

The Content Liberation Movement Recorded at BookExpo 2018

with

- Renee Swank, Ixxus
- Ganessan Paramanathan, Alfresco
- Max Riggsbee, Gadget Software

For podcast release Monday, June 18, 2018

KENNEALLY: A very warm global welcome to all of you joining us. My name is Chris Kenneally with Copyright Clearance Center, and we're here to discuss the sort of provocatively named content liberation movement. We'll hear about content is being liberated for all kinds of new purposes. And even well into the digital age, publishers have persisted – sometimes insisted – in maintaining processes that confine their businesses to a specific format, usually the book, and to a single business model.

Forward-thinking editors, however, expect, even demand, the freedom to reuse and repurpose content in innovative, high-value ways, especially on mobile devices. Content management systems, though, aren't usually fast enough at identifying assets and don't go far enough when assembling new products. Our discussion today is going to identify the digital transformation accelerators that can help you editors and executives break down these barriers.

To help me with that discussion, I want to introduce the panel, starting from my immediate left, at Renee Swank. Renee, welcome.

SWANK: Hi.

KENNEALLY: Renee Swank is vice president and head of the Discovery Practice at Ixxus, a subsidiary of Copyright Clearance Center. She has 25-plus years' experience in publishing content and knowledge management, helping clients support new digital-first and content-enrichment processes.

To her left is Ganessan Paramanathan. Ganessan, welcome.



PARAMANATHAN: Welcome.

KENNEALLY: Ganessan serves as evangelist and solutions architect at Alfresco, an enterprise open-source software company focused on driving the convergence of enterprise content management and business process management to advance the flow of digital business. Ganessan has served in a variety of roles, including product management and sales. At Alfresco, he leads a set of applications for digital information and asset management platforms in publishing and other industries.

And finally, on the far end, Maxwell Riggsbee. Max, welcome.

RIGGSBEE: Welcome.

KENNEALLY: Maxwell Riggsbee is co-founder and chief product officer at Gadget Software, based in Newark, New Jersey. Gadget Software is the pioneer of vPub, a virtual publishing technology that atomizes, enhances and streams book, manual, and journal content to smartphones. His recent previous positions span senior director at Cisco, chief marketing officer at Whiptail, and general manager and CTO at Brocade.

Ganessan Paramanathan, I'm going to start with you, because you're from the furthest outside of what we traditionally think of as publishing – although, of course, in the digital age, nearly everyone's a publisher, individuals as well as all types of organizations. Where Alfresco comes in is helping publishers of this content work with their data, getting close to the content via identifying the data. And I guess the question on many people's minds is does this lead us to a point where data becomes the decider of everything with content?

PARAMANATHAN: Sure. I hope you all can hear us.

KENNEALLY: Please lean in, yeah.

PARAMANATHAN: OK. I hope you can all hear me, yeah. If I look at like three to four decades back, data is all about structured information. The last one or two decades, what the trend we see in the industry is data is not only about structured information. Data is structured and unstructured content. So the new definition of data is content plus structure information. That particular information is really critical to make some informed decision – to get some insight about the



temperature of your customer or your consumer and get the feedback loop from your consumer back to your product development. So that's really, really critical.

At Alfresco, we manage that whole complete lifecycle about that data. Again, when I say that data involves your – all the structured and unstructured content, and we manage the complete lifecycle about that data, starting from creation of the content and then manage the lifecycle about the content – you know, manage the reviews and approvals of the book or the article or what you publish.

And then the most important – back to this question – is publish to your consumer and get that input or insight about the consumption of the particular – the piece of data and get the feedback from that piece of information back to your product development or marketing and then enrich it and go from there. So the data has become very critical to get the informed decision and get the gut feel of your consumers and get the insight into the information and then do the continuous loop back to that.

- KENNEALLY: Right. Ganessan, your biography identified you as an evangelist, so I have to ask you about the number of converts you've had to this. In publishing, how easy is that conversion process to get people to come on board to the kinds of close work with data that you're describing?
- PARAMANATHAN: It is challenging. The technology part I would call is the easy one. There is – some structure, methodology is available to that. But the culture is a key important player in this, starting from the creation of the content and then go through that, digitize it, and go through the review and approval process, and also get the insight back. So the challenge here is the culture and the mindset of that – putting your digital process on top of your publishing – it's a cultural thing. That's where I see a challenge. You have the technology in place, but the culture or the mindset, having the digital layer to start from the ideation to all the way to the publishing and to get the insight to it.

KENNEALLY: Yeah. Renee, it looks like you want to comment on that.

SWANK: Yeah, absolutely. I think you made a good point, because I think a lot of it is around reengineering that process to think about your content in a much different way. Maxwell, I'm sure, can speak to this from a front-end perspective and how users start to interact with the content differently on mobile devices or other. But having your content structured differently and enriched with metadata that can



drive some of those new product and user experiences is really, really key. So I think I see content structure, content enrichment as an enabler to giving new features and new ways of interacting with the content on mobile devices and other kinds of digital platform.

- KENNEALLY: Yeah. Max Riggsbee, in fact and Ganessan alluded to this there's a cultural challenge internally in publishing organizations or any kind of content organization. There are also cultural challenges on the customer side of things. You're thinking about all of that at Gadget Software, and you conducted a survey recently you commissioned a survey, I should say by Michigan State University to understand better how readers are navigating content. Can you tell us what you found?
- RIGGSBEE: So I think before I answer that question, I think both of these points are very important. This is particularly true when we're dealing with reference information. When we think of them in books, generally we think of these things as somewhat static, but they're actually part of a dynamic collection of information. Even more importantly, people don't read that information linearly, right? It's a nonlinear journey from the very beginning. But what's now beginning to happen is we want to connect information – say it might be in a manual – with information associated with a sensor to tell me I probably shouldn't open that cap if the radiator is at a certain temperature. Right? But I want that information to come to me as part of that sort of – that work task, if you will.

In terms of the Michigan State Study, what we were most curious about and what we've been spending time on for the past year is we are all sort of captive, if you will, to a particular output format, which is really PDF, right? And it's great. What you see is what you print. But that's not going to play very well on a mobile device. This is a touch-driven device. The semantics are quite different. The user experience expectation was different.

So we commissioned a study. We wanted to know just something very, very basic, which I think we all do here all the time. I want to simply find and save a piece of information. OK? So we had 30 people. Each person had six minutes. And they had to find and save nine pieces of information. Pretty basic, right? Out of those 30 people, 15 were only able to find and save four pieces in six minutes. They didn't have to read it, just find it – find the section, save the section, and then we had a review component to it.



The challenge was that each document was structured differently. The way information came to them was organized differently. And what they were accustomed to was a completely different type of interaction. That's what they wanted to experience. That's what we do at Gadget. We take that same information, and we structure it more like data – data elements, pieces of information that you can pull together. This is now becoming part of a workforce story – not just a consumer story, but it's an efficiency story. If you have nine tasks to do in six minutes and you can only get through four of them, that's a bit of a problem, right?

KENNEALLY: Right. Obviously, publishing encompasses not only the trade publishing that's important to BookExpo, but there are here educational publishers, scholarly publishers, and so forth. So information processing – it makes it sound very sort of like we're living in a laboratory, but this is key to our work world and to our consumer world.

I want to get back, Maxwell, to the question about data deciding things, because Ganessan sort of helped us understand how important it is, but this is also an industry that has relied for many years on its gut and is proud of its gut because its gut says this book is going to be a hit, and it may resist having data inform decisions about not only what to publish but how to publish it. Talk about that tension.

RIGGSBEE: That's actually becoming fairly interesting, right? We have this really rich new medium through which we can consume information.

KENNEALLY: The mobile device?

RIGGSBEE: The mobile device, whether it's tablet or whether it's phone, right? It's touch-driven. It's driven by a completely different set of expectations. And what's beginning to happen is folks are beginning to ask questions around the data that they want to use, but also at times where they want that data to be accessed.

I'll give you an example. One of the largest requests that we're getting is can you take a publication and make it only available in a certain location? So maybe I have someone that's working in a nuclear facility, but I don't want certain book-related information available if they're outside of the geography of that place. That's data, actually. I mean, I have to now have some information that informs the publication that you're in a place where you actually can't render this as



something that's visible, alongside other pieces of information that might be critical to some decision – you used that term a little while ago – some decision I now need to make at this particular point in time under these scenarios. So what we're seeing is an intersection of information from a variety of sources. You hear the term IoT. That's informing some of that.

KENNEALLY: The Internet of Things.

- RIGGSBEE: Right, the Internet of Things. But what's really occurring is there is key information that I need to understand something to do or how to do something, and then there's information around that activity that must now inform that task. And what individuals and companies and others are beginning to ask is how do I get all of that together and deal with the publishing change, which is when I discover an error a modification or an adjustment in some of the core information, how do I not just simply have a new release? How do I just update the important piece of information in that publication?
- KENNEALLY: Right. Renee Swank, your work is helping educate the publishing community that you work with to think in new ways about content, to sort of think beyond the book, if you will, and to think about content in a granular way, in a derivative way. We've been talking about this for a long time. These are not new topics. So before I ask you to tell us how you help them see that, can you put your finger on a point on the scale where one is people have been living under a rock and 10 is they're so tech savvy they're ready for a NASDAQ IPO where is publishing on that scale?
- SWANK: Yeah, so I work with a lot of publishers, and they have big capabilities that they're putting on their digital platforms in terms of creating new ways of exploring their content and being able to search and things like that. But it's really burdened by the old ways of continuing to publish in a very print-based workflow still creating print first, focused on print-first review processes and content creation process, thinking about their content very linearly, and creating front-to-back books.

I think that gets in the way of being able to open and liberate your content to that kind of capability that Maxwell was talking about, where you contextualize your content, where you can allow people to search in different ways. As your study had shown, people can't find the information that they're looking for when it's buried down in a very deep part of the book. It's just not really made for the mobile



devices that we're starting to see people use, especially in research and that kind of content.

So I think one of the things that is important to do is start to break the way that people start to create their content, so they're thinking about how their content can be explored and used in different ways on the platforms that they're developing today and any new platforms that are going to be coming in the future. So it's really about looking at that process – editorial process – and looking at, instead of creating a front-to-back book – front-to-back set of content that's really about a particular product – it's more creating more granular pieces of content that can be mixed and matched and enriched that devices, whether that might be a web platform or a mobile device, can explore that content in new ways.

Perhaps an end user may want to mix and match that content to create their own books. Maybe it's contextualized based on where they're at. So I think that it's about starting upstream in an editorial process in order to create content agility and giving those end platforms a way to create new user experiences.

RIGGSBEE: There's one thing I want to add to that, which is – I'm certainly using the term tablet and smartphone, but eyewear is on the rise. So we're beginning to see the intersection of three-dimensional images in one part of the space, and then we're beginning to now see, well, where is sort of the supporting information associated with that? So now I'm dealing with voice management of information. Do I have time to sort of look for everything in this linear plane? I need a piece of information associated with a thing that I'm seeing that's associated with real time. That is actually happening.

KENNEALLY: So you're talking about the VR environment?

RIGGSBEE: The VR environment. One of the challenges they're having in the VR environment actually is they don't have the supporting information. I have this beautiful visual of an augmented reality rendering, and I now need some important information, and I'm actually getting a book, but I don't need the whole book. I need some portion, some data, that's associated with the thing that I'm looking at. And as I'm turning, this visual is changing, but this plane of view is remained highly static. This is not a good way to work, and we're still getting to wearables, right? That's also a very real thing.



- KENNEALLY: It's funny we're here at BookExpo, and we're talking about virtual reality and things that really seem very futuristic. I was thinking that, Renee, when you were describing getting publishers to think in kind of content-agile ways, they need to think about possibilities. In some sense, they need to be looking into the future. That's a real challenge. So how do you do that? How do you prepare to do that? I guess that's a question for all three of you, but Renee?
- SWANK: Yeah, getting format-neutral content so that you're not tied to one particular medium, that you are creating kind of product-agnostic it's not specific to any device those are very core foundational practices. Figuring out what granularity that you want to manage your content, what kind of metadata and enrichment that could drive current and future capabilities, certainly having a way of evolving that enrichment over time, because metadata and things like that that devices today might need it will need to evolve over time. So it's about creating content that's structured and that enables lots of pieces to be built in different ways by the applications on top that will deliver it to the end user.

I think a lot of those things are basic foundational things in the way that people just manage their content. Think of it as lots of Lego blocks that can be enriched and built and combined in new ways that today you may think about those in one way but later be flexible enough to be recombined in different ways – maybe to support a derivative market, maybe, as you said, that might be contextualized to a location. Maybe 80% of that book is the same across all markets, but maybe there's derivative parts of that book that can specify and be directed toward certain markets.

So it's really thinking about content creation in a much more lower, granular level, agnostic of any sort of way that it might be represented to end users, so that you leave it up to the devices to grab the things that are relevant based on maybe reader behavior, maybe on contextual location, or maybe you're putting together something fixed for a particular market. So I think it's just making sure that you're creating that content in a flexible way, which really means that you need to be thinking about these things upstream in an editorial process from content creation all the way up the chain, which I think is what you were saying.

KENNEALLY: Yeah. Ganessan, I was going to say about that that with Alfresco, it's an open-source architecture, and that's critical to this. We're not talking about locking people into proprietary architecture software. We're giving them the opportunity to think in those open-minded ways that Renee was just describing.



PARAMANATHAN: Yeah, I think just to extend a little bit on that, the context is really important. I call them as – think about it as content is your muscle. Context is going to be your nutrition – you got to feed the muscle and enrich your muscle. So we are seeing a huge trend, as Renee mentioned – context is going to be really critical to (inaudible) nutrition and feed your muscle.

And then, at Alfresco, what we have done is we created a common platform to capture your – I call them as learning objects. As Renee mentioned, so you have the ability to create as a granular chunk – be agile on how you create the content. As you have that information, keep it as a small, small granular chunk of content. Capture that. And also have a common – I call them as a digital process layer, not as a one digital publishing pipeline for print, another publishing pipeline for print – instead of having two separate pipeline, think about it as having a common digital workflow to manage the print and also to manage the digital version as well. Have the common model. And then have that publishing framework irrespective of your multichannel, omnichannel experience, and don't think about – don't hang up on your format. Have flexible on your format but just work on your common model. And then at Alfresco, the downstream – we take care of that publishing into your omnichannel experience and everything.

Also, at the downstream, once you go further, we can put that artificial intelligence and machine learning into that on how do you repurpose your content which you already have? Let's say you are creating a different version for a different culture or a different set of audience. Use your machine learn and let the machine learn automatically create the skeleton based on repurpose the content from there and then publish it to the downstream and also – take an example of a translation – use some of the artificial intelligence and the machine learning to create that different language version of the content. And again, don't hang up of your format. Just think out of the box.

SWANK: Yeah, I wanted to also add to that, because one of the things – what I was talking about – was not in lieu of print. Print's not going away anytime soon. One of the things that I work with publishers is how do you create a workflow that can support both the flexibility to have the kind of mix and match to support that kind of end user digital experience as well as still print the traditional book?

Single-source publishing is that kind of concept, where you have a single source of truth that can serve both print and different media. So it's not in lieu of the



traditional print. That's still very much part of the kind of projects that I work on and clients that I work – the publishers that I work with is that print is still very much part of the process. But it's about changing the focus of the editorial process. Instead of it being very print-driven and then digital an afterthought, it's about being more content-centric from the beginning, and one of the delivery channels among many is print. It might be tablets and mobile devices, and it might be a digital platform. It's that you have a single source of truth that is where the content is managed, and it can push out to any delivery platform.

RIGGSBEE: So there's one thing I do want to add to that, because the way you phrase that is actually somewhat interesting. So print is a medium, book is a medium, tablet is a medium, smartphone is a medium, eyewear, so is watch. But then there's another component to that, which I think we overlook, which is the medium and the user inform an experience expectation. Therefore, not only are we talking about sort of, if you will – we use the term atomization, but you can also think of it as organizing information by topic – same idea. What can this watch support? I actually can't deliver the same information to a watch that I would deliver to a smartphone that I deliver to a tablet or to eyewear. And what's the context under which I'm using this, right?

So we have to now understand user experience as part of this as well. That's where that looking at it from the downstream first starts to play a very important role. What's the circumstance, scenario, that someone is likely to be consuming this type of content? And what does that mean if they're on device A versus B versus C? I may be able to inform more. I may actually have to inform less, but the less now has to have a lot more bite to it. So this is what becomes very important in that granularity of information. At Gadget, we like to talk about it's very easy to create big things out of small things, but it's really hard to make small things out of big things. It's a very difficult thing to do.

SWANK: That's a great point.

KENNEALLY: And I'm thinking, as I'm listening to the three of you talk about this, that there's an actor in this that we're perhaps leaving out and I want to be sure to include, and that person is the author – the creator of a lot of this work, which sounds like it's a lot more work than it's ever been if they're going to try to think about if I'm writing – you know, leaving aside fiction, anything else, from self-help to topics on any science, whatever it might be.



Traditionally, a book has a beginning, a middle and an end. That's what you learn in grade three. It has paragraphs and sentences and things that are sort of innate structure, grammatically-driven. But what you're all describing is a kind of a content that sounds to me like it's a cat that's gotten out of a bag. It's not contained any longer. It's out of that container. The format, the – what do they call it, the form factor – all these things have vanished, and suddenly the content is out of that bag, is no longer under the control of the author. It's even also not under the control so much of the publisher. How unsettling is it for authors and publishers to start ceding some of that control to the outside world – to the audience, to the reader, to the device makers who drive a lot of this? Do you hear about people's resistance to this notion of the loss of control?

RIGGSBEE: It's an interesting question. I think from an authoring point of view, it is very important for someone to clearly articulate an idea or set of ideas as they need to put them together and not really try and say, hey, you want to maybe instead of using 10 words, use one and a half? (laughter) You really don't want to do that. You want people to express themselves.

I think the trick now becomes in a different part of perhaps the editorial process. When you're pulling information to deliver it to different devices, you have to be aware of what this might mean. There's something tricky here, and there's an intermediary between the author and the consumer, which is going to be that editorial process. That, weirdly enough, could fork a bit, because you could have people that are really well versed in how to put something on the web, don't understand anything with what that means to put it on a mobile device.

SWANK: I think there's a lot of work that needs to go into kind of defining what that minimal level of granularity is, and it depends on the content. We work a lot with publishers to figure out what that right level is. There are certain parts that it just doesn't have any meaning without the context, so down to a sentence level of granularity in most cases won't make any sense. But down to a paragraph level might make sense in the context of legal books, because as maybe a lawyer's looking for a particular answer to a question and they need to do a search and they need to get right down to the point of a paragraph that describes that piece of information. But maybe for another kind of book, maybe educational publishers or trade might need a much, much higher context for that information to have meaning.



So I think it's not a one size fits all of determining what that level of granularity is. There's a little bit of understanding the content itself and where the minimal kind of boundaries are. And then it's also looking at the usage of that. What do users need to search, and at what point in the content do they need to get to an answer? And that will drive what granularity. So I think there's not a one size fits all that you were just talking about granular content, and there's a blanket industry agreement on what that granularity is. I think it depends a lot on the content.

- RIGGSBEE: And there are other tools, right? So something that I might deliver entirely as written in a different context won't make sense to deliver that same information as written. So maybe I deliver it as video, maybe I deliver it as audio, or maybe I deliver it as something else.
- KENNEALLY: And I'm sure that comes as a relief to people that there is a limit to this, right? As you say, Renee, we can't break this down to single sentences or maybe single words. We don't want to go that far. Control, though I want to come back to control, because it's an interesting sort of yin-yang thing, where again we have to give up some control, but also we need to assert more control, and the way that publishers can assert more control is over this data this content that's been broken down. And the way to do that is through the metadata. I wonder if and I see a nodding head, so maybe, Ganessan, you could start by telling us about the importance of metadata and what is you know, is more better, or is it just better is better, if you know what I'm trying to ask?
- PARAMANATHAN: Yeah, yeah. I think that a lot of art and science goes behind capturing the metadata, and it's a journey, a walk. Start with the small. We strongly encourage our customers to start with the small. Also don't throw that content without metadata. Why? You cannot find them later. There is no contextual information. So you need to have some governance around those – as Renee mentioned, the chunk. I call them as a meaningful building block. So have those meaningful building blocks and have some governance on top of those meaningful building blocks to capture and take the metadata and then enrich the metadata using machine learning or artificial intelligence, that kind of modern technology. Not all the metadata needs to be manual. Some of them could be manual. Some of them could be machine learning, which could automatically enrich the metadata throughout the lifecycle about the content.

And those metadata is going to be critical to get some insight. As you collect the metadata, once you publish it to the downstream, you may want to know how many



views happen on the particular content, how many shares happen, how much time – you know, what is the timeframe that a consumer has watched? So you want to capture those information and don't throw it away. Those are really critical information. You want to get that information, back it up and take it and put it back into your metadata. Why? So those information, along with the metadata, would help you to decide what is your next best thing you want to do. The next best thing could be what is the new product you want to go and publish on that? What is the new article you're going to write on that? What is the next best marketing project – you want to do that? Those things will derive from your key metadata information, so it's really critical to capture that along with your content.

KENNEALLY: Sorry, Renee, it looked like you had something to add.

SWANK: Oh, I was just thinking about some lessons learned over the many years in working with publishers about metadata. You had asked about more metadata, and I was just thinking about more is not better in the case of metadata, because metadata has to be the three Cs of metadata – complete, correct, and consistent. Without those characteristics of metadata, it's almost just as bad as having no metadata.

So one of the things – certainly automating metadata can help improve the quality, but certainly any manually populated metadata – there has to be that governance that you talked about to make sure that it's good, because the devices that are going to be using it – if there's bad metadata, you'll get false positives, you won't find the things you're looking for, so your users could be very frustrated by poor quality metadata.

KENNEALLY: Right. In fact, it would almost be worse for you to go down that road.

SWANK: It can.

- KENNEALLY: I just want to back up, because again it's a very futuristic discussion, and I'm enjoying it a lot, but there are some terms that perhaps the audience might want to know better what you mean. Ganessan, you're talking about machine learning here. In the context of publishing, what kind of machine learning are we talking about? Can you expand on that?
- PARAMANATHAN: I think doing some automated repurposing of the content using machine learning, automatically reassembling your content, producing from



language A to language B, English to French, you could apply the machine learning. Also, you could apply machine learning in the analytics and the insight. You have 20 years' worth of content. What you can do using machine – try to do some deep analytics inside the pieces of content. You could look into historical views on the consumptions on the content and find out, what are the key interests? It's very hard to do those manually. So you can apply some of the machine learning to find out, what are the areas of interest? What are the next best thing you could do that? That's one thing.

Or very basic – I always like to go back to basic – you could apply machine learning to automate, digitize your workflow. I've seen customers that still – there are manual reviews and approvals are going on. You could apply some machine learning on that – that's a low-hanging fruit, I've seen some huge success in different customers – and apply the machine learning in some of your reviews and approvals as well. That's a huge area.

- KENNEALLY: Right. Maxwell Riggsbee, I want to ask you about that. At Gadget Software, the analytics piece of this – the work that you're doing to break things down to know which kind of chunks ought to show up in which device in which way – that also does generate a lot of analytics and allows you to understand better so you can improve the process, work more efficiently with your customers. Talk about that virtuous cycle there of change and analysis and change again.
- RIGGSBEE: Yeah, sure. So you're interested in two pieces of information. One piece of information are the things that people select, so one of the things that's very interesting when you make information small, you really start to see what people really like. I mean, this is no different than when somebody does a like, right? It's like, I liked that, right? So this becomes very, very important.

But also there's something that you guys are both saying, but none of us use the word. It's feedback. Bringing feedback into both the authoring and the publishing process is really important. Many people use this. This is great. There are people that are giving us feedback that it seems right, but I keep getting stuck at this particular part in the thing. Like this ability to receive feedback, understand which pieces of data are working, other pieces of data that are not working – some of that is done by machine, which is great. Interestingly enough, though, some of the best feedback you get are still from people who decide that they're actually – pardon the expression – put pen to paper and actually tell you, you know, it might be a little better if you did this instead of that.



So I think from the digital component, broadly speaking, that ability to get feedback from individuals and then automate through machine learning which then can influence artificial intelligence – it all starts to become a loop of improvement. All of those things need to work together. That's the real benefit of when you're using a content management system, and then you have a strong feedback loop, and then you're collecting these analytics – and 100% right on the machine learning to – once the algorithms become smart, it's really amazing the kinds of things that they can do that you go, oh, wow, I never would have put – I can't see through that much data. I never would have put that together. So yeah, analytics plays a huge role in informing the creation of better information.

KENNEALLY: We've been chatting with Ganessan Paramanathan from Alfresco, Maxwell Riggsbee from Gadget Software, and Renee Swank at Ixxus. Thank you all for joining us. My name is Chris Kenneally with Copyright Clearance Center. Have a great day.

END OF FILE