**Pre-submitted Questions**

Would it be possible to for Librarian 1 and Librarian-2 to show the Researchers:

1. The Koha system
2. The E2E process of digitizing an item, if possible

Specific questions for Librarian-2:

1. In the previous work done with a previous researcher, how did Librarian-2 perceive METS to be much richer than MARC?
2. Did Librarian-2, or anyone else from the Library, have any experience of using METS to catalog items to compare cataloguing using this standard with the currently used MARC standard?
   1. If YES, how did using that metadata-type compare with MARC
   2. If NO, what concerns are there about transitioning to a new standard (e.g. access to training, conversion of legacy metadata (MARC to METS), creation of new collections, etc.)?

**=====================**

**Interviewee** Librarian-1 and Librarian-2

**Attendees** Researcher-1, Researcher-2 and Researcher-3

**Date** 12-Jan-2021

**Note(s)** Permission to record requested and received before recording commenced.

**Transcript**

**Researcher-2:** OK, yeah, so that's on there now lads.

Yep.

**Researcher-2:** OK, um so yeah thanks a million for that Librarian-2 and I'm looking forward now to looking at that. So the original questions and lads feel free to jump in now don't have me hogging it here now 'cause I'm inclined to do that so.

We were looking at KOHA system and I know that some of the information you sent on there Librarian-2 was in relation to that.

Yeah, and then there was the E2E process, but you did say in your email - you said that because you know we're not in the college, it wouldn't be possible to do that.

And then one of the questions then was in the previous work done with Researcher-A [previous MSc student] how did you perceive METS to be much richer than MARC?

**Librarian-2:** OK, I suppose I had looked back over Researcher-A initial reports yesterday and it was good to kind of get an insight into it again, and I suppose the first thing is actually something that I didn't realize at the time, but actually in our article we say that there is actually 1400 artefacts in the collection, but in reality that's not the case. It's actually there are 1400 physical books in the collection, and the archive itself possibly... I don't know, I would be guessing would be in excess of two and a half, 3000 artefacts.

Because if you look at the number of slides we have? Is it the two and a half? That is the 50 or 15 millimeter slides or the OR the 20 do you know those small projector screen slides.

There's there's thousand at least nearly Librarian 1. I think it is nearly 1000 alone when we counted them awhile back, you know and then you throw in all the factors in terms of the videos, the books, the individual newspaper articles, reports, letters, photos you see so 1400 is actually now looking back at it is probably a conservative guess. I'll put it like this if only we had 1400 artefacts to catalog but it's good to get that insight.

It came in thirty four boxes, so I had a look at as well with the number of boxes and there was 34 boxes and so again, which is a significant number. So I suppose to break the collection down into two parts. First, the first I would say is the physical book collection, so as I said there's approximately 1400 physical books across a wide range of subject matters and what links them within our collection within the library catalog is the series field “Professor Mike Cooley collection”.

Is it OK? Researcher-2, if I I'm not sure if I can actually share my screen, but I'm going to try because if I can.

**Researcher-2:** I can enable that here. Hold on now.

**Librarian-2:** Yeah, 'cause I just think it might be a good way of actually showing it and what I'll do is I'll show it within KOHA and I'll also show it within the library catalog also.

**Researcher-2:** OK.

**Librarian-2:** OK, I think I'm business, now brilliant.

**Researcher-2:** Yeah.

**Librarian-2:** OK, so hopefully you can see my screen now. Can you see the KOHA system? Yeah OK, yeah OK so this is what you're looking at here is the is our library management system. It's KOHA, it's an open-source system. We've had it probably coming up on four years, I think 3 1/2 to four years. At this stage we used to have. We used to have a product called Millennium which was by Innovative Interfaces a US company and it was proprietary software but and if I'm being honest from my personal perspective and one that is shared by a lot of librarians now is that the product, we had is probably more in fitting with what we're trying to achieve should we say, as an academic library, than the product we have now. It has weaknesses and probably one of the core weaknesses. I believe it has is in relation to special collections and archives, so that's kind of my little bit on KOHA .

I'm going to show you one of Mikes, so Mike's most famous and probably most well-known resource book is called “Architect or Bee” and so you can see I'm searching within the catalog now at the moment and it just gives you an insight. You can see there's seven. should we say 7 entries here at the moment, and what you're actually seeing is actually various editions of ‘`Architecture or Bee”, so not all editions of “Architecture or Bee” will be under the one record, bibliographic record, which in effect is a description of the biblio, the item itself and so there are various different records, as you can see, and the reason being, as I say, is that there are different editions. So, as you can see here, well, I I suppose, just for before I click into a record, if Scroll down here to the bottom, you can see there's a version in Irish and there's also a version in Japanese, so it's we have, we have, uh, we have multiple, it’s a multilingual issues of it, so I'm just going to click into “Architecture or Bee” and here you go. You you see what it looks like within the library management system and if I just click on here, I'm going to click on the OPAC view and what that will do is it will open in a new window and this is what it would look like, if you were looking at within the library catalog.

Can everyone see that. Yes.

**Researcher-3:** Yes.

**Librarian-2:** Perfect.

**Researcher-2:** Yep, Yep.

**Librarian-2:** OK so a few things I just want to point out before I go into the record itself. So, these records were created using uh, using transcript, transcribing, cataloging protocols known as MARC, which is machine readable cataloging.

And in fact, what that is just to kind of give it a quick overview is that MARC allows library systems to to create records, so it allows library systems to record particular facets of information about a resource under structured fields, so the MARC number say within a record so that particular um, that 3-digit code, is understandable across various library management systems and catalogs and so forth. So what it allows- it allows consistency.

So, it means this here “Architecture or Bee” is under a a field most likely being 245, which is the title field within MARC. And if you were to go to other catalogs you will see that those books or resources titles, would generally be under 245 there sometimes there's variance and that could be 246 or whatever, but generally the vast, vast majority of bibliographic records the title information will be contained in the 245 field and that that when we look at the record in a second, that will make more sense. And the second thing that I just want to point out while I'm here is the series field and this can be under 2 MARC numbers so it can be under 440 or 490 you don't really need to know why that is, but just so you know that they are the codes within MARC that allow us to insert a series field. Now why is the series field so important? Well, the series field is so important because what it does it it allows us to add all.of Mike's various books from his library into our main library, catalog. And how do we join them all together? Well, we join them all together using the series field.

So that allows us to link disparate items that could be Mikes. If you look at Mike's collection of books, they are so eclectic, you know you are not going to just find books on IT or so, or should we say computer assisted technologies, or so forth he is actually or human centered technology his reading varies right across the spectrum of Dewey, which is our decimal classification system, so his interests are varied. He is so well read, and he is so knowledgeable with so many interests and, in turn, it's understandable that his library would equally be as varied. So, using this series field, if I click on series it should bring up OK, so it brings up 866 records and these are at the moment what I would call results and in effect what these are is bibliographic records and sometimes there will be more than one copy of a book.

So, but these are, should we say, unique bibliographic records, and you can scroll right down, and you can see these are Mike Cooleys collection of books linked by the series field. So, does that make sense?

**Researcher-3:** Yeah, and so because you're said a moment here Librarian-2 this is.

I suppose the Biblio bibliographic collection, yes, everything else that's in those boxes that came in for the donation - will they be able to get added to this catalog? Can this technology support any of that other items at the moment or is it only going to be the books at their house?

**Librarian-2:** It's a very good question Researcher-3. So suppose what I've been showing you up until now is...

So the Mike Cooley collection is I would consider within two parts. You have his personal library and that's what I've been showing. It has been easier for us to add books to the catalog to KOHA for the simple reason is that we can use MARC and this is tried and tested within libraries.

We've done all our book collections have been catalogued using MARC, so it was an easier process for us to do and we already have structures in place to manage it. Where it becomes, should we say ,where it becomes more challenging is when we move outside the realm of traditional format such as books.

… to the other formats and that was why when I was working when Researcher-A started his research looking at the various standards that are available to libraries ranging from MARC to MODS to METS to to XML Dublin Core, all this that actually that suddenly we could see some of these issues arising and it was very evident that MARC had his limitations when it came to how we were to manage these collections, but.I prefaced that on two points, the first is that it is possible to catalog the various artefacts using MARC, but I don't believe you get the richness of the metadata that you would in a way, if you use the system such as METS.

**Researcher-3:** Got yeah, thank you.

**Librarian-2**: And so, and again, we look at that. And what I will do in a minute is .... So I just want to click in and just want to show you what I term by the MARC, so just so you get an understanding you see here on the screen the codes so you can see here, for here, 020 the one I’ve highlighted. That's where you would put it in the ISBN, the International Standard Book number, which will appear often at the back of books. So it's a bar code and that's the unique identifier. So ISBN numbers will generally be under a 020 field. So with the documents I've circulated and I'm not going to go into too much detail from this perspective but one of the documents I circulated shows you goes through the codes and gives you a sense of what codes we generally use, but as you can see it contains a quite a lot of metadata, potentially within the MARC records, and you can just scroll through the various fields here so we see the various headings here and and so forth. So we have a map link which allows you to see where the book is located within the library collection, but again, that's a bit outside what we what we need to cover here. And finally, you have the items and you can see here we have two copies of that particular edition of “Architecture or Bee” and again and this is the unique reference number for each item. The barcode. So this allows us to differentiate between different copies of a particular title.

At the moment, does anyone have any kind of particular questions? Something that maybe that I've covered here that doesn't make sense that you want me to just clarify before moving on to some of our special collections and how we've catalogued those on the system.

**Researcher-2:** Yeah, I'm just wondering Librarian-2 in relation to METS. How did it appear on the horizon initially like how? How did you become aware of it?

**Librarian-2:** Well, let's we became roommates and as a result I was actually I I hadn't. I hadn't. So Mets, as far as I know, its origins are in the Library of Congress, are a bit, and it's a development on another schema called MODS. And again 3e would not have. I mean generally up until meeting Larry and yourself and working with Researcher-A back in 2017, 2018. I would not have been aware of METS it wasn't a standard we'd ever used. We had used MARC and still by and large used MARC.

But the limitations of MARC were highlighted and identified visually and should we say evidentially by the research in particular of of Researcher-A and also Eoin Dal Researcher-B as well, who followed on from Researcher-A research and as well as the key to receive the key key advantages. Sorry, the key disadvantages that Researcher-A was able to identify in terms of MARC was it's low extensibility in that the limits in terms of the type of artefacts that they could catalog in a rich way. It wasn't interoperable outside the library setting. So again, if we were using a different database for hosting the data MARC really is very rigid in how it can be used.

And it was weak for administrative and rights metadata so these are some of the core fields that make up METSs. So this idea of interoperability an extensibility were key facets, of Mets, that appealed to us and, um, I think that challenge we faced is that we've been restricted by the library management system that we've used, but equally I think as time is going on, I think we've become more increasingly aware that KOA is not the home for our archives and special collections. We haven't. We haven't got a direct home, should we say, up until, this moment in time. But I do believe long term the digital repository of Ireland's will be will become the frame for where we host our archives and digital collections.

So, does that answer your question, Researcher-2?

**Researcher-3:** Are you on mute Researcher-2?.

Under that is, she's after leaving the call after my American please images...

**Researcher-3:** She’s definitely on mute.

**Librarian-2:** Oh

**Researcher-1:** Yeah.

**Researcher-2:** I'm on mute apologies, I'm having a chat here with myself.

I thought for a minute Researcher-2. You were just there. That's it. I'm gonna tell you, know it there, he's gone up from the different tangent altogether. This is all news to me in that for sure. Hopefully not. You know, hopefully hopefully makes some sense.

**Researcher-2:** No, I'm I'm hanging in here.

Uh, no, that's great, Librarian-2. Thanks so much for that. Researcher-3 and Researcher-1 I'm just wondering, do you have any further questions on that?

**Researcher-3:** I have two if you don't mind quick ones. The first one is then is there a cost associated of moving on to the digital repository of Ireland an leaving this KOHA technology and then the second one is?

Um, anything that you've got catalog already in this special collection? How transferable is using that legacy data? You know, and converting it to Mets or what way do you think it would work? Is there an opportunity to reuse catalog items in the DRI?

Right?

That that's really my two.

**Librarian-2:** Yeah, again very good questions and the first thing is in terms of there is a cost to joining with the digital repository of Ireland and we have just signed a contract to join up to the digital repository of Ireland... I think Librarian 1 for three years is the initial contract and we are currently in the process of of we’ve signed agreements, we've been accepted to digital repository of Ireland and suppose the one challenge or the one thing that needs to be defined is the the cost needs to be paid, but the money has been budgeted for so which is wonderful? So we now have access. We will very shortly have access to digital repository of Ireland archive. So that is..

So I would still perceive the KOHA will be still be the main access point for traditional library materials such as books, journals, so forth, but I would see our archives and special collections increasingly moving towards the digital repository of Ireland and maybe even some type of intermediary platform that we might look to host between the digital repository and our own, and should we say and and the library catalog. I think that's how it may work.

In terms of the type of records that have been added already... The best thing to do is - what I want to do is, so the Bonaparte Wyse papers would probably be one of our most significant archival collections. And again, this was purchased by the college back in, I think about 2015, give or take, but this collection is again a very varied collection made up of letters, books and also some sketchbooks, art and this is just one of the items, so this is this is the Munster Catholic Association painting, so it's actually it's a painting that was included in the collection and we've actually created. We created a record for this artefact using MARC and actually all the Bonaparte Wyse papers have been catalogued using MARC. So just to kind of give just in terms of you may know who the Bonaparte Wyses were, but they were a family here in Waterford,who were a merchant family who were very influential in the history of Waterford Thomas Wyse, who would be the most well known in the families, was an MP for this part here in in, in, in in Waterford, in the southeast, and he was responsible for trying to have a Catholic University established in Ireland.

And he married Letitia Bonaparte who was Napoleon Bonaparte's niece and their family and their stories make up the archive. So it's very interesting from a Waterford perspective, and it has a lot of really historical elements to it would be very interesting to local historians and also regionally. So that is the that is the Bonaparte Wyse collection. And again, like with the professor Mike Cooley Collection there, the artefacts are linked using a series field and the 490 or 440 in this case is the Bonaparte Wise collection. So what I wanted to do is this is the Munster Catholic Association painting and what we have done in relation to that is we've created a MARC record but also we have actually digitized the item and hosted it online so that the user can actually go into the record and link out to the digital artefact itself and the resolution they can..they can scan right in or out.

And the gentleman on the second from the left is Thomas Wyse. But um, that's how we've added digitized content into our library management system on the back of using MARC, MARC standards to add the information so as you can see here, this is the MARC information regarding it. So if you're interested I can share this information with you in terms of the MARC record itself, but I think what suppose what's interesting about it is that you get a sense of how various artefacts are catalogued or the information recorded using MARC and I think it's worth looking at a system like Mets and how that is more richer in terms of the capturing of the data and really for 21st century researchers. Its standard that is much more in fitting with what how we would like to develop the collection then just being trying to use a traditional library system like MARC in order to fit our collections into and that's what we've been doing up to now, so the information that's there would probably be used so that information can be used, and it would probably be populated...

It would be, I would say it would be....

We could migrate it over to the digital repository of Ireland once we have associated digital images to go with the records and again, I think this is something in terms of the Cooley collection, which we could look at using as well. I know at the time that Researcher-A used to refer to a particular resource that was the World Digital Library, Researcher-2, that had used Mets as a means of cataloging resources, including books and artefacts, and I think that that would be something that would well worth looking at again. In addition to how? Yeah, it's the World Digital Library, the www.wdl.org and you can actually look at how do other libraries are using Mets to catalog traditional library materials and also traditional, should we say disparate items, some of the more varied material types so so that's kind of whistle. Stop tour.

**Researcher-2**: No worries when you're talking about MARC, then Librarian-2, the difficulty with using MARC is that in a way you'd be trapping yourself in a silo wouldn't you repository wise, whereas METS provides the extensibility not only for disparate items, but for interoperability as well between systems.

**Librarian-2:** Yes, yeah.

Yeah.

Absolutely, and what I think is what I find would be like what we've generally used and what I've been doing a lot of testing on recently is using a tool called MARCEDIT which is which allows you to, in effect market. It will allow you So what I’ve generally used it for would be spreadsheets. Anything that's tabular, anything that with rows and columns that would allow you to to designate a particular column to a particular, should we say code within MARC so you in theory you can actually batch load records up into your catalog, up into KOHA using MARCEDIT once you have tabulated the data into distinct rows and columns that allow you to separate between various records and various facets of information within a record. So if we were to, should we say if we were to create records for the artefact based on METS, then we could decide for METS well, what facets or what characteristics of the METS record would we like to import back into Word? And I I sorry back into to KOHA words in my mind at the moment to tell you.

I think that would be a better way to go, whereas at the moment we've cataloged items based on based on MARC standards, and then we're looking to bring them into Mets, which again is possible, but I think you're right. It's almost like you know in terms of migration where you have loads of rich data and inside and you go to migrate that data and suddenly because of the limits of the new structure you're bringing them into you have to leave so much information behind because it doesn't fit. I think that's kind of the reverse of that is true in the case of what we would be doing if we you know so I think yeah, we can leave some of the rich data out of the library system, but I think it's it would from my perspective make more sense that we use the Mets record to create the MARC Records, rather than using the MARC record to in turn create the METS record because we're losing all that richness of the metadata that METS offer.

Yeah.

**Researcher-3:** And once you've lost it, you can't get it back unless you re catalog it's gone.

**Librarian-2:** Yeah.

Yeah, absolutely and the thing with Mets is - what's really interesting is that what Mets ultimately allows you to do I believe, and again, I'm not an expert on METS.This is very much talking from a conceptual level from my perspective, but what I see the benefits of METS is that it allows you to maintain authority control. It may allow you to maintain provenance, it allows you to manage how the capturing of information within a particular record. It allows you to structure control how that data is collected, when it was collected, and how it links to other datasets, and I think that's so important then and it really isn't because in terms of archives, in terms of the long term preservation of information, ensuring that you ensuring the should we say the provenance of our record is so important and that well, it goes back into. Well, if somebody comes in in two or three years time and make some edits to that particular record, we need to be able to document that.

We need to know why that was changed, because otherwise you're left with...should we say you're left with a record that is actually, that is constant or should we say or is fluid and then you have no way to benchmark? And that is something is librarians, we've had to come to our and I don't know, I I don't know if that is something that is understood widely within the library community, but it is something that is a pressing issue that we do need to come to terms with that the traditional way we do it is not. Um is is is is not consummate, with the way that things need to be done going forward if we want to ensure the veracity and accuracy of our collections.

Because with a physical book, you're very much just describing the book. The book isn't going to change if the book, for instance, if the book gets lost, you mark it as missing. That's the end whereas if you're creating, say a record of an electronic. You’re taking a physical object, your digitising , digitize in order to converting it into another format. You need to be able to in effect benchmark that process so that you can actually compare future versions of it.

And that is, I think that's something that's that's probably is a serious challenge and and in some way the inclination is that you want to rush into the process you want to get as many items cataloged as possible, but actually in some ways taking time...you know sampling and you know working through it in a way actually whereby you find the optimum level actually can be so beneficial. Because in effect, the cataloging isn't the difficult part, it's actually coming up with what you want to capture, how you want to capture it and how that fits in.

Once you come up with those structures, then the information can be garnered from the item and as well as that's that's kind of a big challenge. So in terms one of the questions there you asked was in terms of the videos and Librarian 1, you you've been working on the videos as well in terms of the various fields. If you want to even go over, would you be able to kind of give a rundown in terms of the various facets of information that we've been capturing from videos and so forth?

**Librarian 1:** OK, I'm just gonna nip back just quickly 'cause just following on from what you said Librarian-2 which is really interesting and I think the other big thing that we want to do with the Bonepart Wyse papers rather than just scanning at a later point, we'd like to transcribe those documents, so again, something like METS that when we start going to that further level at the moment we're still in baby steps. We're kind of doing the basics, collecting the metadata and scanning, but the future would be that you can then go into a document and it would be transcribed whether it's read aloud or you can go and search words within that document so.

Right?

**Researcher-2:** Sorry to interrupt, but I saw something really interesting on the Europeana website where they do they do this thing where they enable students to transcribe letters and they run it like a project and what I was thinking maybe in the future, you know down the line maybe we could do something like that where we include the Bishop McHale School in Tuam that Michael Cooley went to and maybe get them to transcribe some items and it might be a very good way of getting them transcribed anyway, but I’ll look at that again and I might send on some links..

**Librarian-2:** It's a brilliant idea and they I've seen the Trinity did something for the 1916 letters and so in the centenary of 1916, Trinity would have done something and actually a really good website to look at is Duchas.ie which is to do with the heritage. So they digitized all these back in the 1940s in Ireland, 30s forties, they came along and they recorded all this old local history and a lot of it was like school projects and they've actually created a forum whereby people can go on. You can log on. You can actually transcribe the documents and then somebody else will verify them as a result and I think that's a great idea about the Cooley Papers in terms of transcribing and it could even be done in terms of the audio you know, even in terms of for somebody. So 'cause we have a lot of tape that Mike would have, actually, you know, should we say recorded and we also have the challenges just in terms of the videos so and Librarian 1 can chat about that in terms of how we would handle it. But in terms of videos itself one of the key issues is the formats VHS, beta tape. They're becoming obsolete so you have another challenge but from a wider perspective, you can convert those materials which Librarian 1 has been working on to, should we say to MP4 or future proof them. But the reality is, is that you can't. You can't make those resources available externally because there's Copyright issues, so Copyright becomes a key player within that, so maybe one way around this would be that you'd actually transcribe the video recording or the MP4 so you take the transcript of the information, which in effect is that you are...you can still attribute this to the documentary, but you're not actually after, say violating copyright on the basis that you've converted from, so that's just something to follow on from the transcription process.

**Librarian 1:** So, so really, just a brief synopsis on the video process. We were just kind of getting into that with our education services in WIT and they had some machinery. I've kind of thrown it out there that we had all these VHS and as Librarian-2 said, beta tapes and even some unusual formats which I think only a place in Trinity or somewhere in Dublin that's got a machine that you can actually translate them from, but we're kind of in , I suppose...I'm still in learning curve because at the moment what I'm having to do is move some firm, say a VHS, to a DVD and then once I've got a DVD I then have to move it onto MP three MP4 and and I can't with with the current software and facilities that are open to me, I guess I'm not paying for, so I have to go through that circuit, which is a bit of a slow process. I'm sure if I find the right technical people, which I think once we start to look into the digital repository of Ireland there will be systems out there that I'm not aware of and I'm sure there are so that I can go straight from a VHS to an MP3 or MP4.

And that would be the easiest way so what we're doing I'm literally just following normal standards, so their standards on every item as to what you should be collecting, such as the date it was produced, the production company, software company? There's the title authors. All of that kind of information. What language is it in? What format is it? VHS, beta or some other format? So the basic information?

An as Librarian-2 was saying earlier about the Bonapart Wyse it is catalog but once once we get to the point where we start digitising then I think we can probably add to that metadata because we're going to be learning, the more we're reading things, I'm sure there's gonna be, they probably won't be called notes feel different, so they'll be called something else. But we'll be able to add so that there will be a richer source when for interoperability, and reusable that you can search lots of other fields rather than just the standard, Title, author, date. Things like that bit futuristic at the moment 'cause? that's the plan, but still like lots to learn I guess around what we're trying what we're doing, and how we go about doing it.

**Researcher-2:** Yeah, I think you're right, Librarian 1, I I think it'll evolve as it's kind of like climbing a mountain and and all you can see is the top of the mountain. When you get to the top you can see a top of another mountain. Isn't it really that it'll evolve as you go along?

An and an even as we learn more and you know, I think we're all kind of on on this on this journey together, which is which is great.

What I'm wondering about is, say in relation to the DRI.

What constitutes a publishable data set? Like how many items would you need to have?

**Librarian-2:** Good question.

What's the minimum number of items you have to have for it to be a data set?

You know, because I've been thinking, say, if we were to describe a number of items, would we have to to in order to get it to the DRI? Would it be 10 items? Would it be thousands of items?

**Librarian 1:** As far as I'm aware, I've only I've only been playing with this recently, but it does. You can add we wondered this. Do you have to do a whole collection before you can up, but it doesn't seem to be. I think you can dripfeed if you like.

Have a certain say if I've got a certain amount of items that I know are ready and I want to get my name out there, I can start putting a certain amount up. Obviously you'd have to think about there would have to be a proper format around that you couldn't just be put in a random that have to be a whole beginning and end to that story of whatever you're putting up.

And but you don't have to put a whole collection, so as long as you've got a story behind maybe at the papers that we are publishing, say as Librarian-2 was saying earlier, we've got a number of sketchbooks in the Bonapart Wise, so you could take them as an isolated part and say we're going to put those documents up with the whole story around where they came from, ET cetera. So it's a unique part, but part of a larger collection. I guess is what I'm trying to say.

**Researcher-2:** Yeah, and and I'm wondering then leads the DRI, then probably have a structure.

**Librarian-2**: Yeah, so I think Librarian 1 you had..so I suppose just in terms of DRI. Just wanting as well as what one of the key benefits of the DRI as well is the fact that it feeds into Europeana. So Europeana is the European Union European Commission's digital repository for preserving artefacts of European importance.

So what really happens in effect is when we upload to the DRI, the DRI intern uploads to Europeana. So what it means is that any artefact are we include DRI will be searchable within the wider Europeana network, which means we can link into items should we see across the EU that may relate to the artefacts that we're cataloging? ...not catalog, but we're including, so it makes us part of the web, so you are part of a web of knowledge.

**Researcher-2:** OK, so the the DRI here then is like our national aggregator?

**Librarian-2:** Yeah, so so as things stand at the moment, METS is not a standard within the DRI but I know having spoken to Larry and we had you know we had spoken to two people at the DRI. They're extremely interested in bringing the METS as a standard so I think you know what we're coming to the party at the right time. You know, in that, I believe that we can actually we can create this discussion now and we can look at it from a conceptual basis within the wider DRI structure in terms of the DRI within the Institute, I've been designated as the, I think that the date is it. I don't know what type to give myself Librarian 1, but I like him.

**Librarian 1:** OK, cool manager.

**Librarian-2:**Yes, just King. I mean, oh sorry. I mean manager. Yes, of of our data within the.... So one of the key benefits of that is that I can approve the upload to the DRI and I have you know we in terms of how to records are created and presented so that is central so it means that we in effect take out another layer of bureaucracy so we have we have more control and it's the library that's driving it and the Cooley lab.

And that's what centrally important is that you know that we you know when we will, you know and I think for items to go on the DRI they will have to be linked to a particular collection or heritage value which the Cooley artefacts are but equally you know you have to possibility you could put up datasets in terms of spreadsheets and so forth. All of that type of data can be included in the DRI, and that's part of a wider discussion, but I suppose that's the potential we have at the moment with the DRI.

**Librarian 1:** On the how you say it now, I probably got this wrong 'cause I've got ahead of me like a sieve, but you know how we have the DOI which the digital object identifier.

DRI, I think it's BIA, but I could. I don't quote me on that. 'cause I probably got it wrong, but they also I put a unique code on the object you put up, which I think that is how it becomes more interoperable that people can then search because they're giving it this unique code with code which links to other areas. I'm not quite sure how it all works, but something around...

**Researcher-2:** Oh I wonder, is it?

I wonder is it a URI Librarian 1? Yes, it's it's like a link it it's like yeah it's it's like a link to a webpage, could say and it'll identify that that big project, yeah?

**Librarian-2:** Yes, that's exactly.

Yeah, yeah, and it's permanence.

Set up permalinks so and that's exactly it. Something similar to what you would see, so it's authority control it's back to this, benchmarking. It's back to you know, future proofing. It's all of that and what it means is that when you are disseminating, show, for instance, if we're referring to particular artifacts, or if a researcher is referring to a particular artefact within an article or publication. They will be citing that resource passed through the DRI, which in turn has permanent link, which means that for all future researchers they are coming back to the same artefact.

**Librarian 1:** So that's it. I knew you’d know Librarian-2.

**Librarian-2:** Librarian 1 has a lot of the documentation they've supplied us with a lot of this documentation already, and I suppose that Librarian 1 also has the those that there's a meeting/coffee morning coming up within the DRI within the next 2 weeks so if there's particular questions that you have in terms of the DRI we can follow up on those on your behalf and if there's particular so for instance, if like you having listened to what we've been discussing like it, you know if you want to come back in terms of saying, well, can you provide us with a copy. It is can you provide us with a copy, you know, and so forth and could you clarify this?

We can do it us, and I think that maybe the because the so my. I mean, I think we've covered so much today and yet, I suppose I’m conscious that there's an information overload, but equally as well as what comes back to is that like what, what, what you Researcher-3 and Researcher-1 are thinking in terms of your projects how we scope what we do and how we can actually give you the information and you know resources that you need to complete your task. So I mean, I think that's that's important from Librarian 1 and my perspective as well.

**Librarian 1:** Yeah, I think just quick thing if you go on a dri website - you can get to all the documents including framework documents and all sorts of information. Now I'm sure we'll get more once we get our official membership, but even at the moment there's a reef of information up there that you can read through.

**Researcher-2:** They have a lot of publications. Don’t they, Librarian 1.

**Librarian 1**: Right like yeah all the time and all. Yeah brilliant thing which they do have as well is remember the last time we were talking about GLAM and FAIR data. There are videos up there where they had meetings. They video a lot of their conferences so you can actually if you go on their site and you don't even have to be a member and you can watch all of that stuff, so there's a reef of information up there already.

**Librarian-2:** I mean, I mean, some of the central challenges digital preservation is the permanent look, you know. So you know where the artefact is going to be stored for long term preservation in terms of as technologies change, how flexible will those systems be to progressive change into ... files would become outdated, become obsolete - this future proofing, which I believe is the biggest challenge, one of the biggest challenges facing archives and libraries is future proofing. How can we ensure that work that has been undertaken now and a digital preservation that has been implemented in 2021 will be accessible in 2030, 2040, and so forth? And I and I don't know.

But I mean in the drive to digitize. I'm not sure that that pause for thought is something that I think that is essential and the DRI offers that cause you are linking into a national network where you host your data is on is on a national repository server and like they provide that expertise and knowledge to you know, to ensure that we are meeting those standards.

And I think that's probably, I think that's you know, it's their their their their key key discussions, and I think that's probably where that's really at the cutting edge. I think that rather than just doing it in isolation that we're doing it as part of a wider schema a wider structure you know, and I suppose that's that's the kind of.

You know, if you had asked me this question two years ago, I'd probably give you a different answer, but it's just as time will progress, and as I've seen, the volume of information that's coming on stream and how we link into the wider networks, I think that the key thing is harvesting the value of what you have and future proofing it for future generations, yeah.

**Researcher-2:** And I, I think here there's a huge move to opening up data sets, But yeah, that that wasn't originally there at all, and I think, really, what we're doing is we're moving from silos to the semantic web.

**Librarian-2:** Yeah yeah, yeah.

**Researcher-2:** You know, and which is fantastic, or for everyone you know the benefit that we can gain from it.

You know and I think Euorpeana as well has what they call an API that people can program from as well that they have an open API. So all that stuff is becoming...it's like there's a tsunami of data sets now coming on line, and it's what we choose to do with them to get good benefit out of them is the thing really, isn't it?

That's fantastic lads and you're you're a mine of information. I have to say and it's it's very exciting because like you, Librarian 1, when you're learning as you're going along and I'm sure Researcher-3 and Researcher-1 are probably the same, we're learning as we go along and more comming to certain points and things that we're confused about then you're you're filling in the gaps you know. So we, we, we go on another little bit, you know and so that's that's fantastic.

An so that's that's brilliant I I think. Yeah, I think we've kind of covered all the questions unless I've missed out hunting. Researcher-3 and Researcher-1 are ye okay?.

Yeah.

**Researcher-3**: Yeah, I I actually just have one little question. If we've got time, you know how you know how you were kind of saying that we're not using the Mets standard at the moment in the DRI. It's kind of new a new concept for them also, and how cataloguing the objects needs to be done the slow way so that you get it right, or you know there's trial and error to get the best results long term.

How are you approaching that process now? If the DRI doesn't have the METs standards that the model yet, and you are trying to figure it out yourselves, where’s the chicken and egg situation. How are you approaching that?

**Librarian-2:** But so I suppose, we're still very much, I believe at the at the exploratory stage in terms of this and so the I would so based on the DRI or DRI structures at the moment I would believe that we'd probably be using a former Dublin core, which is you know which is another standard for recording information, which again has very has a lot of similarities to Mets but, you know is not structured should we say it the way Mets would. One of the challenges and this is the key. I think it it all comes back to even when it comes to cataloging a book. It's the exact same. It's about deciding.

You need to come up with a structure in terms as to what level are you are you going to go to the depths of recording information? So for instance you know so when you and you and that's all down to how you - your perception of the value of an object in terms of which research value in terms of do people really need to know that information you know. You're constantly making judgments regarding what you perceive as being a value. And I suppose that that's a dip. What is value is it, is it is is subjective in that regards, so I suppose where we're at is really is deciding within the time and resource and the structures, we have and those restrictions we need to decide what is the optimum level of recording that we're looking to achieve. So, for instance, is it are we going to decide that we will, you know, we would decide on what particular fields we want so for for for a letter, for a report for a slide we decide on what are the optimal levels that we want to meet. So within METS, and having looked at it myself, you have so many various types you know you can make. There's so many different facets of information you can record, but I think there's only one of the seven categories that you need to definitely include.

So I think our challenge is to decide which of those standards are we seeing are essential to the knowledge value of the item and then deciding how you know, deciding, well saying this is what we want to collate from each of those material types similar to what we do in libraries at the moment in terms of MARC and then coming along and recording that information in a structured way so.

Things have been done I think, based on the factor we've only just joined up to the DRI. I think we've been kind of holding like we've done all the Bonaparte Wyse papers, but in terms of the Cooley collection, we haven't transcribed those sorry cataloged those to any great extent because of the fact we knew in our hearts that putting them into KOHA putting into our library management system was in effects devaluing their value because they get lost, and I think how do you harvest their value you harvest them through a platform that allows you to harvest the value of that artefact, and I think the DRI is the platform. So a kind of a roundabout way, but I think really you need to assess the value of an object. You need to decide what are the minimum metadata you wish to require from all of those particular category of artefact and then ensuring the consistency in the recording of information. And that allows you to create strength structure. What do you think Researcher-2. Does that make sense?

**Researcher-2:** Yeah it does it. It makes it makes great sense and I'm just wondering then Librarian-2 like say if you and Librarian 1 and and everyone in the library say decide on the fields that are important for a specific item.

Will that be institutional wise throughout Ireland, or will the DRI enforce that same structure on all institutions. You know how will it or will it be different for each institution?

**Librarian-2:** I think it'll be different for each institution, and I suppose that there will be shared fields that are consistent. You know a bit like with the library record you know with MARC you’ll have that shared consistency of fields like, like for instance, you will have a title, you may have an author, you have created date and so forth and so forth so they are kind of key, but for ourselves we will decide the structure of how we want to structure our records, and I think that's that's very important, and I think it's probably something in terms of having a wider discussion regarding Mets. I think it would be very beneficial if we were to come along and take because Researcher-A had done it for some of the artefacts, but even looking at it now - maybe pick out a video, pick out a slide. So maybe pick out some of the disperate items within the collection, and not focus, shall we say on that particular item per say? But within the material type of that item and what would be the types of information that we should be recording and and how it should be recorded within METS and I think once we can do that, I think some of it then you may be able to automate the process. So in terms of your Creation Day and the author like their fields. Potentially you can actually set out at the start of the process and maybe they can be hardcoded into the process so that you don’t have to continuously manage those all the time but I think that's how that's a way around that I think.

**Researcher-2:** Yeah, yeah, no, that's that's great. Librarian-2, thanks, many thanks for that and thanks Librarian 1. So I'm just conscious of time. it's.

Just going half three so we've been talking for about an hour, so I suppose unless the the girls have any further questions, I think we’ll maybe finish up there and sure we have our Cooley lab meeting now on Thursday.

And so that'll be, that'll be interesting, but thanks so much, Librarian-2 because certainly and I'm sure for Researcher-3 and Researcher-1 as well, you've certainly given us a lot of information there and filled in a lot of gaps and I suppose pointed us in other directions too which is really great.

...and thanks so much for taking the time to talk to us about this and we're absolutely delighted with that and we appreciate it very much.

OK, thank you very much.

**Researcher-1:** I have a question actually.

I hope you're still free, and it's not too late for you and you're talking about METS standards and moving to METS standards, but after you build the whole system around METS standards, is it something that you guys like the whole system with the METS standards like a catalog that you get from somewhere and insult your machines? Or how does it work? Do you build your system around this all by yourself or?Where is that freedom that you can pick any standards 'cause MARC standards come automatically with the catalog system that you used? That’s my understanding of it? But when you move to my standards, how it's going to look? Is it just something that you install into existing catalog or is it going to be a completely new system?

Uh, how it's going to work?

Do you get what I mean?

**Librarian-2:** I do, of course Researcher-1 and Researcher-2 probably has better understanding in terms of the technology behind this Researcher-1. But yeah, my understanding would be that like METS won't fit into the KOHA library management system which we operate at the moment we have to in effect migrate the data from a METS record so that it would fit in with MARC, so that is probably something if we wanted to implement incorporate the data into our library management system within ... in terms of where they should, we say the frameworks come from as far as I know, all that information is like a tool kit on the Library of Congress website which you can download a lot of data and you can in turn use that to structure your platform around that but you would be very much....I suppose you would end up having to build or I suppose at the moment, 'cause it's not widely available or use a software tool that actually is built on Mets and when you create, should we say a record based on Mets that it would automatically produce the METS into an additional file, which in turn you would have the possibility to export to another platform?

But, that's not something that we have at this moment in time. We don't have a platform that does that, but I know that it's something that Librarian-2 last year, Researcher-2 in his research had touched on in terms of is that correct in terms of to my understanding?

**Researcher-2:** Yeah, what Librarian-2 was looking at was, um, creating a METS description of video and going further with it, then in trying to create a profile and get that profile validated on the Library of Congress website. But I think time you know the time that he had for the project kind of came against him in that. But the idea is that the Library of Congress.

They put it up for viewing on their website and people can comment on it and you create a profile which describes your description. So the profile really describes what tags you're using in your description, and if that profile is validated then that can be used by any other institution to share an object of that kind with you so they look at the profile to get the list of items that you describe, and they'll create their METS description according to that and then back in WIT then we should have a system where we can ingest that into our system. You know, may probably validate against the profile and ingest it into the system, but at this point, it's all theoretical, you know.

And it's interesting too, I suppose, to go down the route with this and because it's new.

And it hasn't been used very much. METS hasn't been used very much outside the US.

Um, so whereas I suppose we can have theories and ideas and you know, uh, about things. What I'd like to see you know in particular, talking about my own region of research is to to kind of tie it down and say, right this is the idea I have.

So now I'm going to create a METS description. Now I'm going to create a PROFILE, now I'm going to send it so proof of concept I suppose really.

I don't know if that answers your question Researcher-1 does it?.

**Researcher-1:** Yeah, so the idea I'm getting or the idea or the image or impression I'm getting is that with the METS standards you're hoping to build a whole new application that will be used just for Mike Cooleys collection. And we'll be storing the data separately in METS format. And you're hoping that this application will somehow connect to the existing library system and to other say European, Europeana system or other systems out there and hopefully that that can become an automatic collection so you don't have to do any manual translation from one standard to another.

Yeah, I mean, that's yeah the only.

**Researcher-2:** Yeah, I mean that's I think the idea being that we could tie it down to its very basics and then we could see the points at which automation could occur.

**Librarian-2:** Yeah, and I think actually the only thing I would say Researcher-1, the only difference I would see in terms of what you're saying is that actually that the standard wouldn't just be exclusive to Michael Cooley archive per se so in terms of when we decide on a METS description we are in effect looking at the material type as opposed to the archive. So for instance, a letter from the Cooley collection should be open to be managed in the same way as a letter from the Bonaparte Wyses as a letter from any of our other archives and special collections. Should we say it's the material type that we're looking to manage as opposed to it being like you know in terms of like there's nothing should we say that a letter ultimately is a letter irrespective of the collection it's in and as well as that's probably what we're trying to move to and that requires us sitting down and coming up with a benchmark. A proof of concept, or a benchmark standard in terms of, well, how do we want to catalog records? ..sorry letters.

And just using letters as the example and we should be able to look at letters from our various collections and see that what we decide as being the fields we want to include that all those various types of letters should be manageable within that framework. I think that's probably what we're looking to achieve, and but again, that's not something we'll achieve in the short term, but it's something that we're working towards.

**Researcher-1:** OK, so what I understand what I'm hearing is that MARC catalog doesn't recognize such a certain items, say video recordings or letters or pictures or anything other than bibliographique marks. Understand recognizes and that's what the issue is.

Because they need to be to include certain type of description or certain values that are not included in MARC standards as they come prepackaged with your system. So say if you type something in and you say if you are adding in a new item if it's a book you have automatically options how to store that book and information about the book bit say with the letter you don't have that kind of pregenerated information that you're picking from this knowledge.

**Librarian 1:** Researcher-1 we do so - even with our system at the moment, which is, uh, a basic library system and Librarian-2 obviously is far more superior to me on this one. But we can. We can say something is a video or we can say it's a CDB, DVD or whatever the item is or a letter.

If you look on our catalog at the moment, the Bonaparte Wyse papers are all cataloged and their letters, their sketchbooks, their books. There's an assortment, so we can we have an item type, it's just and I think what we're trying to say is the MARC record has a certain criteria that you fit into, but with the Mets, you can add a lot more data if you choose to and that's the bit that we haven't got to yet. We haven't said, OK, there's this system is the world but how much of it do we want to use? I think, but I’ll let Librarian-2 take over.

**Librarian-2:** Now you nailed it. No, that's exactly it. Like MARC, allows you to catalog the various material types on MARC, but it's limited, it is rigid and you know they feel you. I don't believe you can. I don't believe you can you future proof or should we say, the record in a way that you can using METS and you lose a huge amount of metadata for archival records based on using MARC and I think it was built at a time that, but I think as the....as you talk about semantic web and in terms of developing wider networks of web, then I think MARC is limited. And while we could use this, I don't believe that.... I think I think...

I think in in a few years time I think the discussion will have moved beyond it and I think METS is definitely becoming standard and I think it's a the more the more I look at its benefits... I just see it as being... that's definitely what we should be aspiring to, you know so yeah, that would be my understanding of it anyway.

**Researcher-1:** Yeah.

**Researcher-3:** Yep.

**Librarian-2:** Yeah.

**Researcher-2:** Very good.

**Researcher-1:** OK, thank you, thank you.

**Librarian-2:** Yeah.

**Researcher-2:** OK, so folks. So last round Researcher-3 anymore questions.

**Researcher-3:** No.

**Researcher-2:** OK lads, thanks so much. That's been so so helpful and as I said before, you're so kind to give us the time and we very much appreciate it.

And I suppose our next our next performance, so to speak, is on Thursday for the Insyte- Cooley Lab meeting, and we’ll be, I suppose, putting in front of you then I suppose a picture of where we're with our research at the moment.

**Researcher-3:** And that's that presentation into your schedule. That's that one, isn't it Researcher-2?

**Researcher-2:** Yeah, the insight Cooley? Yeah yeah, that's it.

OK, and and so we look forward folks to seeing you then and you know just again thanks so much from myself and Researcher-3 and Researcher-1 for taking the time and for being so forthcoming with your knowledge.