Book Sales Secret Sauce – Metadata  
Interviews with  
Rebecca Albani, Publisher Relations Manager, Bowker,  
Clark Fife, Metadata Manager, Macmillan

For podcast release  
Monday, January 14, 2013

KENNEALLY: Once locked away exclusively in the card catalogs at libraries, metadata in 2013 has escaped to the Web, where in digital form, it describes the contents and context of data files. And because almost every form of published content today is created in digital form, metadata is an essential building block of every publishing business.

Welcome, everyone, to Copyright Clearance Center's podcast series. My name is Christopher Kenneally, your host for Beyond the Book.

In a world where brick and mortar bookstores are vanishing like white rhinos or blue whales, the challenge and burden for metadata about books is to drive first discovery, then sales. For guidance on best practices in metadata, we turn to the world's leading provider of bibliographic information and management solutions for publishers, booksellers, and libraries. And joining us today from Bowker is Rebecca Albani, Publishing Relations Management there. Rebecca, welcome to Beyond the Book.

ALBANI: Hi, Chris. Thanks for having me.

KENNEALLY: Well, it's good to have you join us today. Rebecca Albani assists publishers with making their data more discoverable to customers, and therefore, helping increase sales. She educates publishers on the types of metadata that should be supplied with titles, as well as the most effective method of submitting this data in order to improve searchability. And let's remind people too, that in addition to its extensive role in bibliographic information management, Bowker is the official ISBN agency for the United States and its territories.

And Rebecca, what we'd like to ask you to help our audience with is understanding some of the principles around best practices for metadata. And as I mentioned at the top, you know, it is very much a technology business today, publishing, and metadata is an essential piece of that.

Now, we know sort of the basic definition for metadata – it's data about data. But tell us how that all applies to the book business.
ALBANI: Well, so we break metadata down into a few different categories. We have our basic metadata, which is the core fields that are needed by pretty much anybody you're submitting your data to, which is your title, ISBN, author, price, and a few other fields.

And then we have what we called enhanced metadata, which is the more detailed information about the titles, the author biography, table of contents, cover images. These, the enhanced metadata, is not necessarily needed with the titles, but it's obviously going to help make your titles more discoverable, if you're sending in that type of data with your titles.

Along with all that, submitting all the data together makes it more discoverable, and then helps you increase your sales.

KENNEALLY: Right. And, you know, it's probably no revelation that metadata is important, but it is becoming increasingly a complex issue. And as you mentioned, there are some basic fields, but then there are others that really describe in depth what content is available for sale here, and offered.

There's some problems that publishers encounter in more frequency than others. What are some of the stumbling blocks to providing high quality metadata?

ALBANI: I would say some of the stumbling blocks are, like you said, the standardization of how you're submitting data. So if you're going to submit data to someone, you should take a look at the BISG best practices, because there are certain ways that data is accepted, and you know, places like Bowker will take data any way that you send it, but it's easier once it's standardized. So by sending – you know, if you're sending it in an Excel file, or even in an ONIX file, sending the title and the subtitle in separate fields, sending the contributors as last name, first names – if you're doing all that, it's just going to make your data easier for people to find.

KENNEALLY: And of course, the thing you mentioned, BISG, for people who may not be familiar, of course, that's the Book Industry Study Group, and they have released a rather extensive list of best practices for data centers. And data, in the digital realm, of course, is just 0s and 1s, so it can describe almost anything. And so, beyond titles and authors, we're talking about cover images, for example, and a uniformity around cover images is important for their potential discovery and sale.

ALBANI: Right. Sending cover images is very important, because when you're searching online, if you're looking up a title and there's no cover image, you might not necessarily buy that book, because you're not – like, sometimes a cover makes you buy that book. But when you're submitting your cover image to providers, you want to make sure that you know exactly the pixels and JPGs, whatever format they need, because different providers accept different things. So one company that you're sending your cover image to might
take a JPG, where another person might want a TIF. So, when you're sending your data out, you want to make sure you know what people are accepting before you send it out.

KENNEALLY: And of course, what Bowker provides is assistance and counsel in all the aspects of data submission.

When you work with smaller publishers, and indeed, the growth in independent publishing has been tremendous just in the last year, are you finding them not so well prepared as perhaps the larger houses are?

ALBANI: Yes, I would say that, but I would also say that they are getting more knowledgeable in the area, because there's a lot more people leaning towards small publishers or self-publishers. So they are getting more knowledgeable. But I've spoken at quite a few conferences and giving presentations to a lot of small publishers and self-publishers, and they're always very encouraging and very welcome to hear any suggestions that Bowker or anybody has on submitting the data and making sure it's clean.

KENNEALLY: Right. And at a presentation that I was a part of that you gave back in the fall at the annual PubWest conference, you gave us an idea of the best practices for submitting to Bowker itself, and there is a website for all of this, myidentifiers.com. Can you tell us a bit about what's involved there?

ALBANI: So, myidentifiers.com is a website for US-based customers, publishers from – really, from large to small to self-publishers, independents, can go to myidentifiers. They can purchase their ISBNs directly through this website, and then ideally, for the small and self-publishers, they can log in once they've purchased their ISBNs and upload their metadata directly through this website. It is hand-keying your titles into the system, so we don't recommend it for our larger publishers. But for the self-publishers, it's fairly simple. They go in, and they can upload the information. All of the basic metadata fields that are mandatory have a star next to it.

They click submit, once they've entered all the information. They can upload the cover image, and then the data gets automatically loaded into our database within 48 hours, and then after that, it gets sent out into our regular customer feeds.

KENNEALLY: Right. And I think what's important about that is, again, the independent and small publishing piece of the business grows, the whole challenge of metadata is to get the material, to get the information correct. It's not simply to do it. We want to get people over the hump of being put off by the enormity of the task. And so the kinds of things that Bowker is doing is helping to make that easier for them, and to make it a faster process.

ALBANI: Correct, yes.
KENNEALLY: Absolutely. And of course, the point here is, finally, to help drive sales. And so, make that connection for us. Again, it's a philosophical principle that metadata is helping to describe the content in its best possible way, but how does that then make it easier for the book itself to be found? Because the consumer, presumably, is doing searching, and is trying to come with a title that would be relevant to them, and they'll be asking for all kinds of things, and you need to anticipate their need. Isn't that just about right?

ALBANI: Right. So, for us, when you submit data, whether it's through myidentifiers or sent through a file to us, like a large publisher would, it gets loaded into our database, and then it gets sent out into our customer feeds. Our customers are – consist of retailers, online retailers, libraries, schools.

When the data gets sent out, it's basically just – at least, as I tell people, it's free marketing for the publishers, because all of that information that you've submitted to us now about an ISBN and a book is now being sent out to everybody that gets a feed from us. It's going to libraries. So when you're going to a Barnes & Nobles, and you're going to libraries, you're actually pulling back the Bowker data. So by submitting clean data, it's going to make your titles more discoverable, and the more information you can supply about your titles makes it more discoverable, because when you submit an author biography, a table of contents, all of that information is searchable.

So, when people search for keywords, it will actually pull from the author biography or pull from the table of contents, where if you only submit a title and an author, you could be missing a lot of sales, based on keyword searches.

KENNEALLY: Right. And an aspect of this that's perhaps overlooked on occasion is not simply the retail end of it, but the social media aspect of it, with so many sites emerging where readers and consumers promote books to their friends and associates. Discovery is an important part there, and metadata plays a role there as well.

ALBANI: Right. So there's tons of websites. Goodreads.com and librarything.com, where you can – everyone can go on there, and they can review any title they want. It's no longer about reviewing the New York Times bestseller titles that are getting reviewed. Like, I could upload – I can make sure my title is on goodreads.com, and tell everyone I know to go review the title. All of that stuff is searchable, then, in a search on the Internet. So all the social media sites are – come up within searches when you're doing them.

So I guess that the more information you can have on your title, the more places you can have your title listed, the more discoverable your title is going to be.

KENNEALLY: Right. It's all about sort of creating those hooks for the content, and making sure that they're in the right places. And indeed, this is a subject that one can go into in
great depth. You've provided us with a very brief overview. I know that your colleague, Pat Peyton at Bowker, is going to be on a panel next week at Digital Book World on Thursday, called *Making Content Searchable, Findable and Shareable – Tools and Best Practices to Improve Discoverability*.

So while we are providing people with a very high level introduction to all of this, they can get more of the details, and sort of get further into the specifics with this panel that will include Pat Peyton as well as Matt MacInnis from Inkling, Linda Holliday from Citia, and moderated by Neal Goff of Egremont Associates, along with Gary Price at INFOdocket. So we hope that they will join your colleague Pat Peyton there for that.

And with that, by way of background, I want to turn now to someone who has to put all of this into practice, and understands what works and what works best. Joining me on the line from his office at Macmillan is Clark Fife, Metadata Manager there. Clark, welcome to *Beyond the Book*.

FIFE: Hi, Chris. Thanks for having me.

KENNEALLY: Well, it's a pleasure to have you join us, and we'll tell people that Macmillan, of course, publishes trade books and textbooks as well as scientific journals, that prior to becoming Metadata Manager there at Macmillan, you worked as the Product Manager, Product Category Manager at Barnes & Noble and barnesandnoble.com, helping to design their subject hierarchy. And you were previously Director of Metadata Services at Howcast, an online video site specializing in how-to content.

So, you're a metadata manager, and I suppose you get asked a question from time to time as to what that's all about. Why is this so critical at this moment to Macmillan, someone like yourself being in this position and making sure that the data is correct and in the right places?

FIFE: Well, as more and more sales move into the online realm, and the online realm becomes increasingly larger with a huge explosion of titles, it's just, every advantage that you can get to make your book stand out in a very, very big sea of books is – I mean, it's crucial to take advantage of each and every opportunity, and make sure that your book does come up in the searches every time.

KENNEALLY: Right. And Rebecca Albani from Bowker was just giving us a very high level description of the challenges involved here. Can you tell us about any examples where you were able to really sort of move beyond the ordinary level of metadata to sort of – you know, help a book, a title, any kind of content, emerge in that sea of content that's now out there? What other kinds of things do you learn from experience work best?
FIFE: Well, I think it's really about knowing who your audience is, and trying to think about the different ways in which they may be searching for a book, and making sure that each one of those keywords gets included in every title. We've had cases in which we've had great success with a title by just taking a look at the data and saying, oh, it's a book about dogs, but it doesn't mention puppies, and we want to make sure that the book comes up in a search for puppies as well as dogs, and just tweaking the metadata that way, making sure that a book description and an author biography includes all those possible ways in which a customer might potentially search for the book.

KENNEALLY: Right. And you have an interesting perspective, having worked on the retail side of things there at Barnes & Noble. Can you help us understand perhaps what I would imagine is the frustration that many online retailers have with regard to metadata? Are they not seeing the kinds of data they are looking for?

FIFE: Well, I know that in my time at Barnes & Noble, we saw a lot of publishers submitting some data that was of low quality, shall we say. So, I'm aware of the problems that retailers get, but I'm also aware that by submitting bad data, that you run the risk that someone downstream, either a retailer or a distributor, data aggregator, is going to fill in, and make changes to your data that you don't know about.

And that's something as a publisher – working for a publisher, we always want to avoid that and we want to control our data as much as possible. So the best way of going about that is making sure that you don't give anyone downstream any excuse or any reason to change your data, just provide them with a clean and accurate data from the get-go.

KENNEALLY: Right, and Clark Fife, Metadata Manager there at Macmillan, you've been at this for quite a few years, so you've really seen the evolution of metadata as it applies to book publishing.

Where do you see this all going in the next year to two years? Are there any developments ahead that we should be watching for, or is it just really a matter of making sure that the standards are applied, and that people really do think through the kind of tagging that they need to do to be as thorough as possible?

FIFE: I think a lot of that is going to be increasingly more important, and I think that the publishers that don't get a handle on just doing that sort of basic work are going to fall behind in the race. In terms of what's coming up next, I think it's going to be a lot more in terms of multimedia content, where we're doing a lot with video and audio clips that we distribute to different places. We're doing a lot with different kinds of author websites, and getting the sort of social media aspect in there.
And I think you'll just see a lot more of that. Taking advantage of the fact that it's online, you can provide them with extras in content that you couldn't necessarily get in a physical bookstore.

KENNEALLY: Well, that's a great point, and Rebecca Albani, I'll bring you back on and ask you about how Bowker, which, you know, at least at the level of the ISBN, is associated with books, is handling that particular challenge, which is that content today describes not only text, as you might see on a book page, but video and all sorts of other types of data.

How do you manage all of that, and are you seeing an evolution from the book level down to the chapter, or even, so-called, chunk level?

ALBANI: We are seeing that. We do see currently where we'll get people are assigning individual ISBNs to individual chapters, because they're selling them independently. We also do see the – you know, the audio clips, and video clips, and we currently take all the data for that.

So, we're seeing it in our end too, and it's – for us, it's a matter of making changes to our database and to our feeds going out to our customers, to be able to store all of the new information coming out, and being able to send it out to all of our customers, and sending it out correctly.

So, you know, it's one of those things whereas everything is evolving, we have to change our databases to make sure that we're taking everything. So it's keeping up with the times.

KENNEALLY: I have no idea how you do it. In fact, I kind of want to close this discussion off with you both by asking about that relationship that a publisher has with Bowker. I'll turn first to Clark Fife there at Macmillan.

Give us an idea of how frequently you're working with Bowker, and how you are working with people like Rebecca, her colleague Pat Peyton, and others, to develop solutions that are going to anticipate the consumer needs and the needs of retailers as well.

FIFE: Well, I would say that we want to make sure that our data flows to them very cleanly. We pay attention to the recommendations at Bowker provides, as well as the recommendations that come out of the BISG, to make sure that, yeah, we're complying with the current standards. And then, of course, we're always looking for what's the next thing, and what's the next thing. Yeah, a lot of that does comes directly from Bowker.

KENNEALLY: And Rebecca, is that really what it comes down to, that S word there, standards? Is that really the critical piece? Without standards, it's going to be just a confusing mess, I suppose.
ALBANI: I would say yes, because without the standards, people can send their data any which way, and we try to – well, we’ll take data. Like Clark said, like we will go in and make changes to the data in order to have things show correctly to our customers as well. So, following the standards is a very important part of the – getting your metadata clean.

KENNEALLY: Well, we understand here at Copyright Clearance Center about standards, particularly where it comes to rights, and we want to follow this particular issue throughout this year, so we look forward to having you both back at a future point to look at some of the news that we expect to see in metadata throughout 2013.

I want to thank Rebecca Albani, Publisher Relations Manager at Bowker, for joining us today. Rebecca, it’s a pleasure having you on the show.

ALBANI: Thank you for having me.

KENNEALLY: And Clark Fife is Metadata Manager at Macmillan. Clark, thank you for joining us as well.

FIFE: Thank you for having me, Chris. This was interesting.

KENNEALLY: Beyond the Book the produced by Copyright Clearance Center, a global rights broker for the world's most sought-after materials, including millions of books and ebooks, 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