Data Curation Profiles Symposium 9/23/2012

Panel 3: Broader Perspectives

[2:40:42]

JC: Well I'd like to introduce our panel on broader perspectives and our thinking on this is looking at beyond training and education and learning, are there other applications for data curation profiles? and are there other areas that librarians could or should perhaps explore?, sort of similar to our educational panel and looking at the theme of identity and responsibilities, what other potential exists out there for actively growing this capacity in librarians? And we're very pleased to have three folks with us, we have Dianne Dietrich from Cornell University, we have Nancy McGovern on the big giant screen behind us from MIT, and we have Holly Surbaugh from the Stratus Consulting. Nancy, do you want to go ahead and start us off?

[2:41:18]

NM: Sure. I'll jump right in, so um, Scott and Jake provided a couple or set of questions, so I sort of framed mine around that and I'll just carry on, but I'm most concerned about getting to--or interested in getting to the question part, so I'll try not to rush through this so I can hear what you all think. It was really fun to think about how you know having digital curation profiles, what we can do with that as a, moving forward. So, one of the things that I've seen is that research data, we often talk about it as though, it has, is one single monolithic thing, when in fact it can take almost any form, literally almost any form. And even to the point where we have to manage as we do many things across hybrid, analog, digital, and combinations of those things. So, one of the things that I think would be a terrific thing as we move forward is to accumulate data curation profiles, so that we can start looking more systematically at what they tell us for understanding research domains, the kinds of methods that researchers in those domains use, the kinds of results that they produce, the kinds of things we might expect to have to manage over time. So if we had that, and that is definitely on the data curation profile website, that these sort of broader sets—some of things I noticed were that the digital curation profiles may be used as to inform the development of data services and maybe used as objects of research. So I think that that's in the same line of practice that we might explore. So if we start looking across—libraries are always required to be looking across domains as a natural thing. We to collect things, to select information from an array of sources um that gives us a perspective that just about no other domain has and um would allow us to accumulate these models and look for trends in practice. One of the things that researchers in an individual domain wouldn't be required to know about would be practices in another domain, they just get on with their research. So if we were kind of able to look at, for example, someone, if a researcher performs an experiment, the kinds of things that result from that would be observations and some array of results that would um you know could take any form. A survey instrument produces survey results in a range of quantitative to qualitative results. Surveys though take place in a number of different domains, I mean social sciences use them heavily, but other domains use them as well. Instruments produce measurements of various kinds and there are loads of other examples. So when we start parsing apart what does research data mean, we should be able to look at

the kinds of, this range of results, not only within a domain but across domains, so we can get patterns of what kind of services might we build up around a particular kind of research outcome and how might we engage in a different, how might we model behaviors within and between research domains to be able to think about what they most need or how we might most help. Thinking about that information, if we had that information available, we might be able to tackle what is going to be a real challenge and we're hearing things about it already, but the application of records management principles to research data. We're not going to be able to decide what kinds of research data needs to be kept for how long, unless we have a much much better, broader, deeper understanding of the kind of research data that research results that are going, that already exist and that are going to be coming towards us. So thinking about you know, how do we decide determine as we do with loads of other content, records management has been doing this for decades, we've extended it out to different kinds of content as it emerges, but we haven't systematically done that yet for research data. So how will we be able to apply research management principles to figure out retention, lifecycle management, help parallel what researchers are doing to kind of be part of this broader life cycle. Um and based on that understanding, really figure out a set of responsive services that we could begin to develop collaboratively and um at our Individual organizations. So, that's about that uh u m, the sort of needs around data curation, definitely understanding the domains better, so that we can build better services around them.

Some of this sort of practical applications of around data and in relation to data curation profiles is that some of the things that we notice when we were doing the e-sciences institute, Jake was part of the faculty with me, it was a great opportunity to look across what a lot of institutions are doing. But we did notice that there were you know several things that libraries tended to be more comfortable with in terms of management and data services their data curation kinds of activities. The first was to take custody of content. So through an accumulation of data curation profiles, we should have a much better sense of the kinds of things are being created, but we may or may not be the proper custodian for that over time, we may not have the scale, the skills, but it doesn't mean that we can't have a role in the long term management of research data, we just need to think about what that role might entail if we don't have direct—we do, we know a lot about policies and other kinds of actions to help manage content over time and the second area that they were very very comfortable with was metadata. So we seem to be kind of focused in ways that we are used to rather than, I think we are going to have to really look at metadata management, broaden it out and turn it out and research data might require to help manage it for as long as research domains need. Right now we don't really know how long researchers within various domains really need or want to have access to outcomes from their accumulation of research and we need to have a better understanding of that. And finally, um I think that if we were to have this array of data, we could—you know, I was really struck on the data curation profile website by the you know maybe use as objects of research further a better understanding of data types, um if we were to have this accumulation of information, these profiles that are built upon individual data curation profiles, the models and how to better understand it, we should be ready for things, cause it will continue to change, continue to have different types of content types, and to have different demands for us to meet, but one of the things that we really help is to have a series of brainstorming sessions that across any opportunity we can to have for conferences to really look at getting um beyond things that we might currently be really familiar with and really able to do to look across and develop services that

people will realize they really want to have. So anyway, those are just my thoughts and um based on the task we were sent, and I'll look forward to the questions at the end of the period.

[2:49:19]

DD: Okay, let's make sure, is that okay for the microphone? I'm getting the thumbs up, so I'm assuming that's okay. I'm Diane Dietrich, I'm the physics and astronomy librarian at Cornell, and when I thought about this question I thought it might be good to sort of give an overview of my trajectory at Cornell because I've actually been involved in research data curation since I stared at Cornell 4 years ago, but in two separate positions. I'm getting a little bit of feedback so I'm moving that back [referring to the microphone]. Okay, so the first position I had was research data and metadata librarian and that was focusing on data curation issues from a technical services perspective and in that role I did not have a connection to a particular department like subject librarians do, and then two years later I took the physics and astronomy position where a component of my job, in addition to all the things that we traditionally consider as part of a subject librarian's job you know collection development, reference, instruction, was also assessing and understanding emerging data curation needs for the physics and astronomy departments, so that will come up a little bit later, but I just kind of wanted to give that background as to where I'm coming from an institutional standpoint. So for the data curation profiles at Cornell, we actually have a little bit of a different experience. So we've been in, we did data curation profiles Cornell as part of a project that had started quite a number of years ago called Datastar, which was originally conceived of as a staging repository with a semantic web approach to metadata management. We're in the second phase of that project now and it turns out that the metadata management part was a little bit too complex for the time that we tried to fit it all in so that the next phase of the project is more focusing on a data registry for the institution. And we did 8 profiles across different disciplines at Cornell and multiple librarians did them. And the goal of doing those profiles is to look across them to form developments to inform the development of Datastar as it's in its second phase. So when we think about broader perspectives in terms of what can we do beyond outreach and beyond training, you know, if you are building infrastructure in your institution the data curation profiles are one way to sort of inform the development of such infrastructure because it is a way to compare different answers to questions that are important for such things as a repository for instance. You know, what sort of functionality does the software need? What are the actual needs in the department? You know, across the researchers, what are they actually saying? What are they actually doing? Rather than trying to make a guess or, you know, base these things on conjecture, right, so what should the repository support with respect to embargos? That's really important for a lot of researchers who want to make sure they protect their data before their results are published because they're worried about other people getting their hands on their data and fear of getting scooped. What sort of access rights do they want? You know, some data comes from privacy and confidentiality concerns so maybe there's some researchers that want to protect sort of or guard access rights to certain data. What's also needed for long term preservation in the context of you know developing infrastructure and you know one of the interesting things I think we noticed in when we were doing the profiles is that a lot of the researchers made comments about preserving things in the long run, but then didn't quite see the value in some of the pieces that actually support long term preservation. So there was sort of a disconnect

there, right, in terms of what they want to see happen and the components that actually are needed to make that, to put that into effect. You know it's like when you develop software you need your use cases, so the data curation profiles are really good use cases for this type of work because we're asking people to focus on a particular project, we're taking, we're not asking researchers you know when you do the interview. We heard that in the very first panel, on you know the experiences that it worked well when it was focused, when you asked the researcher, "Ok we're going to ask the researcher about this one particular project. We're not going to ask you to think really broadly about your data management practices" because that might not yield the same kind of results as if you ask someone to really focus because then you can ask specifics and then you can really get into what's going on.

More broadly I think, in terms of an institutional context, we can use the results that we get to advocate for institutional policies that support what we're hearing from researchers and what, where the gaps are. So, intellectual property concerns, do we have enough language or do we have people who can answer really detailed questions that researchers might have about safeguarding their intellectual property with respect to data. Uh security, privacy, confidentiality, like a lot of us have been saying, sometimes that expertise might not be in the library, but you know it's always been a function of our jobs to refer people to the appropriate experts, right, so being an advocate for those, for policies that would support the things that are needed is actually, is important. You know, sustainably supporting long term archival, um, how best to cite data. And I was thinking about this a little bit more, and I've decided to put on my subject librarian hat for a little bit because a lot of us have been talking about engaging subject librarians and doing this work and I feel really in a good position because I, that is the same person,--I can work with me right [laughter]

JC: One would hope.

DD: One would hope! Right! Umm but a lot of us in, who are subject librarians, and especially I think I notice this trend in the physical sciences, we have relationships with publishes because we work with them to acquire the materials that are in our collections and a lot of us are on publisher boards and we advise on various policies and so if publishers are thinking about data, and they are, you know they're aware of these requirements, you know, that's also another avenues for us to provide input. Are they thinking about accepting supplementary data or are they are already doing it? What are their policies around it? What file formats? You know, so we have an opportunity to be heard in that venue too because I think the publishers are thinking that they want to become, maybe some of them are thinking that they want to players in this arena too. Do they want to allow people to link out to external repositories and do their systems, you know a lot of them are developing really nice systems for searching through all of all of their journal and literature content, but are they actually providing a good means to link out to external data where people would want to do that? And maybe we can consult on some of the things that we're seeing in our institutions to inform the needs to inform the development of these systems? Another thing that I thought about was you know, a lot of the standards where some of it is it is happening in scientific societies or other organizations and how can we be involved in developing standards with the researcher with the practicing researchers, right? I'm sure a lot of us when we were going through the profile we noticed, what kind of standards do you use or do you know if there are standards available in your fields? And so you know obviously we can help people select an

appropriate standard, but maybe for subject librarians especially who are involved more closely with their disciplines, they can get involved in activities that support the creation of or the development of standards. Um, what else? I had one other things, oh as a subject librarian, I have also found it very helpful to people to cultivate relationships with faculty members, you know to engage them and say you know the library is interested in these sorts of things and have this conversation about it. And I think other people have also talked about the important of having a data base or a repository if you will of profiles where we can compare. We're also thinking about policies that funders are putting in place, requiring the, you know the PIs on projects to have a data management plan. So what can we take from that that might form how we help and consult with folks on data management plans. So those are some of my thoughts, so I'm going to turn it over to my next presenter and we'll have questions later.

[2:58:38]

HS: Hi my name is Holly Surbaugh, I work for Stratus Consulting, an environmental science firm with offices in boulder Colorado and Washington DC, and specifically my role there is to provide information management services to various Stratus clients. I found my way on to this panel today because of one particular ongoing project, so I'm to start by giving some details about that project that will hopefully, well will hopefully provide some helpful context. One of our clients is the Environmental Protection Agency's Office of Research and Development; I'm going to refer to them as ORD throughout. We support ORD in their efforts to develop a scientific data management policy. The intent of developing a policy will be to introduce a consistent organizational approach across ORD for managing ORD's data. This is increasingly important because ORD data is generated in huge quantities and it's very diverse. ORD research covers a broad spectrum of disciplines, ranging from ecology to human health. And as Chuck Humphrey brought up in his presentation, there's a difference between just writing a policy and just doing everything that is needed to implement the policy effectively. So in trying to craft the policy and all of the procedures and guidance and outreach that's bundled with that to make it make sense, and to make it beneficial to the individual researchers and to make sure that ORD as a whole benefits from it, it became readily apparent that we needed from the beginning that we needed to get a clear idea of what we were getting into. How were ORD researchers managing their data? What problems are they encountering? What ideas do that they have to share? How are they actually going to react to a new policy? And that's where the Data Curation Profiles Toolkit came in for us. We did modify the toolkit materials somewhat because we were using them for a very different purpose than the creators intended. We were trying to accommodate a different scale, so we still were collecting the individual researchers accounts of their activities, but we were also aggregating that information to try to create really a profile of the current state of data management at ORD as an organization. So, during the summer of 2011 we took to the field and we conducted over 60 interviews with researchers across the country and instead of producing individual profiles for each one, the interview results were collected into a data base. The other product of the interviews is this fairly substantial guidance manual. It's very specific to the ORD environment and it's designed to help ORD researchers navigate their information needs and their data management needs according to what they have available to them. And the guidance was developed using information and best practices gathered in the interviews. Now after working through the majority of 64 of these interviews, it seems completely evident to me that

scientists really could use the help of information professionals [laughter] when they're managing their data. Scientists are brilliant at what they do, but frequently managing data requires an entirely different skill set. Also I think librarians, sadly, could empathize a lot with the do more with less environment that government scientists are currently working in. The researchers that I spoke to are not just responsible for producing defensible research or advancing scientific inquiry, they're also responsible for increasing a lot of what they term administrative burden. And if you're going to talk to scientific researchers that word is going to come up eventually somewhere, probably repeatedly. Every time that you try to discuss with them managing their data effectively, administrative burden comes into the conversation and the discussion gets wrapped up in that reality. In a perfect world, my personal opinion is that all research teams and especially large transdisciplinary research teams could have an embedded data manager within the team, someone to provided that unique skillset and to relieve the scientist of that task that doesn't always come naturally to them. In the absence of that capability, I think there's still opportunities for librarians to work with researchers to try to resolve their data management needs with the resources available. Particularly in agencies where those resources are in flux, right now new tools and approaches are constantly being developed. It seems like every 5 minutes something new comes down the pike. And I think helping scientists in organizations stay on top of what's available to them is very important. One of the exciting things about this project for me is the opportunity it's given me to collaborate with other people like myself working on the same issues at another federal agencies particularly through ORDs involvement in Cindy(?) activities, my perspective is that it's not just ORD or it's not even just EPA, all federal agencies that I've come into contact with are grappling with these issues to one extent or the other. Scientific data management is definitely an issue, where everyone's talking about it and from I've see it doesn't seem to me that librarians are overrepresented at the table. Uh quite the contrary, and I'm not entirely certain why that is, because coming from someone like myself with a library background it seems like you are ideally suited to help scientists and the administrators that govern scientific activities figure out what they need to do. I would definitely encourage librarians who are working in or around government agencies to be proactive about trying to overcome any institutional silos and really insert yourselves into an important conversation. And so that's, I kept it short and sweet so that there would be plenty of time for conversation, I look forward to your questions.

[3:05:59]

JC: Great, thank you very much, thank you all three of you. Fascinating presentations, I think that I certainly gained a different perspective on how to use the DCP and how you use this in a larger context. I'm trying to sort of suss through an actual question, but I haven't quite gotten there, so I'll start talking and maybe a question will come up as a part of this. So I'm sort of struck by in comparison to the last panel, there was sort of a theme of identity and what is the identity of the librarian in this sort of newish role of data curation and working with research data and I think you three sort of have a theme that's sort of in a sense responds to that somewhat and the theme that I heard sort of is communication. The idea that getting that many librarians to talk to faculty and learn more about what we could or should be doing to help in developing services but getting researchers even to talk to each other and connect with each other. Chuck had this sort of set of slides in his keynote where he presented this kind of

honeycomb structure with some very solid walls between different research projects that were going on in the same program, but It seem liked they didn't really talk or interact with each other very much. And then you got to a later slide where you sort of amorphous blobs that sort of over lapped and consumed each other in some senses. So looking at breaking down those walls and barriers to getting to more of a loosely structured place where communication could happen across different groups and one could learn from each other and not just doing your own thing and siloizing your efforts. And so I heard that theme sort of repeated from the three of you of really the need to push for acting as a catalyst almost to bring people together to leverage what they're doing into making it more substantial or more value added as an output or results. And I guess my question is that sort of what you're getting at? And does that make sense? And how would that actually happen? Or how might you see that happening at the institutions where you're a part of? That was really vague and fuzzy. It was.

DD: I'll take a stab at it first, maybe. Alright, I'll do that. Yeah I'm thinking just in terms of sort of the subject librarian perspective and working with publishers, you know we can take bits and pieces of what I've been learning, you know, from conversations with various researchers and bring it to the publisher level in sort of providing input on things that they're thinking about doing because that would again be another kind of fragmented piece right. So we're all kind of working separately, you know, where the library community is thinking about this one way, you know the researchers are thinking about it another way, and the publishers are also trying to get their—to be in this area and you know if we're not all talking to one another, and I guess as a subject librarian, I am talking to the researchers, I am talking to the publishers, the publishers and the researchers don't necessarily talk to one another, so maybe that sort of like the graph of communication, you know to bring together and kind of aggregate all the bits and pieces that we've been hearing about and make them into a more coherent story. Those are my initial thoughts to your big questions.

JC: Thank you.

NM: My turn? Can I? Sorry, um yeah I think, well I think connecting the dots across any boundaries within domains, across domains, within our own organization, between organizations is always a good thing. I think in addition, I was, you know and I heard aggregation come out of the other two sets of comments as well of we have these individual opportunities to understand a you know this research project in capture that in a data curation profile, but the ability to abstract that up into models of, you know, how can we understand the results of different kinds of methodologies that may be used by many different domains, how can we understand um, so the kind of results that might occur within a research domain and then the research domains are changing, the boundaries are changing between the, even as we're watching them, so definitely use those results to communicate with them, but also as a way of building our own understanding and our ability to develop responsive services and, you know, just be able to work with them more effectively. Because we're not looking at it as, it's all research data, there's all different kinds of things going on, and so how can we specialize our response to be the most effective.

HS: I don't have a whole lot to add beyond that, but I can say that as you're trying to facilitate communication among the various parties, I've found that it's helpful to consider everybody's

motivations, um, so there might be strong for altruistic reasons for the benefit of the work and forwarding science and all that to talk to each other and communicate effectively, but then there are other motivations that perhaps are having a stronger impact on scientists' behavior that's keeping them apart, that's keeping them away from communicating effectively, or not quite motivating them to overcome burdens. So I would look for where there are walls either real physical walls because there's technology challenges or something that's like in a physical way keeping people and those sort of more, uh softer, harder to pinpoint psychological things that are keeping people from having the discussions that they need to have.

JC: I'd like to open the floor up now for questions.

[3:12:17]

LH: Hi I'm Lorisha(?) Hart from the University of Michigan. And my question is mostly for Katie, but anybody can chime in. It's a much smaller microcosm of the big question that we just talked about. But specifically, I'm hearing in different places when I have conversations with people in different institutions, who are librarians who um approach some sort of data services but are not liaison librarians that they need more conversations with liaison librarians, um, and they need some more shared vocabulary in order to understand the other side of the coin. And that's probably true too with technical services librarians. We have some people who are probably gold mines for metadata you know information who aren't necessarily at the same conversations that uh I'm participating in. So since you've kind of worn a couple of hats on either sides, do you have any ideas about besides special interests groups or mandate from administration, how to bring different people together to kind of work as an organization to meet all these goals?

DD: So one of the things that we do at Cornell is we kind of have an informal data discussion group so that's one way that I've found that has been really helpful to kind of bring people together from different sides of the library. So basically we try to make the meetings as informal as possible, right, and we encourage as many people to come as possible, sometimes we bribe with food, that is helpful actually. You know, when we're talking about doing more with less, you know, a lot of us don't, you need any incentive that you can get. Um, but more seriously, I think what helped us initially, so every, we run these every month, during the school year and we come up with a topic either it's a journal article to talk about or it's an issue that maybe we were thinking a lot about and we want someone to elaborate a little bit more on. Sometimes we have short presentations by a number of people, sometimes if we're lucky to have someone who's already visiting, we grab them and we have them talk to the group, and then the idea is to foster a discussion. One of the things that we found really helpful at the the beginning was to connect, to make very literal connections between the work that the subject librarians are already doing in terms of reference and social science data reference is a really good place to start because a lot of the issues about weird file formats and difficulty searching through various vender systems are things that we think about when we're also thinking about data curation and it's really easy to make those connections. So that's what I've found really helpful is to sneak your data curation agenda in there. [Laughter] But to tie it to things that are already happening, but that are still on the original terms, right, so I think we hear a lot about sort of data curation let's say from the data

curation point of view and then being told that this connects to what you're already doing. And then, I think it's also helpful to have the conversation the other way, saying, this is what you are already doing, this is how it connects to this new thing because it might be a little bit less overwhelming rather than throwing a whole bunch of new terms and concepts at people, so when you, both directions I guess, so that that would be my initial kind of thought about bringing people together. Do it really informally and start where they are.

JC: So the question was targeted to Diane, but does anyone else have a thought on that? Before we go on? Nope? Okay, great. Other questions?

SB: This one is from Sally Gore who said there are already data managers in many research teams on my campus and there have been for some time, so what other role can librarians play? Is there a different way to get to that?

HS: I think for me, coming in from a not traditional academic background, coming from an agency perspective, I think it's great that you already have access to what government scientists are sort of crying out for in the darkness, like uh, so it's also difficult like I said to um, basically from an agency background it's difficult to make sure that librarians are, because they're frequently these sprawling, if you think about all of what the Environmental Protection Agency encompasses, and there are so many organizational chart reads sort of like a little rabbit warren, trying to get the EPA library network, which does fantastic work, and just making sure that they're communicating properly the what they are with the office of science information management, which governs a lot of the efforts that we're doing with the data management policy, again it's about overcoming institutional silos and having the conversations, uh, facilitating the conversations necessary. I wish I could offer a little bit more insight from success stories from librarians who were participating actively beyond the embedded data manager role. Does anyone else have anything? Any words of wisdom for that?

NM: I guess I heard a question, as part of the questions, a question about kind of terminology as well because a lot, you know, people have, who say well I'm a data manager or a curator, they may or may not be thinking about the same thing because there's so many different parts to those words, so I don't know that, I mean, libraries I think can play a facilitating role in actually elaborating on, making explicit some of the work that's going on in terms of looking from the research project and moving forward, thinking about lifecycle management, so I think helping a group to understand how the words are used and making sure there are no gaps that need to be filled in terms of moving towards lifecycle management.

[3:19:27]

JC: Any more questions before we go in? Charles?

SB: Can we pause first?

JC: Oh, sure you need to change the tape?

JC: Okay so we're going to pause for a few minutes while we have our folks change the tape, but we'll get right back to the questions momentarily. Apologies for the awkwardness.

NM: Have you timed it? [laughs]

[3:20:11]

JC: Good! Let's role!

NM: Awesome.

CW: I'm Charles Watkinson from Purdue University again. Nancy mentioned and a number of you have mentioned the potential benefits of accumulating the data curation profiles so that we can better understand research domains once as a critical mass. And there was some discussion in the previous panel about the best presentation of data curation profiles and the balance between the current, uh, the fact that they are basically structured interviews and a lot of the richness lies that lack of very very defined structure and then the balance between that and the need for searchable database type presentation. So my question really is, um, you know, if, bearing the in mind, if you had to pick three fields say, that need to be searchable and will promote re-use and the ability to aggregate, what three fields would make searchable?

DD: I think yeah, what I would find really helpful and I'm not going to an—oh man I don't know if I can pick just three. Um, but just to be able to say can I see, like I'm just looking at mine right now, all of the formal standards used paragraphs for all the profiles and just read them all in a row and see what happens, what light bulbs go on in my head. If I were to read each one of those, you know, that blurb, so I think it's more of like a function type thing, I would like to see each part of each part of each profile side by side. Because obviously I've gotten a great deal out of reading the whole profile from start to finish and thinking about this as sort of a holistic thing. And then sometimes I have this desire to sort of want to just take one piece of one and just look at all the profiles in that one section, just to read that as sort a coherent whole as well. So I'm going to not pick three. [Laughter]

NM: I—I guess I would also, I don't think I would, uh the reas—uh yeah. I'm really interested in pattern analysis across, so I guess it's kind of like what you were just talking about. I think we have to have like a more deductive, sort of like we look at the data and look for models that might be the most interesting because they're the most different from each other. I think we know enough to know what those three are yet. Um I'm working on a project that's um, DIPIR.org and it's looking across social science, archaeology, and zoology, and there's some really interesting differences that are showing up from that in terms of how they view metadata and data and, you know, how their long term view towards what they view of their data is. So I think that it would be hard at this point to say—even the term re-use means different things to those domains. So I, I think that we should plug along on trying to see what models might emerge from the information bef—and then we'll have a better sense of which ones might be the most interesting. Anyway, I'd like to see what things turn up that are common and then what things that are different to see how that would inform what kind of services we—or approaches how our understanding of research data builds.

HS: Going with the trend of kind of ignoring your actual question, [laughter] I cross my fingers and hope that I say something that you'll find interesting anyway. It's a little unusual from my perspective because I never created like an individual data curation profile to begin with, all of our interview responses were sort of chunked out and put into a database so we could slice and dice that information according to a variety of sort of questions that we had in mind before we even started interviewing people that were relevant to our needs. A big part of one of our goals was to find best practices that we were interested in disseminating widely among researchers at ORD. Another goal that we had in mind was uh, collecting user needs requirements for potentially planning for our IT infrastructure needs well into the future as an organization and so there was an effort to sort of, regardless of what part we were uh of the interview we were in to sort of tag whenever somebody said something of interest to that. So there was a lot of sort of repeat within our database structure, well this applies to this, but this same comment also goes under here. We've also broken up according to, because one of the reasons why we're initially interested in the DCP toolkit to begin with was because the module structure fit so neatly in line with an overall SDM framework that we were developing that had various, we called them procedures, but they were sort of general topics like metadata or data storage. And so we also tried to um, we used the interview structure so that we could get a lot of rich detail and allow people to go through these digressions and then we systematically and went back though and sort of chopped it up so that we could look and see what everyone we had talked to had to say about metadata and that was very useful for our purposes, I don't know how useful, um it's an interesting question to see if that would be useful for other people thinking more from like the information science field perspective to do it that way, it was incredible useful for use because we actually have to figure out what we're going to tell people about what they have to do for metadata in order to comply with a policy.

[3:26:48]

JC: Thank you very much please join me in thanking our panelists.