Book with Renée Register, who is co-author of *The Metadata Handbook* and founder of a company called DataCurate, which supports publishers and libraries in development of data policies. And she’ll be presenting at Tools of Change for Publishing, coming up on February 12th this week, a special 90-minute workshop, Creating Powerful Metadata.

And joining us on the line is Patricia Payton, senior manager of publisher relations and content development for Bowker, the world’s leading provider of bibliographic information. Pat Payton, welcome to Beyond the Book.

PAYTON: Thanks, Chris. Happy to be here.

KENNEALLY: Well, we’re glad you could join us, and you’ve been hearing what Renée has had to say about the importance of metadata. And of course, Bowker is in that business, deep into that business. And I wonder if you can help us get a sense of the current state of communications, and sort of interpersonal relationships, if you will, between publishers and distributors, as far as metadata goes.

PAYTON: Well, publishers create their metadata, but distributors also do services in the way of metadata. So they might add on to the metadata, they might distribute the metadata for the publisher. So it really becomes a partnership, because you don’t want your distributor sending something that you don’t expect to happen, and you don’t want your customers not to find your book because your distributor didn’t send it to somewhere that you wanted it to go to.

So it really is a communication of the sales and marketing area of the publisher working with this client services area of the distributor to make sure that all the I’s are dotted and everything goes to all of the customers as needed.

KENNEALLY: Right. And Pat, you were a contributor to *The Metadata Handbook* for Renée Register. And how critical is it to have such a book as this out at this particular point, do you think?

PAYTON: I think that it is very critical. At every conference that you go to, you have at least a 45-minute session on metadata. Well, Renée has compiled a lot of research and a lot of background about the industry, how it’s evolved over time, how metadata has evolved over time, and how it’s evolving into the future and what to look for and what to expect in the way of new identifiers, new customers that might come around that need metadata.
So I think that this is an important handbook for every publisher to have and to read through, so that they can see what is their next action plan. And since metadata is evolving, it’s, what am I going to do in the next six months? What am I going to do in the next 12 months? But the energy spent on metadata will continue to be spent over the next several years, because we are moving and changing as we go.

KENNEALLY: Well, Pat Payton, that sounds like a considerable burden to publishers. And I want to turn to Renée Register and ask, is your sense that publishers are investing sufficient resources in this project?

REGISTER: Well, we need to do more studies, more research on that and continue to reach out to publishers as with the book industry study group surveys and things like that. But from the final section of the book that you referred to, which was on metadata and the future of publishing, we did ask questions such as that to industry leaders about the most important problems, outcomes, and solutions.

And one thing that came up consistently was that we need systems that have the capacity and flexibility to handle this dynamic nature of metadata today, and the volume. So to me, that sounds like a degree of investment if it’s not already there, especially for larger publishers. They also brought up integration into the overall publisher process, which means looking at your organization and staffing responsibilities and all of that.

So I think that we may have a way to go, and the future kind of points to maybe at some point more whole-industry solutions instead of piecemeal solutions, but to really look at it together as an industry, and figure out how we can get there without all the duplication of effort.

KENNEALLY: Well, hopefully it’s still early enough in their process that we can do that together. Can you tell us, Renée Register, what exactly is open data? It’s an expression I saw in your book. And why is it important?

REGISTER: Well, there is a trend toward open data that we have to accept, and that is that data is kind of out there and reused in multiple ways by multiple players, rather than being in proprietary systems. Now, that’s a real challenge for us, because when metadata goes out from the publishers to these various trading partners, it does end up in proprietary systems and then goes out from there.

So to really embrace this would involve a good look at our business models and how the current players such as Bowker, who have such a massive store of metadata already and do lots of things to manipulate it, would play in this world.
The other part of that that I can’t get into that much and I’m still trying to learn is the idea of linked data, which you can link different pieces of data to other pieces, kind of recombine and remix for your needs, and that can only happen if it’s sitting out there for folks to play with, so to speak.

KENNEALLY: Right. Well, Pat Payton there at Bowker, can you give us your top 10 or maybe even top six data fields, as far as importance goes?

PAYTON: Sure. It depends on whether you’re looking at discoverability or transaction, like Renée mentioned earlier. But for discoverability, title, author, description of the book, table of contents, the description of the author, their author biography or their affiliations, where they work, who their constituents are.

And then you have things like awards, media mentions, that maybe are put on later in the supply chain, as Renée mentioned, but all of those are important for discoverability, because people are looking for a recommendation from a friend. Well, in an online world, it may not be an exact recommendation from a friend, but a friend may become an endorsement from a review source or from an award.

Then for transaction analysis, or for transacting on the book, you’d want things like the price, where you can buy it, how many copies are available for purchase at this time, when am I going to receive it if I order it today. So I think that those combined make up the metadata record that a publisher needs to have today.

KENNEALLY: Right. And Renée Register, finally, you’re going to be doing some training, some on-the-job training, if you can call it that, at Tools of Change in your workshop, Creating Powerful Metadata. What kind of training do you think that publishers should be considering for their employees, perhaps in that marketing department that you were mentioning?

REGISTER: Well, that’s exactly what I’m trying to work on and evolve now. There really isn’t anything standard out there now. The organizations such as Book Industry Study Group here, BookNet Canada, and BIC in the UK do lots of great Webinars and seminars and provide lots of documentation.

But what I’m working on, and this workshop is kind of the first part of the experiment, is developing training materials that can be used by different departments, by different publishers, kind of to standardize that so that there’s a methodology behind it. And I hope that that will continue also to evolve and get better as folks give me feedback and input on how this works.

But I think metadata is a practice, and you can read all you want to, but until you start playing with it and actually inputting it and analyzing your content, you won’t quite get there. So those are the tools I’m trying to build.
KENNEALLY: Well, I understand it’s going to be a rather well-attended program at Tools of Change. I’m not sure how many people are going to be attending, but from what I saw online, it looks as if it’s going to be a full room. So good luck with that, Renée Register, presenting Creating Powerful Metadata, a 90-minute workshop at the upcoming Tools of Change for Publishing conference held in New York City on February 12th.

Renée Register is also the co-author with Thad McIlroy of The Metadata Handbook and founder of DataCurate, a company supporting publishers and libraries in development of data policies. Renée, thanks for joining us today.

REGISTER: Thank you so much, Chris.

KENNEALLY: And Pat Payton, senior manager of publisher relations and content development for Bowker, thank you very much for joining us on Beyond the Book.

PAYTON: Thanks, Chris.

KENNEALLY: Beyond the Book is produced by Copyright Clearance Center, a global rights broker for the world’s most sought-after materials, including millions of books and e-books, journals, newspapers, magazines and blogs, as well as images, movies and television shows. You can follow Beyond the Book on Twitter, find us on Facebook, and subscribe to the free podcast series on iTunes, or at the Copyright Clearance Center Website, copyright.com, just click on Beyond the Book.

Our engineer is Jeremy Brieske of Burst Marketing. My name is Christopher Kenneally. For all of us at Copyright Clearance Center, thanks for listening to Beyond the Book.

END OF FILE