Arches For Science CoSTAR Presentation

The topic today is data integration for conservation science. And I will play something for you. And then you know what the talk is. Here.
Anupam, that certainly woke us all up. I think, we're now ready for Catherine.
It's a time to disappear Catherine, so the floor is yours.
All right. Hopefully, I'm sharing my screen now. Does everyone see that?
Yes,
Yes.
OK, great. So as Anupam said, my name is Catherine Patterson and I'm a Conservation Scientist at the Getty Conservation Institute. And I'll be talking to you today about, yes, a project we call DISCO at the data or at the GCI to help us dance better with our data. But for your purposes today, we'll be referring to this as Arches for Science. And I'll explain that a little bit later on.
I also have, with me, on the call today, my colleague, Dennis Wuthrich. And he comes at this from the technology side. He's the person who actually helps build the software, so that we're going to talk about today. Which, itself, should provide you an overview of what's so important.
And before we go any further than this, I also would just want to pause for a half second and say, with everything that is going on in India right now, between COVID and cyclones and everything else, we really do appreciate the opportunity to tell you a little bit about what we've been doing to try and extend the Arch platform that you heard about last week into something that works for folks like me, that is practicing Conservation Scientists.
Hopefully, the slide advances here. There we go. So I know that you've been learning quite a lot or hearing quite a lot about what conservation science can look like in a museum-type of environment, which is the environment that I work in.
But just to make sure we're all on the same page, I kind of want to come at this from a sort of high level view. And we'll, progressively, dive in and get a little more specific, and get a little more detail about what a technical data integration system for conservation science might look like.
So just to kind of recap the process of technical analysis, and I will say, since I work in a museum environment, very often, I will use conservation science and technical analysis interchangeably. And what I mean by that is the process by which I'm typically asked to look at one or a body of objects, often in my case, movable [AUDIO OUT] objects. Although it doesn't require that they be movable.
And such as the grouping of images that you see on the screen here. This happens to be a collection of illuminated manuscripts or painted books, and panel paintings that are all produced in a single, 14th Century workshop in Florence, in Italy.
And as a Conservation Scientists, we are asked to take objects that look like this, and answer questions about them through some sort of scientific observation. The types of questions that we may be asked to address are, as I'm sure, you've heard over the course of your workshops, incredibly broad.
We can be asked to talk, for example, about what conservation treatments might be appropriate for an object or a set of objects. Can we be more specific about attributing particular works of art to particular artists or artist's workshop. We may be asked to talk about date or about to understand the change of an object over time. There's a huge breadth of questions we may be asked to answer.

Just for a little bit of context, in this particular instance, the question that we were addressing was one about workshop practice. So all of the objects that I'm showing you on the screen were produced by the same Florentine workshop, which might seem somewhat strange. Why would one workshop produce both books and panel paintings? And that was a curator's question-- is what could we learn by studying the materials and these about how one workshop would have worked in the two different modes-- books and panel.

as I said, to do any kind of study like that, as a scientist, I come at that from the point of view of collecting data. So here, you see three images-- three action shots of data collection on some of those objects that you saw on the previous slide. What I want to point out here is that the way we collect that data, just like the questions we try and answer, is incredibly broad, incredibly variable.

So for example, what you may be able to work in situations like you see in the left image, where I'm in a very controlled environment, using a benchtop, in this case, Raman spectrometer in the GCI laboratories, where I have complete control over everything that's going on.

If we look at the far right picture, sometimes we have a very different situation, where I'm in a controlled environment, say, the gallery at the museum. But being asked to, very briefly, in a very small amount of time, analyze an object. In this case, while it was being de-installed. So I had something on the order of 12 minutes to look at this object between when it came out of its case and when it went into its packing crate. And that necessitated the use of, as you can see in my hand, portable instrumentation to take a particular measurement.

And then, of course, we, sometimes, have situations that are somewhat less controlled. As you see in the image in the center, where, for example, we're actually taking our scientific technologies to the field, working in, in this case, a chapel in Florence, collecting data while the painting still hangs on the wall. Again, using portable instrumentation. So lots of environments, lots of instruments.

But all of them result in some form of data. So I'm just showing you here representative spectra that come from each of those three data collection modalities. What we do then, as scientists, is we take all of that data, whatever it looks like, and from however many objects it was gathered from. And our job is to do some sort of synthetic, analytical work to collapse all that information into an understanding of the objects. In this case, I'm sort of showing you one of the ways that played out for this particular work, which was done, I should note, for a specific museum exhibition called Florence at the Dawn of the Renaissance, painting an illumination from 1300 to 1350 which took place in about 2012.

And what we did for this is all of the information I learned from about 30 different objects had to be understood as a unit, where I built this table that was in the exhibition that shared with visitors, for example, what were the materials that we found in panel paintings, and compare that to what were the materials we found in manuscripts, what materials were found in both, and what did that tell us about workshop practice.

So hopefully, just the fact that I'm able to show you that deductive, it should tell you that, clearly, Conservation Scientists, this is the work we do. And we've been doing it for decades. Most people know what they're doing and how to do it.

So the question I hope you're asking is why do we want to introduce a technology solution into that process that already works in terms of creating new knowledge. It's fortuitous that we're having this conversation now in the time of COVID when it's been made more clear to all of us than it ever has been before, just how collaborative this kind of work is.

So I'm showing you here on the left, for example, an image of two curators one who works at an institution in the UK, one who used to work at the Getty, and was part of this exhibition, actually standing in front of a light box, and trying to understand what X-ray images of the manuscripts taught us about those objects. So we bring people together. We bring data from a bunch of different objects together. And we understand it through a collaborative discussion-type of modality.

That's a lot harder to do if you can't meet in person. So it would be really nice to be able to recreate that collaborative research environment in a more digital function. And as I said, COVID has made that more clear than ever before, that we need mechanisms to do this.

The other reason we might be interested in technology has to do with the sheer volume of data. So I might have just mentioned, offhandedly, that this particular exhibition, I looked at something like 30 objects. So I'm showing you here just one of them-- leaf of The Martyrdom of Saint Lawrence.

And from just this one object, I went ahead and counted it up. And it looks like I took about 450 individual pieces of data, about 300 images. And those images ranged from images of the whole paint [AUDIO OUT], all the way down to individual images taken from the instruments that we work with, that showed individual pigment particles, for example.

And all of those spectra and images came from about 14 different sources-- instruments, cameras, that kind of thing. So that's just a huge volume of data. And if you were to go to my office right now, and see how that data is stored, you'll find what I imagine you'll find at each of your desks-- a set of file folders labeled in a way that makes sense to me.

But in order to find any individual spectrum of that 447 spectra that I took, you have to know where to look in my file cabinet. And so, wouldn't it be nice if we could render this huge volume of data more easily searchable and more easily shareable. And then coming from a pure data and data science perspective, the other thing that it occurs to us is that that data [AUDIO OUT] searchable.

That would enable more complex data analysis. Sort of data science kinds of approaches to the data, as well as making sure that other people could find it, whether or not you know what my file cabinet is look like.

So all of these sound like things, I hope that technology is the various sorts can help us with. And it's why, I, at least, as an individual, began to be interested in understanding whether there was a technology [INAUDIBLE] some of my technical analysis questions.

But first things first, those of you who are at Dennis' talk last week-- Dennis and Annabelle's talk last week, heard them say, several times, that one of the important things that [AUDIO OUT] before you try, and you use or build a technology solution for your problems, that you really understand your problems and [AUDIO OUT] towards the questions you have.

And we wanted to make sure that was true in this [INAUDIBLE] as well. So one of the first things we did was undertake a series of visits to other individuals or organizations that do [AUDIO OUT] heritage research, located mostly in the East Coast of the US and Europe mostly, for ease of travel reasons.

But basically, what we wanted to do was go to these organizations and talk to people who do technical analysis research, and try and understand what their problems actually were. And to understand whether what I saw in my work presented what we saw in the field.

We did that through these interviews that we call Use-case research interviews. You can see on the map, on the screen, some of the places we went in these different places. But basically, what I want to kind of

[INAUDIBLE] here is that we tried to visit institutions that use a small number of visits to get a broad sense of the field.

So we [INAUDIBLE] both quite large, sort of national level museums or archives. We visited small institutions, which might have very different takes on what's required. We visited organizations that had connections to local governments, and ones that were privately run. We talked to museums, we talked to libraries. We talked to organizations that had a high technology sort of baseline, and ones that had a low one.

And we talked to people who work with very different types of data, whether that was imaging, spectroscopy, the kinds of materials they studied. And basically, we were trying to get a really broad view of what the questions that people have. And what would be most important to them in any kind of technology solution.

So what did we learn from those interviews? Well, service starting at the top, like if you have data from objects, what do you want to do with it. We found out that that's probably the most consistent between organizations. Everyone is very interested in taking objects, whether one or more, like you see two of the illuminated manuscripts we've been discussing in the corner.

We want to be able to find connections between those objects. We want to be able to understand what materials are present in each of these manuscripts, regardless of which instruments we had available to study them, for example. We want to be able study how these objects are changing over time or how our knowledge about them is changing over time. So the kinds of uses that we want to make up their data are pretty consistent among organizations.

In terms of the actual data themselves then, we also learned that one of the things that's really important to people is that they want to be able to have access to the spectroscopy. They don't just want a system that tells them in words, in language what this particular Raman spectrum showed. They actually want to see the Raman spectrum. They want to be able to examine that spectrum for themselves.

They're also interested in comparing spectra. So to look at, for example, both the Raman and the FTIR spectrum from the same area. But in addition to the data themselves, they really want to understand, and they think it's critical to know the context in which that data was taken.

So for example, it's not enough to just have access to the two spectra you see on the screen, you need to know that they both come from this little area that's marked by the red box. So that spot map is really important, for example.

So basically, understanding the context in terms of where data came from, why it was taken, might give you information about why the spectrum looks the way it does, for example. And people really want to understand that in order to maintain the fact or to keep the data interpretable over time, in short.

The other thing we learned is that people have a lot of concerns around data access and sharing. That is there's a strong recognition that not every piece of data, that a Conservation Scientist takes about an object, is going to be available for the same level of sharing.

And there's kind of a wide recognition that there's some information that people might be very happy to share with not just all of their colleagues, but even for example, the general public. Things like photos or contact information of who to contact if you have information or you want more information about a project.

There's also other types of information for which they might want to allow only selective access. Like maybe, you don't want everyone. Maybe you don't want the general public to be able to see your individual spectral data, for example.
And maybe there's some information that's actually really quite sensitive, like how you treated something. Maybe the actual uninterpreted data that you took. Maybe you don't want to be able to share that at all. And that some of this may be defined, not by the scientist who took the data, but by, for example, the owner or manager of an object or institution. So people had quite a bit of concern about that.
One of the other things that came up quite a lot in these interviews is we asked people if all of these sort of issues around data use, data access, data analysis, and interpretation are so common among different organizations. Why don't they do this in a digital platform already?
And to make a complicated story very simple, we learned-- if the slide advances, there we go, that one of the main things that's keeping this from happening has a lot to do with data organization and management. In short, we learned that I'm not the only person who has a kind of messy file door with a file organization structure that makes sense only to me.
We all have our ways of organizing our little cabinets or the sort of digital representations of those, whether that's on your personal computer or on a server at your institution. We all kind of do that differently. And one of the things that people really commented on is that that esoteric data management that we all have is one of the primary barriers to even simple data sharing, for example, within your own organization.
It is true that a lot of my data is stored on my laptop. And if I don't come in, one day, and my boss needs a piece of data, she may not have access to it. So the fact that we organize and manage our data the way we do has sort of big implications in terms of how effectively and easily it can be shared and reused, even within an organization, let alone outside of it.
That said, people understand, as a general rule, that having a more uniform way of organizing data in a way that makes it shareable would enhance not only our ability to find data, but to reuse it, and to avoid institutional information loss when data becomes inaccessible. So people understand that that would be useful. But we found that a lot of organizations didn't really have the capability to build something that would meet all these needs themselves.
And so, as we heard that, that's where the project, at the Getty, really started to take shape was we realized that there was a real need in the field for a system of some sort that would help conservation scientists manage their data in a way that would keep it secure, that would allow us to retrieve it when we needed it, that would allow us to compare information, share it with our peers. All of that kind of work.
In short, we realized that we could be useful in building the bottom tier of this pyramid of sort of data and data analysis that I just showed. And if we can build it, we're hopeful we can share that tool with the rest of the field.
Now, here, I will kind of back up to the video that started us out today. You will occasionally hear me say, DISCO versus Arches for Science. And I just want to mention that what we mean when we say, we need to build something, we do want to build something that I can use. And the version of it that I will be using at the GCI, we're calling DISCO.
the version of it that we can make freely available to labs worldwide, that's Arches for Science. So I want you to think of this, this is Arches for Science. And please, forgive me if I slip and say DISCO. We mean the same thing for the purposes of today's talk.

So hopefully, that gives you a sense of what the overall project it is. It's to build this data management platform, which I'm sure sounds fabulous in theory. But now, I want to start diving in a little bit deeper and say, OK, well, what would that mean? What would it mean or what would you need to be able to do to organize and manage the data from a typical conservation science project? And in this case, I am going to use the technical analysis of an object as a description.

And so, one way to do that is to figure out how to categorize the parts of a technical analysis project in such a way that a computer could be used to manage the data. So we'll start with an analysis that I did relatively recently of the painting that you see on your screen, here at left. This is Jeanne or Spring by the artist Manet-- Edouard Manet. And we analyze this, again, for an upcoming show.

Now, if I ask you to sort of categorize this item that you're seeing on the screen, you might say, well, it's a thing. It's a painting. It's a physical painting that exists in the world. And that is one way you could organize it. The other way you could organize it is by knowing that it's actually one of two paintings that we analyzed for this exhibition. That is, it's one physical thing in a collection of paintings by Manet that we analyzed.

Another to get a broader view of the project, you would probably have to know that I did quite a lot of the work. So you see me pictured here looking at the painting under the microscope. But I didn't do it by myself. I did it with my colleague, Devi Ormond, who's a conservator who was working on this painting. [INAUDIBLE] each of us are people. But together, we are the Manet research group. So that's one way you could categorize this important component of the project. What did Devi and I do? Well, we took a lot of data. So we went to instruments, like the Macro X-ray fluorescence scanner. You see an image of it here. We collected data and made observations about the painting through that mechanism of scientific analysis.

The other thing we did, that's very common in this kind of work, is we actually removed small samples from the painting, and looked at them as cross sections, which you see imaged here. So we actually went through a very specific kind of activity known as sampling. And through that sampling, created a secondary, little piece of physical material in the world-- this cross section. That is, of course, related to the painting from which it came, but what you might want to crack differently than the painting itself.

From both the sort of experiments that we do noninvasively and the samples, we collect a bunch of data. And that can take a bunch of forms. So we can have spectra, like what you see in the upper right, which happens to be a Raman spectrum from this cross-section. We may have data that takes the form of images such as the copper, zinc cobalt, chromium distribution map of the Manet that you see there in the center right.

But any way you cut it, these tend to be, in today's world, digital files that I can store in that esoteric file system on my own computer. So hopefully, that gives you a clue that one way we could characterize these are as digital files or digital resources that are connected to the project.

Typically, I would take all of this data, all of these images, all of the information that I learned, all of the interpretations of that data that I make, and I would try and write something down that collates all that information into some sort of final product. In this case, it take the form of publication of the exhibition. So it's a formal book. It may be an internal report for the institution. It may be both, which actually it is, in this case. We got both. But in either case, what I've done is I've created a standalone textual work-- textual material that talks about or includes the information from this project.

So all of those might be components that you would need to be able to describe if you wanted to build a digital representation of this particular conservation project. But even as I type through it, hopefully, it's clear that it's not just those concepts that we need to be able to organize, but actually the relationships between them.

So even talking through it, I felt obligated to say that this particular painting is part of a collection. I, as a person and part of the Manet research group, the physical sample resulted from a sampling activity. And these relationships [INAUDIBLE] things can get ever more complicated.

So for example, we've identified that a person is an important [INAUDIBLE]. And it's worth noting that in this particular case, there is a person or several people, actually, associated with the painting. We have the person of Manet who painted it.

We have the person of Jeanne Demarsy, which is the name of the historical figure who was the sitter. This is actually a portrait of a known person-- Jeanne Demarsy. We have that person that's associated with this physical thing.

We talked about the relationship between, say, the whole painting and a sample removed from it. And so [INAUDIBLE] and so on. Hopefully, it becomes very clear that really, there's tons of relationships between all of these components of a research project.

And if we want to fully and fully describe the relationship or the project, the relationships, I'm trying to sort of indicate by all these red arrows, are just as important as the components of the project itself. And have to be able to be recorded if we're going to create a digital surrogate of the [INAUDIBLE] conservation project.

To get a little more specific, that's actually what Arches for Science needs to be able to do. It needs to create a system to help us secure and create and visualize our data. But it needs to do it in a way in which all of those complex components of a project, and all of the interrelationships between them are recorded and preserved over time. So that's the overall goal and what we're trying to build.

Now, there are technologies that help allow this. One of them is what you hear about in the Arches presentation the other day, that is the use of sophisticated data models-- [INAUDIBLE] with what's known as semantic metadata that helped make the self-describing. That is something that [INAUDIBLE] physical thing is part of a collection or set.

And that [AUDIO OUT] allows material to be ported from system to system, for example. [INAUDIBLE] understand what that means, let's go back to this kind of seemingly simple example of a person. So hopefully, you will take it at face value that I am a person speaking to you. And if I ask you to differentiate between myself and someone next to me, there's a lot of ways you could do that. But one of them would be by knowing that my name is Catherine, that I'm identified by the name Catherine.

You could get even more specific [AUDIO OUT] too, and note that Catherine is my given name as opposed to another type of name, like my family name or a pseudonym, for example. The one way to define me, as a person, I am a person identified by a name, [INAUDIBLE] given name.

However, it's important to note that I have labeled this particular area or arrow that connects the concept of a person and the concept of a name by the phrase he's identified by. But I could have labeled it with something else. Maybe you want to say but this person is, I don't know, related to Catherine rather than is identified as Catherine.

And so you have to be really clear when you try and add this metadata to your underlying primary data. What you mean? What is the relationship you're trying to explain, and what are the concepts you're trying to connect?

Standards, of course, can help with that. And one really commonly used one is the CIDOC CRM or the conceptual reference model. And this is a system which describes, in formal language, the concepts and the relationships between them that might be relevant to cultural heritage data.

Now, if we were to take the simple relationships you see on your screen then, and encode them with the information in the CIDOC CRM, we get something that, functionally, looks the same. We're still saying a person is identified with a name, with a type. But now you see all these little number and letter codes that are sort of attached to that, which make it abundantly clear that we specifically mean a person is identified by this name. And this name is an appellation. It adds all the definitions.

Now, if that seems like a really fancy way to say that I have a name, it is, frankly. But it's a really important thing to do because it provides so much benefit. So I'm not going to read through everything on this slide. I'll let you read that while I talk at you. But basically, what I want you to take away from this is that using as something like the CIDOC CRM to encode our data with the semantic metadata provides a foundation upon, which other projects can grow, advancing the goals that you see listed here the left side of the screen.

So for example, projects like linked art make use of this to sort of create a specific community of users in the art space, and identify terms that are important to those fields. And projects like Arches, and by extension, Arches for Science can then make those things practical for someone to actually use it, like me, a Conservation Scientist to use it by building tools that let me take advantage of those conceptual relationships.

So as you can tell from the list on the left, there's a ton of benefits to doing this and they're all critically important. I don't want to imply that any are more important than the others. But just coming from the perspective of a conservation scientist, I want to point out a couple in particular. One of those is sustainability.

And so it's important to note that this concept of creating the metadata makes it very clear what each component of this project is makes the data self-describing and able to persist over systems over time as long as the CIDOC CRM is maintained.

And that's really important to me because it means that any data I take today may be available and understandable, critically, to someone many decades from now. And especially, as technology becomes a bigger and bigger part of the data collection process, this becomes more important. I mean, how many of us have experienced taking data with a software system that 10 years later, the company stops producing. So we can no longer read that data. Using a system like this corrupts that problem, and renders that data sustainable into the future.

It also sets the foundation for a spirit of openness and community and data sharing because all of the data are rendered roughly the same way right. We've all identified the parts of our data in the same manner. And that allows us to share it. And of course, sharing and communication is sort of the underpinning of conservation science and science writ large, actually.

So what I want you to take away from this is that the reason we're interested in using this is because we see that conservation of the data, that we take about objects in our care, is, itself, part of conservation. By

sharing that information allowing people to make more and better or distinct interpretations about our objects based on data that exists, we hopefully make it easier to avoid taking new data in the future. For if we maintain that sustainability, maybe I don't need to go back and take another painting cross-section out of the Manet if I can access, reuse, reinterpret the section that already exists. Sustaining this data to us is part of conservation.

Now, luckily, at the Getty, using these kinds of semantic technologies is nothing new, as those of you that attended the Arches lecture will well know. I'm not going to go into too much detail about the Arches platform, generally, and how it meets our needs to build Arches for Science. I do want to recap it just at a very high level.

Arches has a series of features that make it really attractive. One is it's got robust data import and data management systems. It has many, many tools for data discovery and searching, including semantic searching that is on those concepts that are designed with the semantic-- or designated semantic metadata. And it has the ability to build in comprehensive workflows to make the life of somebody using the system relatively straightforward.

So there's a lot to like about Arches. And it's why we chose to build on that for Arches for Science. And again, just to hit the point home, part of that is the fact that it's built on standards, like using the underlying CIDOC CRM ontology, the fact that it uses controlled vocabularies, like the art and architecture thesaurus, and uses IIIF or the International Image Interoperability Framework of standards to do cool things like annotate images.

All of that is important in this realm of conservation science data because it means that my data will be as we've discussed-- protected through time and accessible into the future. So that's why we want to build on Arches

So from there, we're going to get even more detailed and specific dive into what does Arches for Science actually look like. And to do that, we're going to stay with this concept of identifying a person. And show you what that really looks like in practice as opposed to the sort of the little model of just the concept of a person-- a name and a name type.

This is what it really looks like in person. If you take this concept of a person, you see listed here, and what I'm showing you here, I should note, is a data entry card for Arches and Arches for Science. And what you see in this blue column at the left are a bunch of the sort of concepts that you might be interested in a recording for a person.

So you see things like their name, their ethnicity, their nationality, their birth, and death dates, gender. So there's a lot of things you could record about a person if you were entering them into the system. But if I play this little video, which I hope works. There we go. You can see that, actually, that's just the surface level. If we expand all of those categories, there's quite a lot more detail buried under there.

This is all data that you could conceive of entering for a person. And rather than seeing three lines connected by arrows, this is what a visual representation of all of those connections actually looks like, and what we call a data graph. So there's a really deep amount of information here. It's much more complicated than what we looked at before.

But what's really cool about that, as a practitioner, is let's say, we've entered a particular person of interest into the system already. So let's go back to this Manet painting that I studied. If you want to enter a new painting-- Jeanne, the painting I studied, and say, that it was painted by Edouard Manet. If he's already in the system, I can just search for him, add him to the painting.

And then what I get from that is all of this rich, semantic information that's already been painted about Manet gets associated with my technical study of the painting Jeanne sort of for free or automatically.
Of course, defining just the concept of person, as we just discussed, isn't anywhere near enough to describe the full complexity of a conservation science project. And so in order to do this, we're creating data models that are just that complex or more, not just for the concept of person, but for many of the components that you need to, adequately, describe a conservation project, including some that we haven't talked about yet, and won't talk about today in any kind of depth.
But when all that is done, when all of those data models exist, what you get from it is a universe of connections that provides, then shows all the linkages in a conservation project. So this is just a very small example, but hopefully, you can kind of see that, for example, if we look at a painting from Manet, we can load all of the connections to that artist. We can see that the painting was part of a collection of paintings that was studied.
If we expand that collection, we see that there was a technical examination. If we load more connections, you see this line form that tells you that I was the scientist that worked on that painting. So that's a very small example of just sort of one painting, and some of its connections. And it's not even all of them. But hopefully, it gives you enough that you can imagine what this would look like as this web of information becomes ever more rich as the systems used.
Now, all that, said basically, there's a ton of complexity going on in this data. But that's what's happening of behind the scenes when the data is in the system-- once it's been entered. So I want to kind of take a step back again and say, well, how did it get there in the first place? How did this data get entered and do I, as the scientist using the system, do I need to understand all of that data complexity?
And the short answer is I don't, really. And that comes from the ability to use these workflows to define commonly used processes. So as again, you saw this in Arches that workflows are a way to sort of describe the most commonly encountered processes that we might find.
We're showing you just a small selection of possible workflows for a conservation process here. And these are kept as simple as possible for the user. As I said, hiding a lot of that complexity. And I hope that walking through a couple of these will make this a little more clear. And so, we can, for example, come to the system and enter a new project. I've been asked to work on a new painting. I'm going to create a new project to define that.
So how this would work is I would select that workflow. And I come to a screen that looks something like this, where we can see cams like we're used to seeing in many interfaces and some data entry areas. And I can say, I can just type into this field that I'm working on-- the technical and analysis of a particular painting.
And I just want to note that this is just a text block. So you could mean this in whatever way matches. For example, your institutional project naming conventions. There's also other pieces of data on this card. I'm just highlighting language. But I just want to note that, basically, each time you see one of these data entry screens, there's places where we're having data that's building some of those complex connections behind the screen-- or behind the scenes. Pardon me.
So if we've entered our project name, and we click through to the next tab, we'd see something like this. This is called the Project Statements. And you, again, see a big open text box. And what this is actually designed to do is to actually allow us to record some of that information that we heard people talk about during the use-case interviews as being important.

That is, why was the analysis done? What were you trying to do? What were the goals? Because understanding that context may help someone understand why you have the data you have, and how to interpret that data if they come to it later.

And I want to note this one in particular because this is the kind of information that is very easily lost in institutional memory. If I leave the Getty, 10 years from now, is anyone going to know why I studied 32 Florentine objects? Is anyone going to know why it was important to look at two particular paintings for Manet? This allows us to actually record that information and make sure that it's not lost.

I'm not going to talk about the next tab, which is Project Timespan. You heard Dennis talk in the Arches presentation about complexity of time. So we'll just let that lie. But I do want to talk about the next tab, which is adding the project team. So again, this is something that can easily be lost in project history-- remembering who worked on something. So we do have the ability to add that to a project.

Here, how I did that was by starting to type my own name in the search box that you see here. And what you see pop up is because the system has been pre-populated with people of interest, you see that to Catherines popped up. So I can select my name, and then associate that with the project.

However, one of the really cool things about this is, remember, all of this is built on that semantic metadata, which means you can also search on things like a concept. What does that mean? Well, in this case, it means let's say, somebody is entering this data for me. Maybe my manager is-- or not my manager, but somebody's manager is entering a project into the system, so that multiple people can work with it.

And they can remember my first name, but not my surname. So they can't remember-- is it Catherine L. Or Catherine P. that's doing this work. Well, what they do remember is that the Catherine that worked on this project is a scientist. And so for example, this person could start to write the word science in this search box. And what pops up is a list of people who are members of the GCI Science department. Because there's been this semantic linkage of these members as being part of a group that has the word science in its name.

And in that process, they can say, oh, OK, so yeah, that was Catherine P. They can select that name and add me to the project. So that ability to search on a concept instead of just a search term can be critically helpful. And again, we will walk through it, but you could do the same thing with for example, adding objects to the study, and things like that. And in so doing, create a new project in this list.

So once you finish that, what comes next? Well, I've been talking about data this whole talk, so let's talk about how we would associate that with this project. So we can select the Upload Dataset tab, and we see a very similar workflow.

And I want to start by noting one of the cool things that Arches does, as a general principle, and the Arches for Science does, specifically, for conservation science is it's been modified to refer or reflect the work we actually do. And that means that the first thing we do here is select one of the different types of data that are collected in a conservation science project.

In this case, that kind of means sort of differentiating between the types of analysis we do in situ on the objects. So the images, the action images I showed you of me collecting data at the very beginning of this talk. And we can note that that is materially different from analysis, where we might take a sample and do something destructive to that sample such that that material doesn't exist anymore. Semantically, those have very different connections behind them. So we can differentiate between them.

So we do that, and then we go through screens that look very much like what you've seen before. We can do things like select the project that a coordinator just made for me on our last example. If there is a project like the Florence show that had 30 objects involved, I can specify which particular object, in that research project, are we uploading data for.

Similarly, you can do the same thing with instruments. If your laboratory, all of the common instruments have been loaded in, you can select that.

To upload a Dataset, that looks like something we've all seen in many, many different systems in the past. We can just drag and drop relevant files onto the system and upload them. But then we come to another part that's really specifically tuned for conservation science.

So one thing we haven't talked about is the ability to do image annotations such that we identify the areas of an object, in this case, a painting that you see here, as areas of interest or areas for which we did take data or make observations.

But that's something we can do. We can select and annotate those areas, and name them. So you can see here, for example, there, are two named areas-- one in the Angel's face and one in the hand of a figure. And they're just called area 1 and 2.

But once that exists, when we're in this data uploads that we have an extra ability, which means we can take each of those areas, and look at the individual data files that we've just uploaded, and associate, select area 1, let's say, the Angel's face, and say, that all three of these files that I've put a checkmark on all came from that area.

In this case, let's say an example where I used-- this is X-ray fluorescence data, so I used it rhodium tube at 50 kilowatts driving voltage. I did it again at lower voltage to try to get a better understanding of the light elements.

And I still wasn't happy with my data, so I did it again with a different tube, a chromium tube, to get a better view of, for example, whether there was any silicon in this area. But all three of those spectra, fundamentally, came from this Angel's face. So we can actually say that in the system and link them together.

Lastly, when you've done that, you can, as we heard in our use-case interviews, people really want to be able to look at the data, so they've been busily building in lots of spectral viewers, so that we can see that data. And actually, you can see here, we can enter specific file information.

If we want to say that I counted for 30 seconds to make the scan, we can enter that information here. And we can interpret it. I could look at this and say, well here are the elements I see in the spectrum and leave that information for the next scientist to come along.

So hopefully, that gives you a sense of what would actually be like to use the system. So I just want to end by saying sort of what we're doing now. As I said, we're building out these data models, many of them, for all of the different concepts that would be components of conservation science project. We're continuing, along with the changes to Arches over time and the maintenance of that system, we're continuing to understand and develop new ways to visualize and search that data once it is entered into the system.

And of course, we're continuing as we build out the data models and ways of interacting with it. We're testing all of that with data from real projects. So that's the work that's kind of happening now and into the future. And with that, I think I'm at about the time that I was slotted to speak. So I want to stop. I'm just

going to throw on the screen here the names of just some of the people who have been really helpful with this project over the last several years.
But other than that, we'll stop here, and open the floor for discussion. And hopefully, myself and Dennis can answer any questions that you may have.
Thank you very much, Catherine. While we open out these questions for everybody to comment on, Dennis, do you have something to say before we articulate the questions that people have put forward in the chat?
Thanks, Anupam. I'll just say, I thought, Catherine did a great job. There's a lot there. And even at cursory glance, it will show you that she's thinking about lots and lots of information, working with sophisticated instruments, creating quite large datasets, and collaborating with lots and lots of people. So I think she really made a great case for how challenging the data management work is in the field of conservation science.
Thank you. Thank you, Dennis. Catherine, we might need you to share the screen once again because there are some requests to see some slides again. If you don't mind, do you think that would be OK?
For sure.
Thank you.
For sure.
One of those that they wanted to see was one of the slides in which you had the US piece.
I'm sorry. Not sure what happened there.
OK. It'll come. So Ruzbeh, you want to ask some questions related to that slide also? Ruzbeh, please.
Nothing in particular. I just wanted to study what options are there [INAUDIBLE] if [INAUDIBLE]-- I think, unless we try, we won't know it. That's for sure. But I think if we have a reference book in hand, they can [INAUDIBLE] and see what [INAUDIBLE] they're missing in the software with those quite expensive [AUDIO OUT].
We'll have a look at that slide once it comes on,
My whole screen had gone black, actually. I'm looking at nothing.
Yeah, It did, It did.
But in the meantime, while that is sorted out, Bhasker, do you want to articulate your question, please, yourself?
There we go.
Yeah. Thanks, Anupam. I specifically wanted to ask of what are ways that-- you spoke about data conservation, you opened a new channel there for me. Thank you. And that naturally, led me to the question.
So if you're talking about data conservation-- data conservation, in my understanding, is kind of generational. That's what we've been doing in the publishing industry-- conserving data from HTML to XML, and then to HTML too, and now, the JSON, and so on and so forth.
So this conservation is actually an industry that tends to be generational is my experience. So what do you say about your data standard, and it's generational limitation, and its potential capabilities beyond it. Catherine, do you want to take that?
Yeah.
So this is an important question. And I think it's important to distinguish the format of the data versus the structure of the data. And we're making quite an important claim here. And it has less to do with the

format of the data and more to do with including what Catherine has been calling the metadata-- the descriptive information that's embedded in the way we structure data in our shoes.
It's right back to the semantics. And it's why she talked about the semantics so carefully in her talk. The key idea here is that when you export data from Arches, you're not just exporting the values of the data, you're not just exporting the interpretation, you're actually exporting a data structure that includes the metadata that lets you, as a person, and perhaps, even more importantly, that a computer, a machine, understand the context of the data because we use semantics.
And not just general semantics, but semantics tied to a specific standard-- the so-called CIDOC CRM. The intent is to make it much, much easier to interpret the context of the data, and not just the value of the data. So it incorporates both the definition of the information and the information itself.
Yeah, thank you. The better and the content are both together all the time.
That's right.
Thank you for that.
Thank you, Bhasker, for that intervention. Ruzbeh?
Yes, sir.
You had a question also. Would you like to articulate that question, until we can get to the slide?
No.
Are you not seeing the slide?
We're seeing the first slide. We're into that first slide.
OK.
OK. In this semantic linkage that you're talking about, can it be done by anyone? Then does anyone can link anyone, like is this an IMDb kind of a scenario if your part of IMDb. It's an open project. It's been taken over by Amazon if I'm not wrong. And it's now a paid project now. And any actor can be linked to a movie he was acting in, or any director or [INAUDIBLE]. Any part of the movie industry can be linked to in that particular movie.
So is this something like that? Because that can happen without permission. Then somebody has to raise an objection, and the moderator has to intervene. They have a role to play over there-- a process of objection and this and that. So is this something if we open the house, like the lady said to everyone, it would be lovely. But if something like this happens, if somebody links wrong Catherine to a wrong person or something like that, then open on, how can-- So that's the question.
Dennis, do you want that?
Sure. I'll take that. So again, there's, I think, two ideas that are kind of mixed together here. One is this idea of linking information. And the example of the linking actors to movies, that's actually, I think, it's a good analog.
Arches has, essentially, the same idea at work, when you use Arches for Science or really, any Arches project, you-- in fact, in effect, you're doing this. You're making these linkages happen. So the example, I think, is a good one.
The big difference here is that Arches has quite sophisticated data security protocols built into it. And it typically wouldn't be available to just anybody on the web. I mean, Catherine's been talking about Arches for Science, specifically, as a tool to help conservation scientists manage their data.
And you have to have an account to even make any data entry possible in Arches. So you interact with the system based on the permissions that your account has been granted. And Arches supports quite

granular level of data security. So you can allow some people-- and everybody has to have an account to see it.
Some accounts are only granted view status. You can only look at data, and only select the data. Other accounts are granted the status of creating new data, including these relationships. So it really is a decision on the system administrator on who to allow access to the system in the first place, and what kind of access they're granted. So hopefully, that's an answer to the question.
Thank you.
And is this slide that you were looking at, Ruzbeh?
No, sir. I think the three slides to the right, maybe. That's the one.
The one that went back here. One more.
Can we get to the back?
One more back or this the one?
Yeah.
OK. While the slide remains here for some time, Catherine, Juhi Bafna-- Juhi, you want to articulate your question yourself, please.
Yeah, sure.
You want to introduce yourself also, Juhi.
Yeah. Hi. I'm a conservation architect. I work at the Center for Heritage Conservation at CEPT University. And we're mainly in the building heritage [INAUDIBLE]. So I just wanted to ask whether the data for even agent can be downloadable into a report or printable format or to store it or to disseminate or engage with, physically?
Yeah. That's a good question. And the short answer is that's something we're working on right now. So it's clearly important to be able to communicate our results, and share them with our clients, for lack of a better word. I'm very often working with a specific conservator or specific curator. And I do want to be able to provide them written documentation. And as you say, to store documentation in written form as well for institutional reasons.
And so we're working towards finding ways to export certain critical pieces of information into report templates. And that's important to me as a practitioner, partially, because if we can automate parts of that process, that becomes something that doesn't take my time. And I can spend more of my time really digging into the details of a spectrum and analyzing it, and letting the system-- letting Arches for Science create the report from that for me.
I don't know if you want to add to that at all, Dennis.
I think, that's a nice summary, Catherine. I'll just say that Arches and therefore Arches for Science supports data export. But I think the question really here is a question about structuring data like a report and supporting download of that. And you're correct. This is the kind of thing we're working on, really, kind of as we speak. This is the topic of conversation with the development team right now
So if this study, if I may just continue, so if this code is open source, it can be worked upon and developed by someone else, and it may be proposed. Is that possible to engage it at the development level?
Yes. That's an excellent point. And it's correct. And it is open source, which means that you are able to engage with the code and extend it or add to it. And we welcome that. We really are hopeful that people will take the code and adapt it to their use.

And your only responsibility is to share your work with others. If you do that, and you're free to do that, you're only really responsibility is to make sure that you make your own code available for others to take advantage of your efforts as well.
And that's the thing of open source development-- to have that community.
Exactly. That's exactly right.
Thank you so much. Thank you.
But excellent. This is very nice, Dennis. It's very good. I think, here, in this Coastal Program, many of us are just talking about breaking silos and getting people to share things with each other, and having a transdisciplinary sort of an exercise. This becomes an important part also.
Yeah.
So in this context, I suppose it answers it in some way. But Angeli, you wanted to ask about access to Arches. Would you like to say that? Would like to ask your question?
Right. That is again unstable because of the winds to the [INAUDIBLE] what we do. Thank you for the presentation. My only question is I know that it was mentioned before-- that DISCO is the part of, basically, the customized version of the Arches software that has been made to use for getting.
And the other version is basically Arches for science. So I'm just wondering whether either of the two part of [INAUDIBLE] download on that. Is that page itself or is that something that you will have to customize at our own? And how does it work?
Shall I take that, Catherine?
Yeah, probably.
I think this is also an important question. It really comes down to the distinction between Arches for Science and DISCO. And fundamentally, they're essentially the same software with some configuration changes.
So let me see if I can make that-- I'll try to make this as simple as I can. Catherine's working in a laboratory with a specific kind of art, and specific kind of objects with specific instruments, and with specific reporting requirements.
And let's pick one example. Let's say, Caherine decides that she likes the report with a particular template. Arches for Science will provide a template. But if Catherine decides that she wants to use a specific template, and perhaps, use a specific controlled vocabulary, those are configuration changes. And they represent the specific requirements of the Getty.
So when we do these very specific configurations, like choosing a particular controlled vocabulary to implement or to create a very specific kind of template, when we do that, we call those customers that those configurations-- we like to keep the configuration separate. And so that's the distinction between Arches for Science and DISCO.
DISCO is Arches for Science, but with specific configuration decisions, like specific control vocabularies and specific templates. So Arches for Science will be publicly available. And DISCO will make the report templates available as well.
You can use them if you want, or not. It's up to you. But it's a very specific decision by a very specific Arches user. So hopefully, that makes clear the distinction between Arches for Science and DISCO, and what distinguishes the two.
Thank you.
Anupam, you're muted.

Sorry. [INAUDIBLE] asked that are there any issues of problems like duplicate data that are generated by self in it or something that risk is not there. Or does one have to be conscious of it at all?
I'm not sure I understand the question. Could you repeat it, please?
Yeah. It says, can we face problems like duplicate data being generated on its own?
So could you to Catherine Pattersons, basically.
I can say that, Catherine, if you like.
Yeah.
So this is fundamentally a data structure and data validation question. And one of the things that lets you do is define uniqueness rules. So you can implement rules in Arches that help preclude creating duplicate data.
For example, you can apply a rule that says that the given name and the family name of a person must be unique. That combination must be unique. You can make that rule. And if that's the case, then you can only have one Catherine Patterson in the system.
But as unique as Catherine is, there might be other Catherine Pattersons in the world. So maybe that rule is overly restrictive. Instead, you might want to say the rules should be that the given name, the family name, and the role of a person in our organization needs to be unique.
And so these kinds of rules can be implemented in Arches and Arches for Science, so that you minimize the chance for duplicate data. But notice what I'm trying to say here, which is for every time you create a rule, there's an exception to the rule.
So this is, in some ways, a sort of a philosophical question-- what is a duplicate data point. I mean, I'm going to say that there are more than one person in the world that Catherine Patterson. So your Arches for Science dataset may well need to have two Catherine Pattersons in it.
Are they truly duplicates? Well, there's only one Catherine Patterson who works at the Getty Conservation Institute. I've met her. And I can tell you, she's unique. So there's the distinction between the data and the thing you're describing. So you just need to think carefully about uniqueness. It's a question that goes well beyond Arches.
But correct me if I'm wrong, Dennis. That is part of one of the things that can be customized-- is those rules.
Yeah, absolutely. Like I said, you can implement those rules in Arches. And you can make them as detailed as you wish. So it's definitely possible. You are absolutely able to do this. But think carefully about what you mean by duplicate.
Yeah. And I would just add to that by saying that's kind of why I bother to point out that for example, things like naming the project, you might wish to implement the use of something like Arches for Science in such a way that it mirrors your existing institutional naming conventions.
Because presumably, anyone for whom this is a problem, has already had to think about it. If you have two objects with the same name, you've had to find a way to differentiate those in your own records. So this can mirror that process.
[INAUDIBLE], you want to say something?
No. Fine, Anupam. In fact, that question can be skipped because the slide has answered it for me.
OK.
Thank you.
Drew, you want to say something, Drew? Drew [INAUDIBLE]?

Yeah, sorry. So I'm just wondering, firstly, if there is any other data management systems out there. And the one that comes to mind immediately is, of course, the Google Arts and Culture Project. You've talked about duplication of data. And I'm wondering whether there's duplication of data management systems out there as well.

Well, that's really up to you to decide, realistically. You know your data management requirements better than we do. So you can make the decision as to whether we're duplicating what Google offers or not. I mean I would say that Arches for Science offers some very unique capabilities. And also makes it really clear who owns the data. You own your own data, and you control who has access to your data in a way that you can be absolutely sure about.

You also have control over how you describe your information in a way that I think is fairly unique. I'm sure there are a lot of data management systems out there. How many of them provide the same set of tools as Arches? I think, it's a relatively small universe of tools. And in fact, it may be that Arches is relatively unique in that category.

I'll just say that if your question is, are there other options to Arches, the answer is very clearly, yes. There are other options to Arches. And there may be options that are better suited to your needs. I don't know. Again, you would know better than we would. But what we're trying to do with Arches is to ensure that you can describe your data, structure your data in a way that is specific to your needs.

Create data sets that will outlive the software because they are going to use semantic models that provide the metadata necessary for a person or a machine many years in the future to understand what your data are describing. And to provide tools to make it easier for you to interact with the data and share with your colleagues.

I will say that's a pretty ambitious goal. And if you're aware of other software that does that, it'd great for you to share that with us, because I think we'd be very interested in seeing how they're doing it.

And I think part of the reason that we wanted to build Arches for Science is because while there are a lot of things out there, I think the other thing to really keep in mind is the audience for the different types of data. And so, yes, there are industrial softwares that chemical laboratories the world over used to hold spectral data, for example.

But those softwares, in general, don't have a way of linking that to a physical location on something like a physical object, something like a painting. The Google arts culture stuff privileges the types of information that we make public. They're not geared towards some of the data that you or your institution might be hesitant to share.

And so it's something that we're building to deal with a very particular nexus of needs for the control is then to set up that information. And to be able to met it out to particular users.

Just the follow up to that, if I might ask. Is it possible for data management systems to interface with each other?

Definitely. And so, for example, for Getty, we use TMS to manage our collection data. And there's always ways to, for example, allow Arches for Science to draw from TMS as a source of data. So that I don't have to individually enter all of the objects of our collection.

Dennis, do you want to expand on that?

Yeah. There are ways for data to flow between different software systems. And I mean, Catherine gives a really good example. That their museums will typically have a software to manage their collections. And

the objects that we study with Arches for Science, ideally, would draw from those collections management systems. And in fact, we can do that.
Typically, there's a thing called an API, an Application Programming Interface, that allows developers to transmit data from one system to another in a secure way. And Arches is capable of both using another system's API to request data from it.
And it also has its own API to allow other systems to request data from Arches. And we could have a whole conversation about how this actually works in practice. Let me just say that these programming interfaces are designed to share information. And some can be quite sophisticated, like in the case of Arches, you can use an external system to request data from Arches. You can also use an external system to provide data to Arches, programmatically.
In other words, automate the sharing of information from one system to the other. Automate ingesting information from one system into Arches, and automating the sharing of data from arches into another system. That's quite possible.
And it is really important in the practical use of the system. For example, so if the museum acquires a new object, that sort of automated ingestion from the collection management system, would mean that that object is available for me to create a project around in Arches for Science, for example.
You're muted again.
Thank you.
I have one thing to make-- a statement to make. So there is a group of hundred here, of us, who are processing what you are saying, and according to our own capacities of taking it or understanding it, and our own experiences, and possibilities of being able to take it further, that's also there, either individually or as collaborators.
Now, from your point of view, having invested your time, and having the grace to also share your experience with us, what do you see for this audience for this multi crowd, so to speak?
What do you see as various ways in which you would, perhaps, if not expect, but in ways which would perhaps like to see the different members of different groups here could engage with what you are presenting to us?
Sure. I think the easiest way to talk about that is to go back to what I was kind of saying-- is what I kind of hope as one of the fundamental takeaways of this presentation and the Arches presentation before this is this concept that the data you produce about are works of cultural patrimony are important. And we shouldn't think of them as transient or things that it's OK if they get lost on someone's desk or just become inaccessible.
And so if you take that as a sort of fundamental building block, depending on where you are, as you say, as an individual or you as an organization are, hopefully, this starts to inspire people to start to think about data management and data organization as a fundamental part of the work we do. Regardless of your job description-- whether you're a scientist, a conservator, a project, or a site manager. Conserving data should be part of your job description.
For people who are in positions of sort of directing staff at a museum, for example, maybe you want to start thinking about is there actually a need for somebody in our organization whose role is to help with data management. So that whenever your organization gets to the point where they are ready and able to implement something like Arches for Science, you have the sort of mental picture of what you're doing and why in place.

And then, of course, if you're ready to implement now, that's a whole different ballgame. But regardless of where you are sort of in that process, I would say, figuring out what data management needs you have, and how to advance that as a process.
Thank you so much. Thank you so much. There were two hands. We had five minutes for the 90 minutes to close. [INAUDIBLE] Steve, would you like to go first, and then [INAUDIBLE].
Yeah, sure. Hi, Catherine. Thank you so much for the presentation. I was just wondering, you were talking about the many specimen [INAUDIBLE]. And finally we reached that one slide that had all of those linkages on. I definitely understood what you were trying to say, but I was trying to read it and your audio got cut somewhere in between. So I was wondering if you could just show that slide once more.
The one with those red arrows.
Red arrows.
Yeah.
[INTERPOSING VOICES]

Yeah, this one. Yeah. Yeah, the anatomy-- finally, when the arrows are held. Yeah. I just had that last part I wanted to have a look at it once. That's it.
And I should note, this is just sort of a quick visualization of the process. But hopefully, what you take away from it is that there are many ways to categorize a project. If you look into how we are currently doing that.
Yeah.
It's not the only way one could. I'll say that.
Definitely. [INAUDIBLE] what's [INAUDIBLE]
So I'm an architect myself, but I'm working with an organization for socialist collective, where we're working with artisans of different communities who are working with materials like oxide and artisanal handmade tiles, and things like that. So I'm primarily an architect.
But another thing that the organization has been doing for the last decade almost is household heritage and personal history. So we're looking at conservation and exhibits of objects at home that speak more about intangible heritage than they do about tangible heritage in that sense. So that's the [INAUDIBLE], which I've been working for the last two years.
Yeah. And you can imagine if you were to try and sort of do the same kind of visual graphing out of what a project around intangible heritage looks like, you're going to end up with different categories.
Definitely. Yeah.
And so I think that comes back to something that I know Dennis said during his talk. And that hopefully I remember to say. It's really important whenever undertaking any kind of project like this to be really introspective about what is the important data to track kind of project. And recognize that it's going to be different for different organizations for different types of projects.
There isn't one answer or one solution for everybody. And so that's a huge part of the work. And that's why I say kind of coming to that sort of understanding as an organization as to what are your data management needs, what is the impact that you need to be able to carry into the future is a huge part of the work. And it really can't be understated how important that is.
I think, especially, when you were talking about that space to record a statement as to why you chose to do a certain project-- pardon me there, equivalent of that would be something of great importance in this

case, especially if say, I'm going to record intangible heritage or an account of intangible heritage, if I may put it that way, then those are the data inputs that become more important to us later on if we were to carry the study forward, right?
Absolutely. And it's worth noting this is also the importance or one of the reasons that it's important to use standards. So [INAUDIBLE] the CIDOC CRM has categories for things that are more intangible. So you can talk about more sort of esoteric concepts like the information content in something. So you're doing personal histories if you're trying to capture what's important about a story that's told through an indigenous group of people over time that's not written down.
There are components of the CIDOC CRM that deal with some of those things. So again, it's just have to be able to define that.
Yes. Thank you.
Thanks very much for that. Thank you. Thank you, Catherine. Thank you, Dennis. Dennis, would you like to say some words to us before I conclude the session?
Thanks, Anupam. I feel like Catherine's done a great job today. And I don't have anything to add. Catherine, anything from your side, just something for the group that is here.
I would just say enjoy playing with your data. And know that even this is something that all of us are struggling with and still working our way through. So it's going to take all of us a while. But hopefully, we all get there together.
And if I may ask Alison. Alison, are you here with us?
Yes, I am. Hi,
Alison.
Hi, Alison. Thank you so much, Alison. Though you did not speak on the first day, but we really thank you very much for helping all of this. And we're very grateful for the support, for graciously having your team with us. Would you like to say something to us before I asked Meena to conclude the session on behalf of all staff?
Well, I'd just like to thank everybody for spending your time with us, and listening to what we think is a fantastic project. And we're always happy to talk about Arches. And I think Catherine did a great job today, and Dennis, as usual. And if you have any questions, just please let us know. We're very passionate about our work. And we're always willing to share and help support others who are interested. So thank you.
Thank you so much. Thank you very, very much. Thank you, Alison. Thank you, Alison. Meena, may I request you to please talk to us for a little bit.
Well, I'm just so grateful to the Getty Conservation Institute. You've generously shared so much information with us. And Alison, I was also just looking in the chat, that you were saying that all of your information, but it's Arches or DISCO is it's open source. And so I think what Anupam had asked is an important question.
And this is to all our participants, and to the organizations you represent or work with is for us to continue to think together that how we can engage with what we have heard in these last-- I think it's been two different-- the public talk as well, that how can we engage with what Getty has exposed and shared with us.

And if it is an internal cost of setting up a system or infrastructure, which is something you're already doing, then I think the question we have to ask ourselves is, how can we pool our resources and make it more efficient. And I think that's something that I have taken away from these two lectures.
And Anupam, I hope that we and all the few participants, we can put our heads together and think about what we can do in India, and for the work you're doing. So my sincere thanks to all of you, again, for actively participating. Over to you, anupam
Anupam, unmute one more time.
Yeah. Sorry about that. But thank you very much, ladies and gentlemen. This has been really, really good. It's been very enabling. And thank you, Alison. Thank you, Catherine and Dennis. And please convey regards to Tim Whalen and Jim there at the Getty. And we're very grateful and we are all aware of the support Getty has provided to everybody around the world. And we are one group in that. Thank you so much. Very grateful.
Thank you all for having us.
Thank you.
Have great night, everyone.
Thanks, everyone.
Thank you.
Bye.
Bye.
And Chelsea and Megan, thank you so much. And ladies and gentlemen, thank you for joining us. We'll see you again very soon. Thank you.