

**Presenters:**

Annabel Lee Enriquez, Getty Conservation Institute

**Moderators and Q&A:**

David Myers, Getty Conservation Institute

Dennis Wuthrich, Farallon Geographics

Note: Q&A transcript begins on page 10

Welcome, everyone to today's webinar. Before I begin, I want to let you all know that this webinar, intro to the Arches platform part 3, data technology concepts, is being recorded. So don't worry about taking super detailed notes. I'll be going over a lot of information today. My name is Annabel Lee Enriquez, and I'm a member of the Arches project team at the Getty Conservation Institute. And I'll be presenting today.

During the Q&A session, I'll be joined by my colleague at the GCI David Myers as well as Dennis Wuthrich CEO of Farallon Geographics, the lead developers for the Arches project. If you have a question regarding anything I'm about to present or any other questions regarding Arches, please do enter your question in the Q&A panel in the Zoom platform, and we'll try to get to everyone's questions at the end.

This is a continuation of the webinar series we started in 2019 as an introduction to the Arches platform. The first two parts covered the basic capabilities of the Arches platform and the data management tools that allow you to configure Arches according to your project's requirements. And today I'm going to talk a little bit more about the data technologies that make Arches as unique as a data management platform.

Because this is an introductory webinar series, we won't be getting overly technical, and hopefully the information presented today will help you all in understanding the value that the platform can bring to your project and project data. Now if you haven't yet seen the first two parts of this series, you can access those via the Arches videos page, and that includes not only the videos, but the slide decks and transcripts of the Q&A sessions. And as I mentioned, we'll post today's webinar along with today's slide and Q&A transcript within a few days. So also I wanted to note that the Videos page includes other webinars and videos that might be helpful, especially if you are new to Arches.

Here's today's webinar outline. I'll essentially be talking about some key data management goals and how the Arches software platform can help a project to meet these goals through the

data modeling process. My aim today is to make you aware of what data modeling is and why it's important for your Arches project or really any data management project. And because data modeling is an involved process, I'll also be going over some documentation and resources that can help. In any event, the presentation portion should be about 30 to 35 minutes, and then we'll have the live Q&A session immediately following.

As a quick refresher or introduction, Arches is an open-source software platform for cultural heritage data management developed by the Getty Conservation Institute in partnership with World Monuments Fund. It was originally developed for built cultural heritage inventories, but the software has since become flexible enough for other use cases and applications such as conservation science data management, museum and art provenance data, and project and asset management.

The software itself is free, meaning that there are no licensing fees. It is enterprise-level software that is hosted and deployed independently by organizations via their own servers and/or infrastructure. And it is supported by the Arches open-source community which includes heritage conservation and technology professionals, many of which have implemented or are in the process of implementing Arches for their own projects and end users.

Since the last webinar in the series, we've released Arches version 5.1, and we've updated this graphic to include more Arches features and capabilities such as project and task management in the form of workflows. As a reminder, part 1 in the series covered the data discovery and visualization features, and part 2 discussed the data management features broadly.

Today we'll be talking about a very specific aspect of data management, data modeling. And in the diagram I've highlighted the functionality associated with data modeling, which are housed in the Arches Designer and Reference Data Manager tools. In part 2 of the web series, webinar series, we did talk about the mechanics of the Arches designer and Reference Data Manager which are administrator tools to help set up Arches for end users. And today we'll talk a little bit more about how some of the-- about some of the conceptual underpinnings of data modeling in Arches.

But first of all, what is data modeling in Arches? Essentially, it is the creation of resource models or in other words data models using Arches Designer to define your database structure, the data entry interface that your users will see, and the overall organization of information in your Arches implementation.

So this is a really powerful ability to have in that an Arches administrator has the ability to configure an Arches implementation that meets their users' needs without custom coding in most cases. And this is why the use cases for Arches have expanded beyond built heritage inventories. Once installed, you can really design a data management system for whatever you wish. And beyond that, data modeling in Arches enables so much more.

And to explore that a little bit, let's answer this question. Why was it a priority that Arches include advanced data modeling capabilities? When we started the Arches project, we realized that one of the most important aspects of any data management software is not the software itself. The most important thing is the actual data. And one of the measures of success for Arches or really any data management system is its ability to generate data that is complete, high quality, discoverable, shareable, and persistent.

And for each Arches implementation, the extent to which the data meets these goals is defined by the resource models. So having the right resource models is key for any Arches implementation. I'm going to talk a little more now about these data goals and show some examples of how the resource models enable them.

In regards to modeling for complete data, what this means is that your resource models will encompass all the data that you need to manage to fulfill your project requirements and priorities. For example, say you are trying to create an Arches implementation that includes data on people, and this table represents your data. You have four columns, one for ID of the person, name, type, description.

All this seems pretty straightforward. But when you look closer at the data, you'll see that there are some considerations to factor in when modeling. First of all, in some of the entries under name, there are two names. For the first entry, you could record Mark Twain and Samuel Clemens as one block of text in the way that it is written in that cell, but by doing so you lose the opportunity to better structure the data, which includes two different names and name types.

Another consideration here is that all of these descriptions were taken from Wikipedia. A modeling question would be is having Wikipedia as a source embedded in the description enough or do you want to be able to record the actual links to the Wikipedia entries for these people so that users have access to additional data that you have chosen not to include in your system? Also too, it might not be feasible to include all of the information in that Wikipedia entry in your system so you would want to have some links to the data-- to that data.

Yet another consideration is whether you want to model dates for birth and death of each person. For your use case, you would need to decide, again, if this information needs to be extracted from the description to better-- to create a better and more explicit structure for the data so that it can be more easily interpreted by machines and then used by you and perhaps other systems. So these are some considerations you might have when creating a resource model in Arches for this data.

And once you have made these kinds of decisions or even perhaps as you are making them, you as an administrator would use the Arches Designer to create and encode the settings for your research model for person. I've gone ahead and created a very simple resource model for person based on the data in the table, and you can see the structure in the red box.

To inspect this model further, on the left I include a zoomed in screenshot of the model structure from Arches Designer. On the right is a graphical representation of this model. So on the left, you'll see that there are nodes which is what we call fields for type; name and name type; identifier and identifier type; description, description type, and description source.

And here's the model with the corresponding data from the table in red. Now the data in blue is data that was added in order to describe the name type, identifier type, and description type. I could have added additional nodes to accommodate the data regarding lifespan if that was important to this project, but I decided not to. And these are the kinds of decisions that can be made when data modeling for your own data.

Now because different projects have different requirements, this means that the resource models can be very different from project to project. You'll see the simple person resource model that we just discussed on the left. And on the right is a more sophisticated resource model for person where the modeler attempted to include every possible kind of data that one might want to collect for a person. The point here is that you can tailor your Arches resource models and hence your Arches implementation to your unique project needs and priorities.

While resource models from project to project can vary, sometimes there are reasons why they're very similar. So for example, often a project needs to conform to an established data content standard which defines what data you need to manage. So if you are in England, the MIDAS Heritage standards define everything you need to collect for a heritage inventory. And in that case, you would create resource models that include everything specified as required by that standard.

Some other examples of data content standards for heritage inventories include the International Core Data Standard for Archaeological and Architectural Heritage, or in the US, State Historic Preservation Office or SHPO requirements for each state. Alternatively, there are other data standards that are not specific to heritage or heritage inventories such as Dublin Core that you might want to model your data according to.

The implication here is that every project that adheres to specific content standard will be able to use the same resource models with perhaps some minimal or superficial changes. I'm going to talk a little bit more about this later in the presentation, but definitely this is a very key thing to note.

Related to having complete data, having high quality data is also important. So now you're using resource models to define the parameters of the data that you are managing. For example, you know that you're going to collect the name of a person, but are you going to allow one name to be collected or allow multiple names for that person to be housed in your Arches system? What would be the data type for that name entry? And are you going to restrict data to correspond to a controlled vocabulary or in other words, a list of approved

terminology? These are all very important questions that relate to the quality of the data that you will be managing.

For an example, let's look at data types. Here are some examples of data types that you can use and assign to the nodes or fields for your Arches resource model. If you work in data and have experience in that some of these might be very familiar to you and some are more of a stock data types that were developed for Arches for specific use cases.

To continue our discussion about how high quality data can be ensured by selecting the right data type, let's talk about this data type or the EDTF or extended date time format data type that helps to capture fuzzy dates. So what are fuzzy dates? Fuzzy dates are essentially incomplete or approximate dates. If you look at the table to the left, say you have a data set where you've got to fill that defines built date for monuments or just a regular date. But in that field the formats vary wildly, and all of the dates are fuzzy except for the last one.

Oftentimes these are the kinds of dates you might find via historical research, and they're important as a historical record. Alternatively, you may have fuzzy dates because the date is too far in the past to be considered exact or accurate. Now in the past, because you may have wanted-- because you want this data to be machine readable as a date in this one field, you may have decided in the past to standardize the date format.

And in this example it's been standardized as just the four digit year. Obviously in doing this, you lose the nuance and important information held in their original data such as the one exactly date you actually had, or that one date might be June 1753. Also the BCE date can't be represented using the four-digit-year time format.

The Library of Congress in the US created a data format standard to address this called the extended date time format specification. And in Arches we have a data type that allows you to manage this kind of data. And you'll see what the EDTF entries would look like in the third table. There is a format for that prehistoric date and for all the other dates that represents exactly what you see in the first table.

Now to add this data type to a field or node in your resource model in Arches, you would specify this in the Arches Designer. Here we're going from this regular date format with-- it's a year date format to the EDTF date format or data type which then gives you the opportunity to create some parameters such as padding for how Arches encodes and interprets this data. And once the actual data is entered using the EDTF data type, Arches will know how to interpret the data, and you can search for these dates in features such as the Arches time wheel.

So using the Arches version 5 demo site search interface, I'm going to search for the tomb of Nefertari also known as QV 66. And looking at this record, I'm going to scroll down here, I see that the EDTF date that was entered is approximately 1279 BCE to approximately 1212 BCE. Now that's a little bit hard to read, so I have it in the slide. So the date was approximately 1279

BCE to approximately 1212 BCE and it was expressed using that notation, the EDTF specification.

Now the beauty of using this type of date and the data type is that Arches does two things here automatically. Number one, it accounts for all of the dates between 1279 BCE and 1212 BC, so that red line there on the timeline. And two, it also accounts for the uncertainty based on the parameters you've set. So you might not have seen it when we were looking at those settings, but the padding for the uncertainty for the fuzziness was up to one year. So how Arches encodes the approximate date then is 1280 BCE, approximately-- so it's 1280 BCE to 1211 BCE.

So if we were to search for tombs or just to run a search for tombs built between 1290 BCE to 1280 BCE, the tomb of Nefertari would still be returned because Arches can account for the fuzziness or uncertainty of the date based on the parameters we've set. Let's run that search very quickly and see the result.

So let's use the time wheel to search for resources that have dates between 1290 and 1280 BCE. Let's click on that, and you see I have 17 results. And I don't see the QV 66 yet but I will. You see that it has been returned because Arches can account for the fuzziness of the date. And so because the data was of a higher quality and not just alone four-digit year, Arches was able to better interpret it.

Moving on to modeling so that your data is discoverable or searchable, we just looked at it one way to optimize the discoverability of your data. But in addition to that example, typically you can design resource models that optimize search by normalizing the data structure to reduce redundancy and improve data integrity. So for example, what I mean by this is that you would not repeat the same data in different nodes or fields or be inconsistent in how you design your resource models.

And then once you've done that, you would want to employ controlled vocabulary to standardize the terminology. I also mentioned controlled vocabulary when discussing data quality because they also inform uniform data entry. But let's look at controlled vocabulary now more in how they relate to search.

Here's the Reference Data Manager which manages controlled vocabularies in Arches. We're looking here at the record for the concept of residential single family house and you'll see that it has several different labels, one's for house and different languages as well as synonyms for house. So that you can search using any of these terms and return the same resources. So I'm going to go to search interface and check this out.

So let's first search for house. And I'm going to go ahead and use the concept search. And I have four returns on that, four resources that have been categorized as house using that concept. Now let's type in house or the German for house, same for resources, and then residents. All right, so same for resources.

Let's say you neglected to add a term such as casa or house in Spanish. You see it's not here anywhere. It's just represented as a string. So let's add it. Let's add that as a label to house in the Reference Data Manager. I'm going to search for that record, for that concept, so search for house. And so casa is not there, so I'll add it here. And it's the-- the pref label for Spanish.

So we just added that label on the RDM, and now we don't have to go into every single record to make that change. So now when I search for casa, let me do that. Let me type in casa and so like concept because it's running in concept search and we still return the same for our resources. So this is a very quick example of being able to easily increase the discoverability of the data in Arches by adding a new label in the Reference Data Manager.

Moving on, in regards to shareable and persistent data, resource models enable data that is interoperable and help to ensure data longevity by encoding semantic metadata into the data so that the data is self describing and portable. Now, as a note, because Arches is not proprietary software and as open-source, we use nonproprietary data formats such as CSV or JSON-LD to export the data out. And it exports out with that metadata.

And typically you would want the semantic metadata to correspond to an established standard that is regularly maintained and widely adopted such as the CIDOC-CRM or conceptual reference model, which is maintained by the International Council of Museums Committee for documentation. So basically the goal here is to make sure that the data is encoded with metadata that describes what it is and that that metadata corresponds to a known standard.

Here, again, is that graphical representation of our simple resource model per person. And here is the branch in that resource model which describes the name of a person. Now in other words, you can read this branch as a person is identified by a name which has a type where the data is being collected for name and name type.

Now here's that same branch with semantic metadata corresponding to the CIDOC-CRM. Now the statement is more explicit about the nature of the data and more machine readable. And why is this important? It may seem clear that name is an accolation in this case, but what if the label for that node was something else that was not as explicit? For example, handle, what if it said handle instead of name?

In English sometimes people use the word handle instead of name particularly in the tech world. Now, if you are managing an Arches system in which you've used handle instead of name, then without the semantic metadata describing it as an accolation that a person is identified by, it might not be clear to others who are using your data what the data under handle actually is.

Because of the semantic metadata, it doesn't matter what you call it in your system, and ultimately what matters is the data. Here are examples of actual data that is being described by the semantic metadata. So for this person, Mark Twain is the pen name and Samuel Clemens is

his birth name. And upon export, this data will include metadata that lets you know that Mark Twain is an accolation for this person and that that accolation has a type of pen name. So in that way the data is portable and self describing and can be more easily ingested by other systems that use the same standard.

So let's come in to the Arches Designer to see how we can embed the semantic metadata. Person or an E21, according to the CIDOC-CRM, or in this case, it's the link of Arches version of the CIDOC-CRM. It's probably a good time to note that Arches comes with the CIDOC-CRM loaded, but you can load other ontologies if that better fits your use case.

So we're going to add that name branch to this new person model. So first we'll add name. And let's not start with property. Let's start with the ontology class, which is as we saw E41 accolation. And then that restricts the properties that we can select and to the ones that are defined by the CIDOC-CRM.

And so now we have E21 person is identified by E41 accolation. So let's add now name type to name. That would be an E55 type, and P2 has type. So now this reads E41 accolation has type or P2 has type name type which is an E55 type. So I'm going to save that. And so now all of that semantic metadata is encoded in the resource model.

Let's go ahead and change that to handle. You'll see that it doesn't change the semantic metadata that gets associated with these nodes. So that regardless this data will be-- the data will be defined as an accolation with a name type or with a type as defined by the CIDOC-CRM. I'll also note here that the complexity of the CIDOC-CRM is hidden from your end users. So they will be entering data in a data entry form and will have no idea that they're producing-- that they are creating semantically rich data.

Now I've given you a lot of information about resource models in Arches, and here's what we've learned. Resource models are sophisticated data models that you can create without coding using the Arches Designer. They define the data structure and overall organization of information in your Arches implementation. And as you are going through this process, they also generate the data entry forms for your end users.

Also Arches resource models enable data that is complete, high quality, discoverable, shareable, and consistent so that once you are done with the modeling process, you'll have a data management platform that is tailored to your own use case and end users. And at this point I'll also note here that unless you are constantly changing the nature of your Arches implementation, you typically only go through the modeling process once in the beginning of your project.

But still given everything that data modeling determines in Arches or really any data management system, there's a lot of responsibility on the part of the administrators who are setting up an Arches implementation initially. It requires a lot of time and consideration. And



two of the things that I think are essential here are number one, and in-depth knowledge of your own data and use case. So you would definitely need that to start modeling, and then as well as some experience or training in knowledge organization, and not just knowledge organization but knowledge organization within Arches.

Now we can't help you with that first thing, but we've been working on some documentation and resources to help you with data modeling in Arches. And when I say we, I'm referring to the Arches Resource Model Working Group or ARM-WG, which began with the aim of facilitating the sharing of CIDOC-CRM based modeling patterns amongst different Arches implementations.

The Arches Designer was built to enable modeling without coding, and the ARM-WG was founded so that modeling could be made easier for organizations that wanted to implement Arches but had no background in knowledge organization. So to make things easier, the working group had the following goals. Number one, create a package library that includes resource models and branches that could be shared and used by different organizations. Number two, create documentation for data modeling with our suggested methodology for Arches resource models. And three, facilitate modeling discussion within the Arches community.

Now the first of these goals is the Arches package library. A package in Arches is a configuration bundle that can be loaded into an Arches installation, and it includes resource models, branches, vocabularies, ontologies, and sometimes actual sample data along with other configuration settings. So then the library would be composed of links to and information about submitted packages that can be used and adopted by anyone in the Arches open-source community.

So if you are feeling a little bit overwhelmed by the information on modeling that I just gave you, just know that if there are use cases similar to another Arches implementation, you might be able to use the same resource models with some slight modifications. And here's the beginnings of that package library. All of the web resources that I'm showing you today are works in progress. So while we had made these pages live, we are still adding more documentation and more information.

Currently there are three packages that are available to be loaded into an Arches installation in this library. And one of these package is the one that is currently being used for the Arches version 5 demo site. And you'll see here that we've included information on what resource models are included in this package.

Now in any event, we hope to add more packages in the future, including ones based on historic England's Arches resource models which align with the MIDAS Heritage standard as well as a package for Arches for science, which will include resource models and workflows for managing conservation science data such as object image annotations using IIIF and instrumentation or measurement data.

Now the documentation on Arches modeling includes foundational information on structured data and data modeling for those who have no background in these topics, suggested Arches modeling methodology, and additional modeling resources for more information. Again, this work is still in progress, but we've organized the information in the following categories; basic concepts, methodology, which is coming soon, benefits of modeling data according to these principles, glossary, and additional links with more information.

We definitely invite comments or contributions from anyone to help add to and enhance this documentation. In fact we're inviting anyone who is interested to become involved in the activities of the working group as we're looking to increase discussion around resource models in Arches amongst people who are looking to implement Arches or those who are interested in data modeling in general. Also we we'd love to expand the package library and include more packages that represent more varied use cases. And as I just mentioned, any help in enhancing the documentation around modeling would be very, very welcome.

So again, that was a lot of information, but in a few days to a week we'll be posting the recording, slide deck, and transcripts. Coming next in the intro to the Arches platform series is part 4 on Arches information architecture. We don't have a date for that one yet, but be sure to follow us on social media or sign up for the Arches mailing list to be notified of a date for that webinar. And that concludes the presentation portion of the webinar. So I'm going to ask David and Dennis to unmute themselves and join me for the Q&A.

Great. Thank you, Annabel.

That was great.

Yeah, excellent. This is David. I'm going to be moderating the Q&A. So thank you, Annabel. And great to have everyone who's joined from different parts of the world. And for those of you who-- you can-- as Annabel mentioned, you can submit a question through the Q&A panel and also see other attendees questions there. And one other thing you can do, if you see someone else's question that you like, you can click on the thumbs up icon at the bottom of the question to bump it up in the list in terms of answering it. So we'll try to get to everyone's questions.

And I see one already submitted, which is, this is a general question, not necessarily specific to data modeling. Could you please elaborate on how Arches has so far been used in the UNESCO historic urban landscape undertakings? Which cities or organizations have used Arches and what challenges you at Getty's Arches project, I think, have heard of or observed the challenges encountered?

I'll take a first stab at this question and others can jump in. I'm not aware of specifically implementations that are focused on the UNESCO historic urban landscape approach or methodology. But to the second part of your question, there are a number of cities that have implemented Arches or in the process of implementing, including city of Los Angeles in the US,

city of Lincoln in the UK, greater London is in the process of implementing true historic one, and city of Philadelphia is in the process, and also city of San Francisco.

Challenges that we've heard. I'm not aware of challenges specific to cities. I think that's actually been a geographic scale that there's been a good deal of uptake with Arches. And from what I understand, cities like Los Angeles and Lincoln have been quite pleased with what they've been able to accomplish with Arches. Do you have anything to add, Dennis or Annabel?

No, I think that covers it pretty nicely, David.

Yeah, I don't have anything to add to what you mentioned, David.

All right, the second question, does Arches support open research data? If not, are you planning to do something in this respect?

So in regards to this question, So, Nevin, if you could elaborate on what you mean by support open research data. One thing I want to note here is that when you install Arches, it's pretty much up to you to how you want to model your data. And so if you're talking about supporting a framework that corresponds to open research data, then that's something that anyone in the community can do, can implement, but in regards to any specific implementations that we are aware of that support open research data, I'm not exactly sure in which-- in what way you're referring to that. So nevertheless, feel free to elaborate in the Q&A panel if you like. And Dennis if you-- or David if any of you have anything to add to that, that would be great.

I think I mean I'm not familiar with a specific open research data framework if that's what Nevin is referring to, but I think you're spot on, Annabel, like part of the data modeling effort in Arches is to think about how you want to build integration points with other systems.

Yeah, OK, so we're not going to mark that answered yet, so if you want to elaborate, Nevin, feel free to do so.

All right, in the meantime, we do have another question. "Thanks a lot for an amazing presentation. For those who are familiar with ArcGIS or QGIS but want to start using Arches, would you have some suggestions? I mean would that familiarity help or do they operate with different logics?"

I can take that on, David, if you like.

Sure.

So this is a pretty common question. And it really I think boils down to Arches and how Arches can work in tandem with a traditional GIS platform like ArcGIS or QGIS and is there some

synergy between the two systems? And I think the quick answer is Arches I think works really well and is a good neighbor with a traditional GIS system.

So what does that mean? It means that you can import and export data from Arches into a GIS system with minimal effort. Maybe more interestingly we're working on a project right now where we're supporting really real time direct integration of Arches data inside of ESRI's ArcGIS pro product. And what do I mean by that, I mean that we've developed an add in or we're developing an add in, just about to finish it actually, that allows you to access Arches data as native as refeatured services within ArcGIS pro, make edits to geometry and upload that geometry back into Arches and then directly within the ESRI software access Arches data entry forms.

So it's quite, I think quite an intimate relationship between the two systems. And we have similar plans for QGIS integration as well. So that's one way to look at it, like there's a really good-- there's a really good working relationship between Arches software and traditional GIS software.

The second half of this question, well, I think is really interesting, which is does knowledge of familiarity of those two platforms help with Arches? And I'd say that in general it does because you'd be approaching your modeling effort from the perspective of thinking about the geospatial relationships between things.

But there's a really big difference between how Arches models data and how ESRI and QGIS both model data. And we can have a whole webinar about this, but I'll just maybe cut to the chase and say that traditional GIS software systems really have some quite significant constraints on how they manage data.

A traditional GIS person would think about defining or modeling information as a feature class, that is to say a table with prescribed attributes and a specific geometry type, so a point or a line, or polygon. And Arches does not require you to make the same constrained-- constraints in your data model. Arches will allow you to create models that are less plainer, less linear, sorry, less like layers and more like objects. And your models can support all the geometry primitives.

So you can have a model for historic monuments or historic buildings, say, and there's no need to define one for a point feature, a line feature, or polygonal feature. So there's a lot of additional things that you can do in the Arches modeling world that free up some of the ways you manage data when compared to a traditional GIS.

And to add to that, Dennis, so from like if you are used to working with data in a traditional GIS, and one thing that I will add here is that one of the ways in which-- one of the ways in which we developed Arches was so that people who didn't have experience in these traditional GIS software platforms so that those people could actually use the system and they could create

geospatial data without understanding how to use ArcGIS or QGIS or even knowing anything about GISes. So for end users creating geospatial data, working with geospatial data is I think relatively easier using Arches.

All right.

Hopefully that answers your question. Let us know if it didn't and you can add another question in the field.

I'm going to move now to-- Nevin sent some clarification of his question about, does Arches support open research data. He added what I mean is when we define a digital object, some other repository can take it.

OK, yeah. So I think this is-- at least the way I understand this is in the way that we think about linked open data or just linked data in that in Arches you can model your data so that it be-- so that it is portable and-- so self describing so that it can be ingested by another system, whether that's through just an export process or through some sort of in synchronous way using an API that as long as the two systems, the receiving system and the publishing system, are using relatively same standards, the same standards. Dennis, do you have anything to add to that?

Yeah, I would just add that this-- to me this question hints that data interoperability. I think the key point here to make is that if you're diligent about how you model your information in particular thinking through the way that you build semantics into your models, you can do that in a way that is consistent with linked open data.

And Arches supports import and export of all this information as linked open data and specifically as JSON-LD, so JSON representation of linked open data. And that data should be ingestible by other systems that know how to process linked open data. So that I think is one of the key reasons why-- it goes back to what Annabel was saying in the presentation that key reason in thinking through and putting the effort into thinking about how you build your models is to support these kinds of data interoperability capabilities.

Semantics are necessary to build linked open data models, but they're not necessarily sufficient. You do have to think about the structure or the patterns that you use in structuring your models. And if you do that properly, you can build models that support valid linked open data and that means that any system that can process linked open data can take Arches information either as an export or via the Arches API and interact with it.

All right, thank you. Here we have a question. Are there users who are using Arches to document damaged due to historic monuments during earthquakes?

Well, I'll jump in and say that I'm not aware of anyone who's thinking about it in specifically as earthquakes, but I do know that there are several Arches projects, Arches implementations

explicitly concerned about managing and tracking the condition of cultural heritage objects and their risks and their status so that in the event of an earthquake they'd be able to quickly look at the inventory of significant heritage objects and quickly-- and use Arches to update the condition of those objects by doing a field investigation.

And we didn't explicitly mention this, but one way to-- one way Arches supports that is through the use of Arches Collector which is a-- it's an application that you can load on a mobile phone, either in iOS or Android phone, and that has the ability to log on in a secure way to your Arches implementation and access your Arches data.

I'll just add that there is an Arches implementation focused on Nepal that was really implemented in the aftermath of the earthquake there a few years ago. And I believe it includes information on damages to structures there from the earthquake which is called the-- it's called downhome the-- it's through the Nepal heritage documentation project.

Also both the cities of Los Angeles and city of San Francisco, I know they are, of course, in earthquake zones and I know that certainly in Los Angeles, which I'm more familiar with that implementation, I know they're certainly concerned about being prepared for the next earthquake and having information in Arches is already on the significance of different structures.

Another implementation that it's not earthquake damage but following on what Dennis said about tracking conditions. The Florida public archeology network through their heritage monitoring scouts program, they're tracking conditions of archaeological sites across Florida, including looking specifically at impacts of sea level rise there.

And I'm just going to chime in here and let you all know that David's been mentioning some implementations of Arches, and you can find links to some of these implementations. Some of them are not live, but there are others that you might want to check out on the Arches website at [archesproject.org](http://archesproject.org) and you can look for the implementations of Arches that have some examples of projects that are using Arches.

All right, another question here from Celine saying, Annabel, you have invited the audience to connect or collaborate with Arches project, could you elaborate on the ways to further the collaboration. For example, could individual heritage researchers or practitioners have defined and sustained relationship with Arches or specific research project or Arches application to an urban area?

OK, well, that one is directed towards me so I'll take it. So just to clarify the ways in which I was referring to collaborate was in regards to modeling, but in general if you are working on an Arches-- so first of all, Arches is an open-source project and it's open-source software and we're not involved directly, so meaning the Arches project team at the Getty Conservation Institute is not involved directly in every single Arches implementation or Arches powered project, but we

invite collaboration within the Arches open-source community on several different things, and for this case, we're talking about modeling.

If you want to collaborate with us to talk about how to model for your particular use case and share the results of that with the rest of the community, that's one way that you can collaborate within the community. Other ways-- so for example, right now we have a documentation effort that is actually entirely-- pretty much entirely powered by the community. It's community led, and so we're not necessarily directly involved in that effort. And it's an effort to improve all of the documentation for the Arches platform. Those are some examples.

Now in regards to-- for your-- noted an example, could individual heritage researchers or practitioners have a sustained relationship with Arches for specific research project or Arches application to an urban area? In that case, if you wanted to let people know in the community what you're doing and collaborate with others in the community, especially if you have-- especially in regards to your application to an urban area, that would be very helpful to others who might want to do similar things using Arches. So those are some examples. Again, this is really about interacting, being part of the Arches community to help improve the community and to help others in the community to help them with their Arches implementations.

Can I chime in, Annabel?

Yes, absolutely.

Because I'd like to-- I mean to me the answer to this question is a pretty unequivocal yes that you can choose to collaborate really on multiple levels. I mean Annabel gave some really good examples. But really-- and I think the key point here is Arches, we want-- you should look at Arches as much as a community as a software project, which means that you can work with Arches on multiple levels.

You certainly can just be a user of Arches, but you can increase your level of collaboration by engaging with the community on multiple levels. One level is through documentation like Annabel just suggested. Jumping in and helping with improving or extending the documentation is a form of collaboration.

But you could just as easily take advantage of the Arches discussion forum. So there's an Arches group for Arches users and there's an Arches group specifically for Arches developers. Both of those are avenues through which you can ask questions, but also spearhead initiatives if you choose to.

And collaboration in that sense could include creating or extending code if you're a developer or generating discussions about modeling questions or approaches or research ideas if you're

more of an Arches user. So to me really the point is if you're interested in collaborating, there's a way for you to do so.

One additional thing I'll add, in the UK there's been a self-organized UK regional user group that I believe they have had a couple of meetings and they're planning a third one this spring. But that's something that different organizations and individuals out there in the community decided, hey, there's a critical mass in that region. So they've been organizing and sharing experiences and I think trying to help each other out. So that's another possibility.

Yeah.

All right, to the next question. Is there an Arches app for mobile devices?

The short answer is yes, there is an app. I should be really clear here. There's an app for mobile devices called Arches Collector that allows you to log on to an Arches project and access and create data in the field. And it's-- there are native applications for iOS and for Android that allow you to do disconnected editing and syncing the data that you've collected in the field with an Arches project. It's secure. You have to have an Arches account for this to work and you have to log on to a specific Arches application. But the short answer is, there is absolutely a mobile app that allows you to access and update Arches data.

And just to add to that, there is-- so that app is it's a free app that you can download for your iOS or your Android device. There are no separate licenses. So you can have as many people with the app and you can define projects within Arches data collection projects that you push to that app for your users. So if you have things like crowdsourcing project or just survey projects, you can do that through the mobile app.

All right, let's push through. We have a few last questions and I know we're running to the end of time. Just one other comment really saying that attendees in the past found Arches Designer challenging to get to know, but noted that the new resource modeling documentation makes it much easier to understand. So I'm just noting that that seems to be a very useful resource. Next question, Annabel, did you want to say anything there?

Yeah, I just want to say thanks. I'm typing that as well so that's fine.

Next question. In the Arches Designer you use dropdown, a drop menu where you added casa, and this drop menu for a short while I saw different languages, including Cyrillic based proposals like Russian, where can I see the list of these predefined languages and are they available for all fields in Arches Designer?

So I think-- and, Dennis, I'm going to give it to you in a second. I think what your question is alluding to is internationalization or localization with languages. But just to answer your specific question regarding the Reference Data Manager, those languages that were in that dropdown, I



define those labels for those languages in the Django interface Arches, and so that those specific languages were available. I could have added more languages, but I didn't. Dennis, I'm going to kick this question over to you. So you might want to talk about actually how Arches has dealt with languages currently.

Yeah, so this is a good question, and it hints at a much broader-- really a much broader topic which we would call internationalization of Arches. So what Annabel showed in the presentation today was the ability to support multiple labels for a concept and that those labels can be managed in multiple languages. And in fact those languages can have multiple scripts. So the scripts can include Latin based scripts but also Cyrillic and other scripts, so Arabic, or Chinese, or Japanese.

So Arches code base is completely capable of supporting various languages. And right now currently as of version 5.1, the part of Arches that supports multilingualism is the Reference Data Manager, so the part where you control what appears in the data entry form dropdown.

We're just starting to think about extending the internationalization of Arches, the ability of Arches, the user interface, so data entry forms and all the on screen labels that you would use to interact with the software to support multiple languages as well.

That currently isn't in place in Arches, but we are beginning the effort to at least explore how to support that. And while nothing has been promised, I am pretty optimistic that we'll be working on at least some support for alternate languages in the Arches user interface. And that will probably be coming along in the next year or so, sometime in 2021.

All right, next question. Sticking with Annabel's personal example from the presentation, how would the data look when exported into CSV format, especially in terms of the multiple name types or multiple descriptions?

OK, so this might get a little bit technical especially since I have nothing to illustrate this. So when there are multiple name types for a particular resource, so that example for Mark Twain, usually what will happen in the CSV format is that you'll have two lines then. The first line will have the data for Samuel Clemens and then there will be a second line that has the same resource ID as the primary key. And then that will have the additional information. So one line with Samuel Clemens and first name and the other in line with Mark Twain and pen name. So definitely we can go into that more, but this is-- probably during another webinar we'll talk a little bit more about data import and export, but that's what that is in a nutshell. You basically have multiple lines for the different entries.

And, David, I'll just answer the second question for Christina. Can videos and our 3D models be stored and displayed in Arches? Number one, yes, they can be stored. How they're displayed, the videos or 3D models, depends on what exactly you mean. We have several Arches

implementations that have embedded video and 3D model viewers within the report templates. So that's an example of how they can be displayed.

But generally Arches, if you are-- you can upload most types of file formats into Arches. And the question is whether or not-- or whether or not and how it is displayed within Arches, and that can-- it might entail some customization depending on what you're talking about. Dennis, did you want to add to that?

Yeah, I think you're correct. There's no issue in uploading digital data, digital files. And some digital files can be displayed natively, so out of the box in Arches. Some digital files are specific enough in nature that you need a particular bit of code to be able to actually display those file formats interactively. Arches allows you to embed that code if necessary.

So I'll just say that we support the display of a small set of 3D file formats natively. And if you are uploading data in those 3D file formats, then that's not a problem. If you're uploading data, 3D data or video data in ultranet formats, formats that aren't natively supported by Arches, it's just a matter of including the decoders for those file formats and embedding that in your Arches report formats.

It's typically a very straightforward thing to do. We've seen developers include the appropriate libraries when necessary with just a little bit of work. So maybe that's a bit more detail than you were looking for, but yeah, the short answer is yes, Arches does let you do this.

OK, great. Hopefully that answers your question, Christina. And since we are over time, I think I'm going to end the webinar now. Thank you all for attending the webinar and making it this far. And you can share your feedback about this webinar at [contact@archesproject.org](mailto:contact@archesproject.org). And we had mentioned our Arches forum. And that Google here's the link there, and then the webinar repository, the [archesproject.org](http://archesproject.org) videos. Again, thank you all for attending. And I hope you have a great rest of your day.

Thanks, everyone.

Thanks, all.

-----  
Users do not have to request permission from the J. Paul Getty Trust and/or the copyright owner prior to using any of the materials but must comply with the terms of the Creative Commons license noted at the bottom of each didactic material. Please note that teaching materials created by different instructors within the same course may carry different Creative Commons licenses.

Materials that indicate "CC BY-NC-SA" **may be adapted** and you may use them to create derivative works non-commercially, if you credit the Getty Conservation Institute and the original instructor or author of the material and license your new creations under the identical terms. In this case, please use the following citation format:

Based on original teaching material created by [name of instructor/author, if he/she is identified] for the course [course title] presented by the Getty Conservation Institute.

License: [Creative Commons BY-NC-SA](#)

©2019 J. Paul Getty Trust



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).