



**Minimum Viable Metadata  
Recorded at the 2019 London Book Fair**

**For podcast release  
Monday, March 25, 2019**

KENNEALLY: Welcome, everyone, to Olympia Hall and the London Book Fair. My name is Chris Kenneally for Copyright Clearance Center.

Our panel today is going to discuss Minimum Viable Metadata.

David Weinberger is a coauthor of *The Cluetrain Manifesto*, a highly influential treatise, now 20 years old, on the ways that the Internet has changed markets and marketing. Weinberger holds a doctorate in philosophy, has taught at university, and was a comic strip gag writer. What he has to say will often sound profound and preposterous all at once.

The cure to information overload is more information. Transparency is the new objectivity. The smartest person in the room is the room.

Concerned as he is with the essential elements of the information age and what he has called the new digital disorder, David Weinberger has also given considerable reflection to how we organize our digital world, which brings him to observe the following about metadata.

“The next Darwin is more likely to be a data wonk than a naturalist. To a collector of curios, the dust is metadata. And finally and crucially, metadata liberates us, liberates knowledge. Publishers, take note. Metadata liberates us, liberates knowledge.”

For publishers undertaking initiatives to upgrade or update content workflow and knowledge management, metadata becomes the great gremlin. As Weinberger has noted, our choices will reflect not only the world but also our interests, our passions, our needs, our dreams.

Such an encyclopedic approach seems overwhelming when resources and time are limited. A shortcut to success can begin with defining Minimum Viable Metadata – the set of bare minimum information necessary to describe each element of content. The MVM itself will reflect a mix of internal and external factors, from IT systems to compliance requirements.



Our panelists this afternoon are looking forward to sharing their advice on how to establish an MVM model quickly and sufficiently for the challenges at hand, while not allowing the perfect to be the enemy of the good. And I want to introduce them to you now.

Starting on the far end, to my right, is Brian O’Leary. Brian, welcome.

O’LEARY: Thank you.

KENNEALLY: Brian is Executive Director of the Book Industry Study Group, a US-based trade association that works to create a more informed, effective and efficient book industry supply channel. Brian is the author of research reports on the use of metadata, territorial rights in the digital age and best practices in digital exports.

And then immediately to my right is Marie Bilde Rasmussen. Marie, welcome. Marie comes to us today from Copenhagen in Denmark. She is an independent consultant in the publishing industry, where she works with publishers, retailers and aggregators on metadata and supply chain issues. She represents the Danish ONIX and TEMA user groups in editors’ steering committees.

And then to my left is Ian Synge. Ian, welcome.

SYNGE: Thank you.

KENNEALLY: Ian is a Principal Consultant at Ixxus, a subsidiary of Copyright Clearance Center, with particular specialization in knowledge management, taxonomies and characterizations. He has a longstanding enthusiasm for knowledge organization, and his PhD focused on the taxonomic interpretation of naval diplomacy during the last Cold War. So we may get to some of that, but that will be for later on in the discussion.

And finally, Joshua Tallent is at the far end there. Joshua, welcome. Joshua is Director of Sales and Education at Firebrand Technologies. He’s a metadata expert and a well-regarded teacher and guide on digital publishing. He serves on multiple industry committees and is a regular guest at many publishing conferences around the country and – around the world, I should say, as we are here in the UK.

Ian Synge, I’d like to start with you. And so the phrase Minimum Viable Metadata is the key to our discussion here. And I guess we have to ask, what do we mean by viable? What are we trying to achieve?

SYNGE: Absolutely. And I think let’s get the boat, naval diplomacy reference out of the way straight away.



KENNEALLY: It'd be great. Let's do that. (laughter)

SYNGE: So when we started to think and we started to talk about what do we do in this session – and I thought, what is minimum viable when we talk about this. And topically, let's bring British politics into things – and not about Brexit for a change. It's the 50<sup>th</sup> anniversary of Britain's continuous at sea deterrent. And if you can get through the headlines, you might find some references to it in the press, etc.

And Britain has always adopted the approach that they want to have a minimum credible deterrent. And they said, well, what does that mean? How much do we need to have?

And then someone came up with a blinding insight that said, actually, what Britain thinks with this is entirely irrelevant. It's the recipient of the deterrent that is the person who judges what's minimally viable.

So don't ask the Royal Navy, don't ask the House of Commons – ask the Russians, ask the North Koreans, what does it mean for them?

Now, what's all that got to do with publishers and metadata? Well –

KENNEALLY: I think I have an idea. It's about the customers and what they need.

SYNGE: There you go. There you go. Because ultimately, we as publishers need to think about our users, and what is it that metadata allows them to do? And that comes down to what sort of publisher are we?

If you're a trade publisher producing fiction, maybe all you need is author, title and something which enables Amazon to go and surface your content pretty readily. If you're an engineering publisher, maybe you need to go into a lot more detail, so you can wrap all of your content to make it smart content that could be integrated into systems to make it much, much more powerful.

KENNEALLY: Well, so for the listeners here today and for the audience on our podcast series, tell us about the challenge. Because, as I was describing with the references to David Weinberger, metadata seems a really overwhelming subject. It's something that can be as – well, we are speaking about minimum, but it can go as large as you want it to go. Is it true that publishers and others who have media content have more metadata than they think they do?

SYNGE: Absolutely. It's almost invariably the case. So when I sit down with a publisher who's trying to do something, and it's usually I can't find my stuff. I



can't control my stuff. I don't know where my stuff is. Help me. I think a great analogy –

KENNEALLY: And the stuff, we should say, is their various content assets. That can be images, text, entire works and so forth.

SYNGE: Yeah, exactly. And a great example – last year, was working with a publisher who – their main content was pretty well controlled.

But a lot of the ancillary material that was around it – so, in particular, their video content and their imagery was out of control. And they described a situation like they had a warehouse where the lights were turned off, so they had brilliant stuff in there, but they couldn't find it. So a lot of the time they just recreate it.

And we sat down with our media manager and said, OK, what do you have? And he goes, we're hopeless. We have nothing.

And that's always really scary. You present someone with a blank sheet of paper, and they're always going to – you know, it's terrifying.

So we said, well, actually, we think we've probably got more than you think, so show us some of your files. And he'd bring up a file and say, look, there's no metadata. It's just a picture.

And I said, well OK, right-click on it and say view properties. What do you see? Well, you see a data created, so that tells you something. There's a geocode on it because you took it with a camera that's put a latitude and longitude onto it, so you can tell where it was taken, and that tells you something. With video, you can tell how long it is, and that tells you something. And from those, you can start to infer things. And that started to get him towards this principle of minimally viable.

Now, it's not earth-changing. It's not something that unlocks all of the value there, but it gets you started, and it meant that they could start actually start actually surfacing this content rather than just endlessly reinventing the wheel.

KENNEALLY: Right. And as I was hearing you tell the story, I was thinking we should make sure that everyone in the room understands metadata. Metadata is essentially data about data, and so you were describing that data of the video file, and the subtext that you were able to uncover was that data about the video itself.

And that's of value to many of the publishers, particularly in scholarly publishing today, because more and more, they're incorporating video and there may be, as



you say, a catalog of work that they think they can't use, but it may be more viable than they realize.

SYNGE: Absolutely. Absolutely.

KENNEALLY: Well, I want to turn to Marie Bilde Rasmussen today. And you have your own consultancy, and you call it Pruneau, which is a really sweet name because it means plum. And I guess it's sweet but good for you, so is that a reference to data? Is that kind of a metaphor for data for you?

RASMUSSEN: It is. I think the sweetness is just to charm people – but that it's actually good for you, even though you're not very willing to get on working with the metadata.

KENNEALLY: Yeah, but the work that you're doing, it sounds like it's really hard work, as good as it is and as sweet as it might be. You are now working on a fascinating project of a topographical dictionary of Denmark.

And this is something that is not a trade book. It's not really typical scholarly publishing. Tell us about the approach you're taking to making decisions about the kind of metadata that will be applied there, because people are going to be searching for any manner of subjects.

RASMUSSEN: Yes. You could say that, as a printed work, it's just one reference book. It hasn't got a lot of metadata in the trade supply chain. But when we want to publish it digitally, we chop it up in all the little articles in there. And to be discoverable, these need other kind of semantic metadata, where the most important of all is its location.

We want to supply every little chunk of content with its geodata and a type of place – is it a lake, is it a church, is it some archeological site – which will allow users to either read about all the churches in Denmark, find out, if they are on their mobile phone, are there anything in my interest near where I'm standing right now to go see. Something that's – where can I go to the next location of this type? I found this Roman church extremely interesting. How far away is the next one, and what can I read about it?

KENNEALLY: But when you're describing a country, which is essentially what you're trying to do, this gets to this question of Minimum Viable Metadata. As you say, there are certain elements. It's a lake, it's a church, it's a town square. There's history involved. It must be that you have to go to the minimum, you have to make decisions about where you start, because otherwise it's overwhelming.



RASMUSSEN: It's true. It's true. These are editorial decisions, of course. What are the interests covered by the encyclopedia? Will they describe – there's biology, geology, archeology, history, architecture, art – like main categories – so each piece of content will have a subject classification as well as location finding.

And as we know also from other both encyclopedia publishers but also, for example, Danish Broadcasting Corporation, who streams all their content, what people will look for are places, persons and events. And that's what we're going for marking up in our content too.

KENNEALLY: Well, that's the point that Ian was making – think about the receiver of the data, not necessarily the professor whose entire life is focused on a certain town square in a small village. They have to think about why that would be important, why it's relevant to the reader.

Tell us, in your career, why you've come to understand that metadata isn't about how, but about why.

RASMUSSEN: Well, I think that maybe touches back to why you need to understand that this is good for you. It's hard work. It's by many people.

I've been working in large Danish publishers for a long range of years, and people who know the content see the metadata work as an annoyance – something they just have to enter into some system to publish the book. And if they give the job to somebody else, they will do it, because it's their job, but they do not know the content as well.

So you have to make it clear that this actually benefits you. It's not just the key to be able to push the publish button. It will actually either drive your sales or it will reduce superfluous work or duplicated processes.

KENNEALLY: So when publishers are thinking about this challenge, who should they send in to the room to discuss it? What are the kinds of titles that you see?

RASMUSSEN: Yeah, you have to talk to more management than people on the floor. You always follow to the money. Tell the money story to get the activities going.

KENNEALLY: But you also need people like yourself, those who have the metadata expertise. You can't take a guess at it. You have to some kind of direction.

And the metadata in the case of the dictionary – the topographical dictionary – it is going to not only drive discovery for people who are doing research on their personal history or whatever, but it's about sales too.



RASMUSSEN: It is. Yes.

KENNEALLY: And so how will that help the sales chain?

RASMUSSEN: Well, it makes data products or text or content discoverable for potential customers, and it allows them – if you have a lot of content, which is very well tagged or described by metadata, to aggregate the exact or the perfect amount and the right content for you.

KENNEALLY: OK. Well, Brian O’Leary at BISG, you have – you, as an individual, Brian O’Leary, author, but also Book Industry Study Group – written about this topic for a number of years.

And a question I’m going to ask you and some others here’s is why are we still talking about all of this? This is something that has come up in the past. BISG did some reports as far back as 2005 in trying to identify some critical elements. In fact, you have identified at least 31 critical elements.

Tell us about the long journey here to get publishers to understand – and maybe, as well, how that journey has gone in unexpected directions since you began working on this.

O’LEARY: Sure. Well, you’re right about the 31 elements. That’s actually the thing I’ve been holding here. And they do date back about 20 years. They were updated when ONIX was moved from 1 to 2 and 2.1. And they’re still really valid.

But I think the thing – to answer your question directly – why has it taken so long, metadata is an investment, and the investment has to see a return.

The thing I found interesting about your conversation with Marie a moment ago is that the premise of a print product is one set of decisions about metadata. You’re describing essentially a move across platforms, so that you’re making the data – the information – more useful to somebody who’s at a location or seeking something that’s similar in an area. That’s an entirely different level.

And going back a minute to Ian, this notion of stripping out metadata to create a print product has been a consistent theme for publishing really for the entire time we’ve had book publishing.

KENNEALLY: Define that for people who may not realize what you’re saying. When you say it’s stripped out – so someone may push out a really well-crafted file that can then end up being useless.





O'LEARY: Well, what we wound up doing with print is creating a very rich file with a lot of context around it – often the editorial history. You could put in geolocation data, etc. But when you render it to print, you flatten it. It's just two-dimensional. There's not the ability to query it.

We now have tools and platforms – mobile devices are a good example, the Internet's another – where it's possible to collect a lot more information and still keep it valid and usable for any reader or consumer. But we're actually still kind of working with print workflows.

I think the thing that minimum viable to me means is not just these 31 elements, although they're really critical, but what's the minimum viable for the purpose intended? And that's what's changed.

KENNEALLY: And there are clearly multiple purposes. You could be putting together metadata that's critical to the marketing, but also you need a set that would be important for the product itself.

O'LEARY: Sure. Well, I think virtually every piece of metadata that you put into a product is critical for marketing, ultimately. Metadata is marketing, but it's also a component of consumption of the product – the engagement – and that's the piece, I think, that we've missed.

In the last 20 years, not just with mobile devices but really with the growth of the Internet, we've essentially created an entirely different way to consume and engage with content, but we've still got a metaphor that fundamentally is two-dimensional in producing a book.

KENNEALLY: Well, Joshua Tallent, I want to turn to you, because it's the same question I just asked Brian I want to pose to you, which is why do we still have to be talking about this? And you've been talking about it for some time. Maybe you've asked yourself that question too.

TALLENT: I think Brian's right on track. And something that I think we're missing as well is that this is going to continue to be a topic in the future because new technologies are coming,.

So it's not just about, you know, why are we still talking about it – we've been talking about it – but publishers still haven't really completely grasped this in some ways, and there are still pieces of metadata that we're not collecting well, we're not engaging well.





But the next technological leap that happens for publishing is going to require more metadata, different metadata, the types of things that we haven't been collecting or the things that we need to be thinking about collecting now.

And those publishers who have engaged that and are starting to actually think through how do I engage it from the beginning – like Marie was talking about – from the very beginning of the process. Those are the ones who are going to be able to take that technological leap and get into that new market or engage that new consumer base and make those new sales, because they're going to be prepared for it. And the other ones are going to be catching up, just like we've been doing for the last 20 years.

KENNEALLY: Right. Well, that was what David Weinberger said. Right? The answer to information overload is more information. If you think it's tough now, you're going to need more in the future.

Talk about data strategy and how necessary that really is from the very beginning.

TALLENT: Well, that's the thing. I think Marie made the valid point there. We have to be thinking about this from acquisitions all the way forward. And publishers that engage it from that level and start thinking about, you know, from the editorial perspective, from the marketing perspective, from the authors' perspective, who is the metadata intended for, who is the book's end market, the end consumer of that content – and be thinking about that from the beginning.

If you can have that process built in as just part of the process of publishing, it makes it a whole lot easier when you get to the end of the line and you're like, now we're going to go and put this book up for sale. What do we say about it? How do we engage that market? It's going to be a lot easier to do that if you have that process in place from the beginning.

KENNEALLY: Right. And yet it is not a set-it-forget-it approach to things. Some metadata is evergreen of a sort, but I believe you suggest to publishers that they go back and do some spring cleaning.

TALLENT: I do. Yeah, I usually recommend that publishers look at what's going on in the marketplace and the books that they have and take what time they can to engage that again and come back to that evergreen nature of the data.

Every six months, go back and look at your older titles and see what might be necessary to upgrade and refresh that data. If something happens in the political world, or something happens that's related in some way to the topic of the book,



it's good to use that as a benefit to you and your metadata. Update the data with that new information.

There are lots of stories – anecdotes of publishers who have engaged this in some way or another. You know, an author wrote a book. Twenty years later, he comes out with a new book that's actually a bestseller, or he or she engages – becomes a writer on a TV show or does some other kind of thing like that. All of a sudden, that person becomes more popular, and their older works actually have more value to consumers.

And so if a publisher can take advantage of that and use that new information to go back to the metadata and refresh that metadata for that author, it'll make a difference in the total sales of all the author's works.

Those are just examples, but the idea is really just to think about how your data might be interpreted and used by consumers and how you can surface that information more easily and surface that book more easily in search functions.

KENNEALLY: Right. And so the question of what's a minimum isn't a static one either, because what was minimum yesterday could need to be revised tomorrow.

TALLENT: Right. And keywords are a key example of that. If you look at how keywords were not really very useful in the ONIX metadata 10 years ago because nobody was accepting them – there was no “market,” for keywords, because none of the retailers or aggregators were actually using that data.

But now, keywords are actually really important, and we've seen, especially with Amazon, where they're using keywords to surface content to consumers. If you can engage those new areas as they start to come out, if you're ready for them when they start to happen, then you'll have the upper hand in meeting that consumer demand.

KENNEALLY: Right. And something else that's come to the front of everyone's attention recently is accessibility. And that means really looking at existing backlists and making sure that it's accessible.

TALLENT: Yeah, and trying to go back and fix older files that are not accessible is a lot harder than just fixing them from the beginning. Right? Creating them in the right way from the beginning.

Those publishers who went through a massive update of – you know, a massive creation of all of their e-book files, and didn't think through the considerations of structure and the things that actually are the baseline elements of accessibility –



those are the publishers who are going to have the hardest time going back now to 2,000 files or more and having to update them and try to fix them, whereas, if they had taken a little more time at the beginning or even known about it at the beginning, it would have actually been easier for them.

KENNEALLY: Right. And Ian, the audience here may be thinking about their current list of new book titles or even journal titles, but one of the things about metadata is not just that it creates these larger works, but it also makes possible going granular. And so how does the minimum viable metadata approach help go in that direction?

SYNGE: Well, I think it starts people thinking about what's possible. The most terrifying thing is a blank sheet of paper, as we sort of said. If you get them to the Minimum Viable Metadata so you can start to do things, you can start to then think what's next.

And I guess it's interesting – the challenge I often have is persuading people to take that first leap. And it's a double-edged sword. If you're at the leading edge, the floor is usually yours. But the publishers that took that leap, created those e-books didn't think about accessibility nonetheless were pathfinders. And you can't get to the stage of doing the powerful design it properly without getting it wrong first.

And I think, as people start to go down the route, you can start to craft your thoughts in different ways.

So let's take an example. When Erskine Childers wrote *Riddle of the Sands* in the early 20<sup>th</sup> century, the concept of metadata was probably just putting his name on the front of it. But as he wrote it, he was curiously specific about dates and locations. You read it, and it's really precise about what's happening at every stage.

Now, in a 21<sup>st</sup>-century way, wouldn't it be brilliant to wrap all that with locational metadata, time metadata, so you can then go link that with a whole host of other content, make it a totally immersive experience for somebody, make it accessible in lots of ways.

But you can't expect – you can't blame Erskine Childers and his publishers for not doing that in 1905 or whenever it was.

KENNEALLY: Sure. And ultimately, it's about authority. Who are the authorities for the metadata? It's a question I may ask the others here with you – but start with you, Ian – who really is the one who gets to decide all of this? I mean, authors may think they know the best, but –



SYNGE: Well, you know me, Chris. It's impossible to go and have a conversation without me bringing French philosophers into it – and the whole notion of the death of the author.

But I think, increasingly, it goes back to put the power into the hands of the users or the readers. You know, everybody reads things in a different way. They bring their own perspectives to things.

And increasingly, I think, especially when people go back to works of literature, they'll read this in hypertext. They'll go and follow a common thread. They'll read in a particular way. They'll be interested in a particular perspective.

Now, you can't expect the reader to create all of the metadata to support that. That's just unrealistic. But in terms of the authority and who's got the power to go and do that, I think it goes way beyond the traditional, I am the author, I know my subject, I know best, and I will just tell you.

KENNEALLY: Even I'm the publisher as well. It's beyond that. You really do have to go outside of your silo, so to speak. And maybe, Marie, I could ask you about that. For you, in that project – that wonderful project of the topographical dictionary of Denmark – who are the authorities? It seems to me there'd be a lot in that meeting.

RASMUSSEN: Sorry, could you – ?

KENNEALLY: Who are the authorities? Who gets to decide, ultimately, that the project is where it needs to be?

RASMUSSEN: Well, that's the editorial team, in cooperation with the development folks, and then also with the other similar projects in other countries, so we know that we are aligning with them. And we have also a user group and a scientific council to help us prioritize too.

KENNEALLY: And can you address too the way that metadata will help publishers to do the kinds of things that we were talking with Ian about – to go granular and then to aggregate content in new and different ways?

RASMUSSEN: Well, I think having in mind what – I think you should always go talk to your vendors – where's the data going – because they will know their users. And remember that the metadata has to be presented in new and interesting ways – maybe putting something like schema.org, instead of just it ties the content to make it discoverable and also to let it be combined with something that's not in, for example, the book trade sphere, you could cross the borders.



KENNEALLY: And Brian O’Leary, the kinds of readers that we’re talking about here, I guess we sort of default to thinking of them as book readers, consumers. But we’re here at the scholarly publishing end of Olympia Hall and the London Book Fair. How are readers in the scientific world different? And what kind of considerations does that make for this topic?

O’LEARY: Well, I think that the focus shifts from monographs or long-form text to data sets and the ability to mine large data sets for the kinds of either patterns or information that would influence whatever the project that they’re involved with – I mean generically.

I think the thing that’s missing, though, is a lot of the structural elements Joshua talked about, with respect to accessibility – not specific to accessibility, but structural elements that make the data minable.

I think that the project that Marie’s been talking about, there’s a lot of decisions that were made for how to tag a variety of different dimensions. That doesn’t exist across the superset of scientific or medical or technical information.

And in a sense, that challenges the – not so much challenge the notion of Minimum Viable Metadata, but I would kind of phrase it more to say, particularly for large data sets, what’s the maximum viable metadata? What can you afford to put the time and energy into, with the understanding that, sooner or later, you’ll have to update it? It’s like painting a bridge.

KENNEALLY: Well, we’re talking about the decisions that get made here. And as we all know in our own lives, every decision we make isn’t the right one.

What about that, Brian? What about what I believe your report has called the problematic use of metadata? This is not a perfect world, and so we do find issues with metadata that can cause real business challenges. Talk about that.

O’LEARY: Sure. Well, this problem existed a decade ago. It exists today. We have nonstandard applications of standards, so there’s a definition and then there are the ways that providers, whether they’re publishers or intermediaries, interpret those standards and what recipients do with it. That’s a real roadblock to effective deployment of a larger metadata scheme.

This is – I talked about this, actually, at a conference for Firebrand last fall – this is an industry that is particularly resistant to standards or at least to the standard applications of standards, and that’s a real problem for vendors, who have to customize heavily, as well as for intermediaries, who struggle just with their own workflows.



One of the examples we had was page count. And when we did the study in 2012, recipients were changing it, on average, about three-quarters of the time, which meant that they were not – publishers were not following the standard three-quarters of the time.

Page count is not going to kill you. But if you can't follow that one thing consistently well, imagine what you're going to do with more complex elements.

KENNEALLY: And Joshua Tallent, then, I guess you've seen a fair amount of that, that nonstandard use of standards. How do you make the point that standards are really critical?

TALLENT: Well, I think publishers run headlong into this all the time, and they hit that point where they're like, oh, my data's not being accepted the way I wanted it to be or the way I thought it would be, or that e-book got returned and can't be sold because it's not following a standard the way it needs to be.

Those things – usually, you just run headlong into them, and you have to engage them in some way or another. Unfortunately, that takes more time and takes you away from – it's more of a reactionary approach than a proactive approach.

My hope would be that publishers can engage in a proactive approach to standards. They can look at what the standards currently are. And standards change. But in general, the foundational stuff is going to stay the same.

So if we can hit standards where they are now, at least that minimum viable standard. Right? Hit the minimum viable level we can, it'll make it easier down the road, and fewer of those reactionary events will happen.

KENNEALLY: Right. And in the UK here, a counterpart of sorts to the BISG is the Book Industry Communication Group. And I understand – you were telling me that they have an ongoing sort of developed – what's the word – kind of like a breaking-news metadata map. It's always being updated.

TALLENT: Yeah. Actually, they're supposed to be doing a presentation tomorrow about that here at the London Book Fair – that basically, they're doing a study, a multiyear study, of metadata and where it gets sent out into the trade, how those files get manipulated and changed, how the data gets changed and what's happening to it.

And I think that's the kind of thing that we, as publishers, need to be looking at. The industry needs to be looking at what's happening to the data, how are we





engaging it – and because, so often, we’ll – you know, a publisher will put something out, and then it gets overwritten by somebody else or it gets changed by somebody else, and they don’t even know about it.

This was something that came up in the BISG report is that 35% – I think, Brian – was 35% of publishers, 36% know – actually go and look at their metadata and see what’s going on out in the trade, but 95% of them know it is being overwritten.

So there’s a very large percentage of publishers who just don’t have time to engage that issue, or don’t have the resources to engage that issue and go and look and see what’s happening. But they know it’s happening, and they know it’s a problem. And it’s one of those things you just have to – we have to figure out, as an industry, how we’re going to fix it and keep it from being a problem.

KENNEALLY: So Brian, can I follow up on that? The need for publishers to be watching their data, not just creating it, making these decisions, but ensuring that it’s in the right place and in the right context. Talk about that.

O’LEARY: Yeah. First, I would do a shout-out and follow up on Joshua’s theme. If you’re interested in this topic, Peter Mathews will be talking tomorrow at Building a Better Business seminar that BIC is putting on from – I think it’s 10:00 until 12:30. And it’s just on the other side of the hall.

So the question is essentially, how do we figure out how things are going wrong and work on them? It was an interesting thing that, when we did the study seven years ago, we were finding something like only half or a little bit better than half of the updates that publishers were putting into the supply chain – metadata updates – were actually getting through to retail sites.

And it varied widely by retailer, so there were a couple of retailers in the US who were particularly good at it.

KENNEALLY: It’s like flipping a coin. Half the time, it’s not getting there.

O’LEARY: Exactly. And that’s a structural issue in the United States as well, because we’re feed-based. There’s not a database. There are other issues.

But some of those problems, I think we’ll see, are also true in the UK, and tomorrow’s report or update from Peter will touch upon that.

I think it’s the ultimate supply chain issue. You have to think about creating good metadata, but you also have to have systems in place to make sure that the updates actually flow through to the people who would benefit in terms of discovery or use.





And I think, in various markets, the sophistication of those systems is probably not up to the level that we need them to be these days.

KENNEALLY: Yeah. And Marie, can you tell us about the Danish market? How well are they doing compared to the experience that Brian was just describing?

RASMUSSEN: Yeah. As you know, we're a very small market with a – very small also because of our language – we are five million Danes – so it's crucial for us that everything is actually interchangeable, because our vendors are very small. High complexity from different channels will actually kill them.

In Denmark, the print book distribution is legacy. We are trying to change that. But with the e-book distribution that has been around for about 10 years, we have started an ONIX feed, so that we have a very smooth supply chain, so that's where we learn from and can take our experience to (inaudible).

KENNEALLY: Well, Ian Synge, we'll close out the conversation and give an opportunity for the audience to ask their own questions, but maybe I can leave it here – that what we're talking about here in Minimum Viable Metadata proposition is we're trying to take that content, which we might describe as dumb content – those audio or video files that we just don't know how to use – and make it smart.

SYNGE: Yeah, absolutely. I think that's the journey we need to start going down – to get content, to liberate it, to make it powerful. There's the whole notion that intelligent content is something that really allows it to go and be harnessed in lots and lots of different ways.

If you're a technical B2B producing standards or specifications, maybe you can have something that somebody doesn't have to read and re-key that information into a system. Make that data flow into that system automatically.

And I think, as people refer to Boeing aircraft, it's no longer just an aircraft, it's an information management system, dealing with all sorts of contents in lots of ways. And that gives us a huge opportunity to open up worlds way beyond just the reader opening a book.

KENNEALLY: Well, you have opened up my world regarding the topic of metadata. I want to thank our panel today.

Joshua Tallent, Director of Sales and Education at Firebrand Technologies, Ian Synge, Principal Consultant at Ixxus, a subsidiary of Copyright Clearance Center, Marie Bilde Rasmussen, an independent consultant with her own firm called



Pruneau from Copenhagen, Denmark and finally, Brian O’Leary, Executive Director of the Book Industry Study Group. And thank you all for joining us.

(applause)

END OF RECORDING