

Brian Graney: Before we begin talking about how we follow from here and what kinds of projects we might want to consider as pilots, what kind of collaborations have taken seed here that we might pursue down the line, Mike Mashon from the Library of Congress has a brief presentation for us.

Mike Mashon: I just want to show something. As we are thinking about collaborative projects moving forward this is sort of a follow on a little bit from what we were talking about during the morning session. I'm just going to show you real quick. I mentioned this to some of you at lunch. This is a scan of a D.W. Griffith film out of our paper print collection called "The Light that Came."

We have, a lot of you are familiar with the paper print collection. These are the earliest films registered for copyright at the library between 1894 and 1912. All of those films, they were transferred to 16mm film stock in the 1950's by a man named Kent Niver.

In the 90's our laboratory went back to reprint some of them on 35mm. Those are the ones that are available online. If you go to the library's American Memory website you will see very, actually they're pretty poor quality and postage stamp sized files of about 500 titles in the paper print collection, of which there are over 3,000.

So we have a scanner now at the Packard campus that allows us to scan these 16mm Kent Niver negatives. We can get through them pretty quickly. If the guys are really popping it's about a dozen a day. We have a lot of 16mm silent elements in the collection.

So I just wanted to show this real quick. You're not going to get much information out of the physical element on these. There wasn't really a tremendous amount there on the original paper. We're still over-scanning. You can still see the sprocket holes on that and you can see above and below the frame line but that's kind of immaterial to what I'm talking about now.

The image is very, very jumpy. We're not stabilizing it. This is quick and dirty digitization. The reason why we're doing this is to be able to make it as accessible as we can as quickly as possible. This is a D. W. Griffith film that as far as I know is not available on any DVD set. It's not even available on YouTube so it just as well not exist.

These are ones that I've prioritized for digitization. So we're plowing through it but it's not just Griffith. There's plenty of other things that we've got but I did just for illustrative purpose start with about 15 Griffiths that I knew weren't otherwise available out there. But we've got several hundred more that we could get through.

One of the points that I want to make is the library, our ability to digitize content has far outstripped our ability to make it available online in any sort of meaningful or discoverable way. For this film we actually have really good cataloging on it. If you go to

the library's online catalog and you look for "The Light that Came" you will see an incredibly exhaustive cataloging record that had literally every actor who appears in this film, even in the party sequences. It's amazing and way more information that most people are ever going to need.

But it's also offset by the thousands and thousands and thousands of titles that we're digitizing for which we might only have a title. Maybe we've got the year. We've got very, very little. We got a film in the collection called "America's Treasure." That's literally the cataloging on it "America's Treasure." You have to look at it to find out that indeed "America's Treasure" is the Library of Congress. We have another film, it's like "America's Golden City." What is that going to be? And it winds up being Miami.

So it's that sort of thing, so the thing as we get discussing here, the things that I'm really interested in particularly from scholars we're looking for ways to be able to engage the scholarly community in helping us enhance our cataloging records but we're working the ways through which we can make that happen.

One of the things that I'm doing is we're participating in this media ecology project that's coming out of Dartmouth and Columbia and its sort of this secure space where people can, archives can contribute a variety of content. WGBH is contributing some things from their open vault initiative. We're just giving, I'm just giving them all the public domain content that they could possibly ingest. They have tools within media ecology allows scholars to interact with it, create new content from it.

What we're looking for too at the library is a way to use crowd sourcing for us to be able to enhance our records. I want to be able to find a way in which we can put; right now I have a lot of undifferentiated content. I don't have very much information about it but I'd love to be able to put it online and find ways for scholars and other people to interact with it, to tag it.

One of the issues that I've got to deal with at the library is authoritative data. There's got to be, we have to be able to find a way to be able to control this. There's got to be controlled vocabulary so we're looking at ways of invoking Library of Congress subject headings but all of that requires a lot of human interaction. So then we would have to sort of go through and look at what people have tagged our content with to see if that's really appropriate for the particular image.

So it's a struggle for us but we're certainly looking for ways to be able to get you to interact with the content. Maybe there are other ideas that we'll have coming out of our time here together. So, thanks.