

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

**Interview with Martha Anderson, Director
National Digital Information Infrastructure and Preservation Program
Library of Congress**

**Recorded at [NFAIS 2011 Conference](#)
Philadelphia, Pennsylvania**

**For podcast release
Tuesday, March 22, 2011**

KENNEALLY: In Philadelphia, the home of America's great first innovator and inventor, Ben Franklin, this is Chris Kenneally for CCC's *Beyond the Book* podcast attending the 2011 NFAIS conference, and joining me right now to talk about her presentation Martha Anderson, director of the National Digital Information Infrastructure and Preservation Program with the Library of Congress. Martha, welcome to *Beyond the Book*.

ANDERSON: Chris, thank you very much. I'm glad to be here.

KENNEALLY: It's a pleasure to chat with you because you had a great program just earlier today looking at the ways that the Library of Congress is going after what you have called at-risk digital content. And I guess I'll ask first to define what that means, at-risk digital content, and if it's at risk, what's it at risk of becoming?

ANDERSON: At-risk refers to a lot of different characteristics of digital content. For one thing, we know Web pages come and go. Sometimes we lose files on our own personal computers. We understand that. Many of us may have lost personal photographs. That's a personal reference, too, because our hard drive crashed, we forgot and left our camera somewhere. So there are these sorts of things.

But digital information cannot just take care of itself. We can't put digital information on a shelf like a book, and the book might last if the temperature is right, the lighting is right. Over time, it would last a really long time. But digital information is at risk, but what's important about it is much of it is now being published and disseminated by government. It's being published and disseminated by people who have important things to say. It is recording our public events, world news, things like that that people have always studied and looked back on. It also is capturing this kind of social change that in the future, we think cultural historians will be quite interested in.

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

KENNEALLY: It's a very ambitious project with a very ambitious name, the National Digital Information Infrastructure and Preservation Program. It's a mouthful, and you have one way of looking at the acronym for that, but why would someone care down the line? You say that the future is going to look back and sort of see a lot of social change. Tell us how one particular aspect of this tremendous archive could help historians in the future.

ANDERSON: Well, for one thing, I'm going to refer to the Twitter archive.

KENNEALLY: One second, now. The Twitter archive?

ANDERSON: Yes. Last April, Twitter announced that it would be donating its archive from 2006 forward to the Library of Congress for us to take care of it. And at the time, we had a lot of discussions about why would the Library want something like Twitter? And we talked to the staff at Twitter and said, what do you think this is like?

And they said they like to think of it as broadsides, the kinds of things that were put on walls and fence – lampposts in Colonial times to let other people know what was going on either politically or socially or even personally, like lost and found notices, that sort of thing. So, they said, it's like broadsides.

Well, the Library of Congress has collected broadsides. We have a huge collection and it's used quite extensively by history book publishers who come to look for illustrations for different events. But it's also used by researchers who want to understand cultural, political – economic, even – effects over time. That's the sort of thing these resources would be useful for in the future.

But they're already being useful because the lifespan of many digital files – things that are distributed on the Internet – is so short that they disappear very quickly sometimes, and I'll give a very concrete example.

After Hurricane Katrina, there was a congressional – there are congressional committees that study different things, and the committee that was looking at disaster response had posted on its website reports from FEMA about the response to Katrina. Just a few months later, our Congressional Research service was looking for those reports and they could not find copies anywhere except in our Web archive. We archive the congressional websites once a month and that's where they found them because they needed to look to them as a basis for another report they were doing.

So those kinds of things are quite useful even now.

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

KENNEALLY: Right. And you mentioned that a great deal of the digital creation that's happening is taking place in the public sphere. Public policy is being documented, government actions are being documented. How is that information then being reused today and in what ways is that helpful not just for the public good generally, but for the good of researchers and others?

ANDERSON: The research community around public policy information is vast and it does include citizens who want to know how their laws are made, how were decisions thought through, that sort of thing. There are many Web resources that provide that kind of information.

At the Library, we have a Congressional Research service who uses those resources daily because they are constantly providing Congress with background information.

I would believe that almost no decision is made across the country at any level of government without someone saying, what did we do before? And so looking back at precedent, looking back at the experiences and the decisions that were made earlier, be they to be sustained or be they to be overridden, it's very important to know what happened before.

KENNEALLY: And the kind of information can be adapted, it can be converted, it can be digested into new forms, and one of the pieces that surprised me as part of your presentation was the way that what you call geospatial information is really growing – ballooning, if you will. It's the old atlas, the old map, but now done in a whole new way. Talk about that and why that's such an important piece of this collection.

ANDERSON: Geospatial information at the beginning of the program was just considered another form of scientific data, but in the last seven or eight years, we now each encounter geospatial data every day of our lives. We use the Web. Many of us don't go anywhere without consulting a map request on line or looking at our smartphones to see where are we and what's around us. So there's that personal location.

But there's also a lot of data in geospatial data that contributes to an understanding of a lot of different social and economic circumstances. And the illustration I showed in my presentation was a map that was prepared by someone who works for our Congressional Cartography service. And they prepare maps for Congress to use in their hearings, in their thought process around certain issues.

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

Geospatial data includes demographic data, it includes economic data, not just map locations. But then the applications bring them together and it gives us a picture of place and time and levels of effort and how the place and time may affect a particular circumstance.

KENNEALLY: And the relationship of one set of data, one set of information, to another set that perhaps someone hadn't thought of putting together before, but then someone, because they asked a new question, put those two sets together and created something new out of it.

ANDERSON: That is the real, I think, the real value of this for the nation ongoing, for each of us as citizens, for each of us as students or researchers or just people who are interested in life is that ability to see things come together from different viewpoints, from different kinds of disciplines, but yet have a whole new discovery there.

I think that is the potential of this program for the country, is it will help drive this kind of innovation that we want. We want new thought. We need new ideas about how to solve our problems and that's where this data comes in.

KENNEALLY: Right. We're talking with Martha Anderson, who is the program director for the National Digital Information Infrastructure and Preservation Program, which I think you call In Deep.

ANDERSON: Yes. The acronym is referred to as NDIIPP, and so many times, people will just call it NDIIPP, not Digital Preservation or National. And one time, the librarian of Congress actually said, well, we could say that we're in deep in this topic, and we are. We actually are very, very immersed in the topic and in the effort.

KENNEALLY: And I would imagine it's changing the Library's relationship to publishing because the sort of traditional model was every time someone published a new book – an author or a publisher – they deposited a copy with the Library of Congress. They got the copyright registration. Of course, that matters to Copyright Clearance Center. But today, that whole notion of publishing is of course itself undergoing tremendous innovation, and what it means to be a publisher, what it means to be an author. And so therefore, the kinds of material you're collecting are changing too, and it must be changing your relationship with those creators.

ANDERSON: It absolutely is. One thing about copyright that has always been so valuable to libraries is that through the publishing process, we would understand

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

that something that has been published and has been registered for copyright has had some process of vetting, some quality applied to it, that sort of thing.

Now, with materials and information being published just freely on the Web, it becomes much harder for libraries to ascertain what is authentic, what is valuable, what really should be saved to last. And at the same time, we're also trying to understand how do we deal not with just an organization like a publishing company, but with individuals who may create some very valuable information.

An example of that are legal blogs. Our law library a few years ago realized that legal discourse was happening not in journals anymore, but actually in blogs, and they've identified over 100 blogs that they consider authoritative from people who really are great legal minds who are writing these blogs. But yet, we have to now negotiate with the blog writer to get the rights to collect their blog and make it available through our collections.

KENNEALLY: What a tremendous challenge you have now. There's a tremendous challenge in the amount of information, in the amount of publishers. It comes down to being a communications challenge and it goes well beyond the Library itself, that great resource. It goes beyond the Beltway. Tell us about that. How are you addressing that communications challenge to get this into your catalogue?

ANDERSON: From the very beginning, even in the legislation that created the program, it was understood that this would not be an effort only by the Library of Congress, that we needed to engage stakeholders from the publishing community, from the technology community, from the academic and other federal agencies as well, to bring those communities together.

So a big piece of this program – there are actually three pieces. One is the content. One is the network, as we call it, this network of partners. And from the very beginning, we began to encourage organizations to work with others rather than just on their own on these projects, on these problems, on these challenges. Then the third part, of course, is the technology, the tools, the kinds of services that people need to do this work.

But this network has grown to be over 180 organizations, mostly libraries, archives, government agencies, but some private sector. We have the Motion Picture Academy as one of our partners. We have photographers associations as our partners. We have music producers as our partners because we do collect that kind of information at the Library, and through copyright, we have collected it. So working with these organizations to understand how do we take care of their digital

Beyond the Book

A Podcast Series on the Business of Writing and Publishing

information. How do they take care of it until they give it to the Library or give it to a local archive is part of the challenge.

So it's a community-building effort as well as a content collecting effort.

KENNEALLY: It's a tremendous task and we wish you well with it. We've been chatting for *Beyond the Book* today in Philadelphia with Martha Anderson, director at National Digital Information Infrastructure and Preservation Program with the Library of Congress. Martha, thanks for speaking with me today.

ANDERSON: Thank you, Chris. It's been a pleasure.

KENNEALLY: And for all of us at Copyright Clearance Center, this is Chris Kenneally wishing you a great day.

END OF PODCAST