Transcript Summary 2
Individual Presentations by Workshop Participants:
Perspectives of Stakeholder Groups

Workshop on User-Centered Design of Language Archives
20-21 February 2016
Co-Organizers: Christina Wasson and Gary Holton

Transcription by Heather Roth

15 July 2016

Document Information	
Title	Transcript Summary 2: Individual Presentations by Workshop Participants: Perspectives of Stakeholder Groups
Author(s)	Christina Wasson, Gary Holton, Heather Roth
Original Creation Date	15 July 2016
Version	1.0
Date of Current Version	15 July 2016
Revised By	
Description	Transcript of second part of Workshop on User-Centered Design of Language Archives, held 20-21 February 2016 at UNT
Rights Information	© Christina Wasson
Licensing Information	This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0). To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/.

Revision History	
Previous Version	None
Changes	None

Table of Contents

Explanation of Links Between Transcript Summaries and Video/Audio Files

The transcript summaries of the workshop have been linked to the relevant video files. Hyperlinks are provided in *blue text*, and time code for the video is noted periodically.

At some moments during the workshop, the videocamera failed to record. For those spots only (which are in transcript summaries 2 and 3), we have provided hyperlinks to and time code for the relevant audio files.

See the **Table of Contents: Transcripts and Recordings** for a full mapping of the relationships among transcript summaries, video, and audio files.

Language Communities

This portion of the workshop took place on February 20, 2016.

Each workshop participant was asked to provide a 7-minute presentation from their viewpoint as a stakeholder in one of the five identified major stakeholder groups. Although each participant may belong to more than one group, we asked them to speak to the one they identify with the most and answer some questions. PowerPoint slides accompanied their presentations.

Start Video 2

[00:05] – **Loriene Roy**

Loriene is a member of the Anishinabe Tribe and she begins by sharing what the workshop gathering makes her think of, which includes: her students at both the University of Texas at Austin and the University of Hawai'I Manoa, her tribe, and the tribal schools and colleges she has worked with.

In regards to language archives, Loriene says she is "most concerned with the Native voice," working with tribal communities, and protocol. She notes various particularities about working with Native individuals, these are: disclosure of sensitive topics, spirituality, and understanding protected knowledge. In addition she discusses love and forgiveness as integral aspects of these types of collaborative projects with Native communities. "We forgive you for taking our words, we forgive you for recording them sometimes without our permission, and we love you for having helped us move to the next step."

Loriene then shares information about her own Native language, Anishinabe, and that "it is not dead," but "based on federal policy, some of us were never allowed to speak it." So she asks, "how do you reach everyone?" especially in a dispersed community.

[03:44] – **Daryl Baldwin**

Daryl says he is representing the Miami Tribe of Oklahoma as Director of the Myaamia Center and Co-Director of Breath of Life, but will stick with the community-based efforts of the tribe.

Since Daryl has moved to a leadership role as Director, he is not into archives as much anymore, but his staff and students are. Although he and his staff understand the complexity of archives, "the average community member is not necessarily the one going to these archives." So one of their tasks is to help community members and researchers find research materials as well as "provide easy access" because access is a problem. At this point they are trying to build their own digital archive, "we as a tribal community have very specific needs and we can't necessarily expect archives to make huge changes on our behalf." Therefore, Daryl says we (his tribal community and the Myaamia Center) feel a responsibility to create something meaningful for their own purposes.

Daryl provides an overview of the various archives he and his staff have accessed since the late 80s:
- Some archives have been easy to navigate

- Sometimes they find language materials mixed in with other languages they didn't know existed
- Attorneys have gotten involved to allow sharing of materials between the Miami Tribe and another group who did not want to share information to any US institution

Daryl sums up the reason for accessing these other archives as an overarching goal: "to get high-resolution copies, digital surrogates, of all known Miami-Illinois language materials so that we can create both the research tools and language learning tools to do the work that we've been charged with doing."

He notes the difficulties of getting access to good technologies for this kind of work. "Something as simple as being able to search a 300-page manuscript, and in a particular entry, to be able to recall the actual page that entry relies on is very important to our work." This is exacerbated by the Miami language's status as previously dormant but has now been in the process of revitalization since the mid-1990s. "There is a huge research component to that," when a community bounces back from losing its language. Thus, Daryl and his colleagues are piecing together 280 years of language materials, including works from the Jesuit era, which is what the Miami-Illinois Digital Archive is currently helping with as an NEH-funded project. Now that they have access to such a tool, he is "excited" about what the future holds but they need to be "strategic."

Lastly, Daryl relays some issues among archive use. For instance, people don't know how to find or choose materials. In regards to technology, there may be a difference between generations, but in general most of his tribal members have access to and know how to use, technology. An archivist as a guide "would be a huge benefit." Right now they're working on relabeling things. He also talks about "web presence" and connecting archive materials with additional resources. He says they are also trying to engage the younger generation because they're very tech savvy. Finally, the cost of digitizing materials with high resolution can be anywhere from $3 to $7 a page, which is a challenge to keep up with.

[11:35] – **Edward Alexander**

Edward is a member of the Gwich'in Tribe and he works with Dinjii Zhuh K'yaa, which was formed in 2010 at the Gwich'in heritage gathering. The Center is located in Fort Yukon, Alaska, a fly-in only community of about 700 people. Dinjii Zhuh K'yaa means "Gwich'in language" and a group of elders helped form the organization with Ed to represent each of the Gwich'in villages in Alaska. It was formed because "we perceive a threat to our language, and we perceive a threat to our continued existence."

Ed shares how the gathering in 2010 revealed a communal desire to have an archive. "They wanted it to be user-centered, even though they didn't have those words." Instead they said they wanted "a place to share." He sees the center as a source of innovation, where they make new things from the past. They also think widely about accessibility to make materials open "for generations to come." They are using innovative techniques to record knowledge and that is then "imprinted on the land using augmented reality" so elders and younger generations can explore information. For instance, their building is scannable; by using your phone you can pull up windows of information. His vision is to eventually have the people who use these programs to record and scan their own materials for others to explore.

Ed describes how he found a recording of his grandfather, before he passed, in the Alaska Native Language Archive, that he didn't even know existed. "It was a central moment for me." He has also gone to the University of Hawai'i at Hilo to see how they used their archive "as a source to generate new materials" which he found "inspiring." Lastly, he has used the National Archives in Washington, D.C.

Ed explains the importance of sharing materials and providing access to materials to his community because the majority may not realize that their ancestors have been recorded and/or documented. He also believes in providing the original source material of information for the sake of research. So Dinjii Zhuh K'yaa serves as their own "storage house" as well as a place to create new materials from the old.

Some of the issues they encounter are:
- Lack of funding: their facility doesn't receive funding
- Little manpower: he himself is a volunteer
- Spotty internet access: only have access at two or three public terminals in the villages

Some of the goals they have are:
- To ensure the archive serves as "an integrated hub"
- Use a radio signal to broadcast their archive holdings
- Collect data in a more streamlined way
- Allow users to be able to collect and store information independently of their help
- Collect additional materials, for example, beadwork/sewing patterns

[23:05] – **Wesley Leonard**

Wesley is a member of the Miami Tribe of Oklahoma and is teaching at Southern Oregon University. His work has mainly focused on "language reclamation [of sleeping languages] using archival materials, particularly older archival materials." He distinguishes language reclamation from language revitalization "as a broad effort by a community to claim its right, to speak the language, and set associated goals in response to community needs and perspectives." More specifically, language reclamation includes what is typically associated with language revitalization, "but very much within a frame of decolonization." A sleeping language may not be actively known but exist in documentation and/or claimed by a heritage community. Therefore, an archive associated with a sleeping language can be critical for the community that claims the language.

Wesley criticizes a top-down approach to language, for example, if viewed through "the categories of linguistics." He prefers a bottom-up model, citing L. Frank Manriquez as an example of such an approach, she says, "language is a basket that holds all our culture…" among other things. "Archives support…but can also work against reclamation if they're organized or operated in ways that demote indigenous worldviews."

Some of the common challenges he sees are:
- "Mismatches between categories that guide archiving practices and those that guide reclamation efforts"
- Interpretation of archival knowledge
- Issues of access: including economic challenges and restrictions

- People who just "don't get it" and may mistrust Native community members, although this is not the norm in his own experience

Something Wesley would like to have happen in language archives is "to be invited to share [his] definitions of language and related perspectives up front." He concludes with another challenge "language archives of *tremendous* importance to Native communities, and they must be better funded." Citing high fees for users and bad logs as just a few of the issues that spawn from a lack in funding, he points out the 1990 Native American Languages Act as a way for communities to gain additional funding. The legislation aims to "preserve, protect, and promote Native American languages," so instead of focusing on one specific aspect of language reclamation work, this policy should actually serve as a larger umbrella that could encompass all sorts of reclamation efforts.

[31:18] – **Michael Shepard**

Michael is a cultural anthropologist that has worked with several language communities but he is not a member of a language community. He shares a story about a trip he took with Lummi elders to visit the archive of a Native Hawaiian language community. The Lummi elders kept asking questions about access and use of archival materials as well as educational materials. Michael says the Native Hawaiians suggested that free and open access to materials was necessary for learning, which "was a really challenging statement" for the Lummi elders. Since that experience, Michael has been examining the tension that exists between the interests of language communities to achieve goals in preservation and education, and their "core values in political and cultural arenas."

Through participatory research using Mukurtu CMS, Michael has observed how people use language archive platforms and how they "conceptualize their linguistic heritage in the linguistic sphere." Protocol, sovereignty, and protection of treaty-based rights are all facets that may affect how "heritage materials are documented, accessed, and preserved." Another concern, of the communities Michael has worked with, is how existing materials are mobilized for educational purposes.

When viewing the archive as a "reservoir" Michael wonders how an archive can assist in other processes, "whether that's education, land claims, or other efforts to affirm sovereignty and treaty rights." He also asks, "how are we engaging in ongoing documentation with our archives?" In his dissertation, Michael interviewed managers at some major archives and also tested Mukurtu CMS with some language community members.

Some of the challenges Michael has encountered are:
- Protocol
- Misuse of cultural knowledge
- Cost
- Personnel and technical support of archives
- Outsourcing vs. centralization of services

What Michael would like to see more of:
- Greater facilitation of educational materials
- Greater contextualization of archival materials
- More technology use such as mobile phones, tablets, and QR codes (locative media)

4

Some community members have also expressed to Michael they would like to be more on the back-end of managing the archives so they can spend more time teaching.

Linguists

[39:50] – **Andrea Berez-Kroeker**

Andrea is a professor in the Department of Linguistics at the University of Hawai'i at Manoa as well as the Director of Kaipuleohone, a language archive founded in 2008. Since her start there in 2011, it has grown to about 35 collections. Andrea also teaches a course on community archiving at CILLDI, the Canadian Indigenous Languages and Literacy Development Institute. She is also active in DELAMAN. Andrea believes she fulfills roles as both a depositor and as a researcher. As a depositor Andrea is speaking to work she has collected personally. As a researcher she is speaking to her role to write scholarly pieces about language.

She uses a language archive, primarily, as a depositor, to preserve her own fieldwork materials and have a corpus of data for others. However as a researcher, "I must admit that I don't do that much looking at archives to get other people's data."

Again Andrea breaks down aspects of an archive that she would like to know about. Therefore, as a depositor she would like to know:
- Archive policies to understand how protected her materials are
- The archive's long-term plans to facilitate trust with them knowing how they will handle materials in the future
- How the archive handles requests for access
- How straightforward metadata creation and editing is

Interestingly, Andrea does not access her own materials via the archive, rather she just uses her own computer. However, her students do use the archive as the primary way to access their own deposited materials.

As a researcher she would like to know:
- What's available? Is it media or transcripts? Basic search functions
- Reliability information

Problems she has encountered as a depositor are:
- "opaque" procedures for how to submit metadata
- Poor communication in the "black box period" which is between when you submit materials and when you hear back that the materials are recorded properly in the archive

Problems she has encountered as a researcher are:
- A lack of context surrounding the whereabouts of an archived recording or item

Andrea shares how Kaipuleohone, the archive for which she serves as director, is well-endowed to move forward in terms of staff, plans, and preservation. However "the front end is terrible. I have no real control over what kind of information gets displayed, on how things can be searched. It's really geared towards traditional library publications, not media." Another issue is changing access conditions over time.

Therefore, Andrea would like more automation of services, tracking features, and visualizations of data. She would also like "better integrated citation tracking" and easier access. Perhaps through a "smart" program that can pull from an article what it is referencing in the archive so you are automatically linked to the items you're reading about.

[47:18] – **Justin Spence**

Justin works with California Native communities to create recordings and will interact with archives as a depositor, however he wants to focus on "using archival materials for more traditional linguistic research." He wants to discuss whether or not the field of linguistics is "on board with doing linguistic documentation." He feels as though the case must be made that archival materials are important to things linguists do, even when the reasons for documenting a particular language is mainly for language revitalization. However, "culturally-related questions" do not come up often in linguistics articles.

Justin introduces documentary linguistics as the solution to such issues. In documentary linguistics, languages may be documented for one purpose, but research questions can formulate completely separate from the original intent of documentation, and long after the fact. He admits that "many of the archives I have worked with have found it very challenging to probe existing archives." Older materials can be "very haphazard" but the creation of new materials can be keep the goals of documentary linguistics in mind to be "much more systematic."

He has used various archives in his dissertation work, but interestingly he mentions archive.org as a good source of information. Justin has also worked with languages in which the documented linguistic record is all that exists. However the majority of his work has been done with "Hupa [Whilkut] and other languages in the Dene basket that do have at least some materials that are available."

Commenting on abstraction, Justin says, "sometimes as a researcher I am interested in language variation and change, and you do actually want that information [referring to fine-grained phonetic descriptions] to go back to the archive and try to understand some aspects of variation that can be recovered." Justin also notes that what goes into a linguistics publication is not all of the relevant information, for instance, names of speakers and additional comments are often left out. He shows some fieldnotes as an example to "confirm and expand, get more information about a parameter of linguistic variation or dialect variation that is recorded [and?] published." In another example he demonstrates how he wants to understand "what people's actual lived experiences were of the languages that they knew or were interacting with…when a lot of these processes of language shift were in motion."

Finding a piece of "sociolinguistic contextualization" is often "fortuitous" or serendipitous, not commonplace. Other types of "sociolinguistic contextualization" Justin would like to have are:
- Names (as stated above)
- Dates
- Locations
- Gender of speaker
- What questions were asked as part of the elicitation?

Justin would like the ability to "find things" without having to go through "a million pages of fieldnotes." Similarly, he would like to be able to go to a specific point in a recording to find

exactly what he wanted rather than listen to the whole thing. He explains the term "de-archiving" as "a very searchable representation of archival materials." Justin wonders if this is the responsibility of the archive or researchers, "but I do think having this sort of access to very particular sets of information in archives is something that we need, and maybe a pipe dream."

Other problems Justin has encountered are:
- Distance from where he lives and where archives are
- Cost of obtaining materials
- Inadequate metadata
- Multiple transcription systems – "having to process many different writing systems"

What Justin would like to see:
- More fine-grained searches
- "One-stop shopping" in which there is a landing place that describes where every piece of documentation exists for a particular language so you know where to find it

Justin also shares that gets folks asking him about recordings of their elders, "what should I do with them?" and the answer he gives is "I don't know." He laments the options that are available, "they're not good" especially if the individual doesn't trust their own tribal government or public archives to be "good stewards of the materials." In the end, the materials sit in these peoples' attics, and Justin needs help to tell them what to do with them.

Archivists

[1:01:30] – **Mandana Seyfeddinipur**

Mandana works at ELAR and acknowledges Sophi Salffner, the full-time digital archivist at ELAR. Mandana explains that ELAR has a website and anyone can browse the collections, "but if you want to see anything or use anything you need to register." She says this gives them a little control and helps with avoiding spammers. ELAR contains a system with four levels of access: everyone who is logged in, researchers, community members, and upon request. She says most of the restricted collections belong to Native American and Australian Aboriginal communities, and "user statistics of those, which are restricted to community members, is actually zero, which is heartbreaking." This makes it difficult to justify keeping such collections online.

As a depositor to ELAR, there exists the possibility "to upload files yourself" as part of a new archiving system. Unfortunately the process is "quite complex" and the interface "is so unfriendly you want to cry when you see them." As a depositor you can also set varying aspects of access and track downloads.

Issues users encounter when using ELAR:
- If ELAR funds a project then the recipient of funds must also archive with ELAR. However the interface is in English only and internet access can be a problem because they fund projects all over the world, sometimes in remote areas.
- The metadata is "really scarce and limited" even though they "force depositors to give us a human readable description" of materials.
- "Contextual information is often lacking to understand the collection better."

- Information organization leaves much to be desired, as do browsing options. For instance, if you want to find dictionaries they are not labeled as "dictionaries" which means you cannot find them.
- A depositor may not respond to requests to access their material because they have passed away, dropped out of the field, has switched languages, has changed location, etc. Even so, if ELAR cannot contact the depositor or the delegate of the collection, they have a "sunset clause" to open the collection. Some collections are "locked down" because there is no licensing agreement and there is nothing left to do besides go to the language community and ask someone for assistance.
- Incomplete records, for example, loads of recordings but not transcriptions.

Mandana does admit a lot of these issues stem from items deposited in the past, "they are very hard to access in terms of *finding* things." Now they have a metadata editing tool which makes things easier.

Mandana admits the interface is not sufficient and "we do not have the funds to develop a friendly user interface." She shares that ELAR has a broad set of users. Through questions asked upon registering for an account, they have learned how users are connected to endangered languages. For example, a person may work with people who speak an endangered language or sometimes a person finds a recording of a passed relative speaking in an endangered language. Mandana says there are also artists who would like to use the materials for artistic projects. One of the issues artists encounter is the excitement of finding something for their work followed by the disappointment when they realize they are "locked out" and cannot gain access to the item. Instead of the user in this instance understanding why they are locked out, they think of it as an error on their part, as in, they cannot correctly find the item.

Some questions users ask are:
- What is X?
- Where do I find X?
- What does the hold on language Y?
- What does the archive hold on oral histories?
- Where are the recordings of speakers X and Y?
- Where are the recordings where they speak about the old days?
- Which deposits contain languages that are most endangered?
- Which deposits have video images and teaching materials?

In regards to collections, Mandana says they don't know "how much of that material is transcribed and translated."

Kinds of materials and formats users look for at ELAR:
- Students want dictionaries, transcriptions, translations, video, and audiorecordings
- Community members want to see their own collections, recordings of family members, and materials for language maintenance and revitalization. "They want to see if there's a recording they can use for something."
- Artists want to see "aesthetically pleasing materials," materials for background research, storytelling, and outreach.
- The public wants general information, songs
- Journalists want to know "what's the most endangered language," recordings of a last speaker, or background information for stories.
- Computational linguists want to compare language families. However Mandana points out they don't stick around because they don't have access or "because materials are

not provided in any sort of searchable or usable form for any kind of quantitative approach or any type of systematic approach."

Mandana says they are always thinking about how to make their archive more accessible to varied users but the "long-term sustainability" of the archive is always a priority. Right now ELAR is part of the library at the School of Oriental and African Studies (SOAS) University of London, but "the collections are not even used in teaching at the university itself." There is no cross-over between the professors who teach language documentation and the actual archive. Mandana would like to see students use the archive so they can learn early on how important rich metadata is to a successful, usable collection. In turn, Mandana could provide information from users about what they want students to collet for language revitalization projects. "We could learn from each other" but working in silos prevents this. "We are colonizing everyone away." Keeping up with the digital age is difficult especially if linguists are not trained in creating digital collections.

Other things Mandana would like to see:
- Provide a user-friendly interface that is not just in English
- Connecting the metadata in a parallel multi-lingual thesaurus
- Virtual guide tours on topic areas or families or "off the shelf teaching sets" – here Mandana comments that many linguists have collected data to "give back" to the language community for revitalization efforts, but they are often not used or their existence is unknown.
- More funding for the efforts outlined above
- Mandana has also thought about crowdsourcing for resources. For example, holding a "tagathon" so people can generate richer metadata.

[1:18:10] – **Felix Rau**

Felix is a representative of the group at University of Cologne that is going to reconstruct the DoBeS archive. "We are just starting with the construction of the archive and it has been a side project for the past three years, but we got proper funding in October, so we are just in the phase where the answer to all the questions is, 'I wish I would know' and our archives are in a very specific context. Basically, in the context of the demise of the Nijmegen archive, right? The Nijmegen archive is being scaled down, the group that runs that, the software is in a transition phase. And we are starting basically in a new, in a context with very concrete assumptions about what an archive has to be…and I think we have to rethink a lot of things, but it's been a very difficult context because there is this urgency of we have to set up an archive within the next three years so that there is something productively running in three years, but also we want to rethink a lot of things. The Nijmegen archive is the archive, well the main part is the language documentation data from the DoBeS project, but they have-Mandana knows actually better perhaps."

Mandana: "in 2000 the Volkswagen foundation started a big granting program for language documentation projects and also supported the Max Planck to develop a digital infrastructure for language documentation." She says these were all large grants ($350k - $500k) so "you basically had the crème de la crème of language documentation" and big groups worked together with technologists to develop the digital archive. There was a cycle, twice a year a researcher would go to Nijmegan and work with the beta-tester groups and the developers.

Audio 2 [1:18:00]
Mandana: The result of having such sophistication means it was tailor-made for linguists but "the interface is a nightmare." Due to an end of funding, "the whole thing is now moving into a data center" and they are trying to salvage what they can.

Felix continues by saying this is a good "test" to see what can happen when an archive is "dominated by the people who programmed it." Most of the design decisions that were made came from "the view of the producers." For example, when displaying something the system focuses on how to display the project rather than display the data for a "consumer." He shows an example of what the interface used to look like and what it looks like now. He says they are now in the process of finding a set up for their computer center to work with other entities for community-based, open-source solutions. In this period they are trying to "understand more of what we don't know than what we know."

A majority of users are depositors, but they also use the archive to teach graduate students how to deposit. "It's bordering to the ridiculous sometimes how the whole thing is focused on the producer side." They are also working with a group of linguists who are building a cross-language corpus. He says this has been helpful because the perspective is different and "curated in a form that you can compare" various languages. Therefore a main issue is just how to deal with the software. Another challenge is working with so many stakeholders (funders, departments, depositors) that have an influence on how things are designed when they have differing interests.

Audio 2 [1:28:10] – **Susan Kung**

Susan is the sole employee at the Archive for Indigenous Languages of Latin America (AILLA) "which is one of the earliest digital archives." Much of their material was not "born digital" so they started out as a "digitization service" to make recordings and other materials accessible online. The founder of AILLA was Joel Sherzer, an anthropologist, who "decided there needed to be a clearinghouse" to house and share materials. AILLA, then, is a joint effort between the Department of Anthropology at UT Austin and UT Libraries. "From the get-go we had this promise of sustainability" because even if AILLA were to cease existing, the materials collected up to that point would be housed within UT libraries.

Susan brings up "post-custodial archiving" which is when you do not take material out of its birth community and just digitize the material. She says this is what AILLA has done, being 100% digital, but the "host community" is often an anthropologist or linguist, not the actual language community. AILLA has portals in English and Spanish. People at AILLA will work to translate any metadata they receive in Spanish to English and vice versa. When material comes from Brazil it is often in English, but they will ensure Portuguese fields are also included. "Anybody who has access to the internet can access AILLA [for free] and read the metadata…however if they want to access any of the files themselves they have to create an account and agree to AILLA's terms of conditions and use."

Problems users encounter are:
- Lack of keyword search – although there are collections organized by language or country.
- Pre-set filters do not include the names of people.
- Some materials have access restrictions – much material was collected prior to Internet, so public accessibility/restrictions were not built in.
- Technical problems – when creating an account for instance, people who have accents in their names may have the account rejected or sometimes they have firewall issues or broken links.

Questions users ask:
- Why can't I download files? – users do not understand that materials are restricted to them instead of it being a technical issue.
- How can I get permission to access restricted files? – often they cannot.
- Users forget their passwords/usernames – and there is no automated system to help with that.
- Questions about how to search for something.
- Why can't I upload files and metadata myself? – they are planning to create a self-depositing system. However UT Austin has a very restricted security system, which is why a self-deposit tool has yet to be created.

Susan defines the users of AILLA as:
- Academics doing research on a particular language
- Students interested in learning something about a language – these can be learners of the actual language or students who need to do research
- Heritage speakers trying to learn the language or use it for language revitalization projects
- Language community members who want to hear their language
- Depositors
- Some artists

Materials users are looking for are generally fieldnotes in a particular language, audiorecordings, videorecordings, or other text.

Susan says she is interested to learn more about how UCD is done and has some ideas to implement changes to AILLA, but currently she is stretched for time. Currently AILLA is going through a migration, they're using Fedora with Islandora on top. She is hopeful it will solve various technical issues.

Audio 2 [1:38:55] – **Jennifer O'Neal**

Jennifer is the University Historian and Archivist at the University of Oregon. She says her archive is much different from the others presented because there is no "core language." However they have two projects and collections related to endangered languages in Oregon/Northwest. She also says their collections "are not digitized or online" but they're working on it.

Jennifer says the finding aids to collections are online but users have to physically come to the repository to view the collections. This was a purposeful decision to provide a "barrier" and

make it more difficult for outsiders, those not from the tribal community, to access materials. One of the projects she is referring to is the Southwest Oregon Research Project, which is an interdisciplinary collaborative of people who are contacting archives and repositories in Washington D.C. to collect language materials that belong to Oregon's tribal communities. Another group they're working with is the Northwest Indian Language Group. Although their collection is not transferred to her organization, Jennifer is providing assistance and helping them use Mukurtu to manage their own collections.

In some cases a collection is digitized but not online. Sometimes that is because they are still testing Mukurtu "to see if that's the best software system to use." In general there are no "problems" since collections are physical, "but that is the problem! They're just physical collections." Therefore their current goal is to make the materials digitally available to the communities with appropriate access restrictions and protocols as dictated by tribal communities.

Questions users have are:
- Who collected these items?
- Why are they here?
- Who initially created them and why were they brought back?
- Why are they in a repository?
- Who has access to the documents?

When these questions are asked by researchers, Jennifer sees it as a learning opportunity "because we provide the context" of how materials were collected by an outsider (such as an anthropologist), removed from their original language community, and are now brought back through their efforts. She can also discuss access restrictions at this time with the researcher.

Due to the focus of collections about the Northwest, Jennifer says they often get cultural history writers, anthropologists, and ethnographers. The formats they collect are in all formats: "paper, records, manuscripts, photographs, recordings."

Due to the nature of her position and the organization in which she works, Jennifer highlights the following concerns:
- Issues can be present in all collections, not just language archives.
- How do you prioritize the work to be done when many collections require work?
- Is there a dedicated curator?
- How do you find resources to dedicate to collections that need a lot of work?
- Issues of funding/coordinating
- Convincing individuals in positions of authority why it's important to serve tribal communities who need assistance for grants in language documentation.
- The infrastructure to create systems and manage them takes a lot of resources. Jennifer often has to argue with the higher-ups about why another system is necessary.

Lastly, Jennifer ponders her role as an archivist and also as a member of her own tribal community. Both are very important to her, especially ensuring she meets the needs of tribal communities.

Andrea: have you had to turn away people who were requesting letters of support for the grant?

Jennifer: we had two come in just about a month ago from two different tribal communities. It was difficult because they came in at the same time…and we had just been asked to be a partner for a grant.

Jennifer shares how the Dean questioned the situation "why do we have to be the ones doing this?" and "this is taking away from University work." She says it's "a huge challenge for us" but she appreciates the opportunity to discuss the importance of this type of work. She feels that most people don't work with communities and therefore don't understand why a user-friendly system is necessary instead of a more technical, back-end system like Hydra.

Audio 2 [1:50:44] – **Alexander Wadsworth**

Alex works at FirstVoices.com, which is a project run by the First People's Cultural Council. They serve the BC First Nations, which is 203 First nations with 34 languages and 61 dialects. The organization is governed by a board of 34 First nations, which make up a sort of advisory committee. They also have a Board with 10 representatives from the First Nations and three from governmental organizations.

Alex says 60% of the languages in Canada are located in BC. They officially launched FIrstVoices.com in 2003. It was created by the local tribal school, and it grew into what it is today. The reason for creating FirstVoices was initially for funding language projects of the 203 First Nations. Instead of spending money on providing tapes, it seemed like a better idea to have a place online to store materials. The site also includes some languages from Australia and the US.

Users access FirstVoices through a community portal, where they can go into the language and listen to some quick facts about the language and people. There are also games, phrases, songs, and stories. They incorporate language learning with the archives by populating games with words communities choose to archive. FirstVoices was approached by Health Canada to create a free reading site for learners. The community chooses what content they want to share, and it is now developed into FirstVoices Kids. In one example, Alex shows an image where children can click on colors and listen to the language. They also have a dictionary app for languages who have chosen to do so.

Other resources on FirstVoices.com include:
- Grammar and punctuation
- Language tutor – teachers can create lessons using the data that's published
- Language Lab – a local server that will talk to some iPads and allow learners to access the language, even if there's no internet.

Start Video 3

Audio 2 continues on

- Chat app – developed out of a request from community members. It supports 150 languages within Canada, the US, and Australia. It allows for Facebook and Google chat, it is also free. Currently the app is in re-development to include iOS and Android keyboards.

13

Originally the site was created to support dial-up users, but Alex says they're in the process of re-writing. He says they're having some issues with Flash. The system is currently server-based, but they are moving it to the cloud and plan to open-source it to allow for more collaboration. They're also looking into other ways to present information, such as radio.

Questions users ask usually come in the form of email:
- How do you say word X in [insert language]?
- How can I type this in my language?
- Can I use this on my mobile device?

Alex will direct users to a separate contact for answers and tech support. Common types of media users are looking for include the words, phrases, apps, audio, lessons, and the keyboards.

End Audio 2 [2:00:00]

End Video 3 [03:50]

User-Centered Design Experts

Start Video 4

[00:23] – **Crysta Metcalf**

Crysta is a design anthropologist who worked for 13 years at Motorola using anthropological methods to help inform new product innovation and design. During that time she worked with people in many different occupations who use various technologies in software design. Sometimes the focus was to design a new system or make revisions to new ones based on user needs, but she was always working collaboratively on multi-disciplinary teams. She says she loves working on teams with folks who have different perspectives because it helps create a better end result.

Crysta shares that she has undergone several challenges when making technologies easier for people to use. However, for this particular project, she foresees the biggest challenge being the diversity of goals for different user groups. The workshop thus far has been eye-opening, for instance, she wouldn't have imagined artists using language archives. Maybe you have different user interfaces for the different purposes, but who knows? Crysta says the focus on "users" in user-centered design is the key to success. "We don't think about what we want to use, we don't focus on one specific population, we think about all of them and we talk to people in what we like to refer to as a dialectical design process." Crysta says creating prototypes are easy and inexpensive – even a drawing can be incredibly valuable and cost-efficient.

Crysta images the process of UCD with language archives might look like:
- Understand the process by talking to people
- Do some contextual inquiry – watch people use the tool and ask questions while they use it
- Create some designs
- Circle back and iterate

"Nobody knows everything, so I'm always open to new ideas. That's how I normally think of user-centered design."

[06:05] – **Santosh Basapur**

Santosh is a faculty member at the Illinois Institute of Design which has been a design school since the 1940s in Chicago. He works there as a user-centered design expert, mostly system design. His work experience has been mostly corporate research. He was a mechanical engineer interested in human factors and began to apply ethnographic methods. Some of the projects he has worked on include voice interfaces in cars for hands-free calling, "it's always the future of something!"

Santosh explains the usual process of UCD:
- Research in the context
- Analyze the data of all kinds and understand insights
- Synthesize solutions
- Develop concepts and prototypes – this can be a physical product

He provides an example of how he has conducted UCD research. They took off-the-shelf products and put them in different living rooms. They also put cameras in people's living rooms because "people don't talk if you put them in artificial situations…we can do different things to get people moving along and testing different concepts." This project resulted in the Lab's Social TV concept, "not just communication, but watching TV and socializing around it." Something interesting they found was people surfed the web while watching TV to find answer to questions, for instance, where was a particular movie filmed? "All different kinds of concepts start coming out when you do research which is situated in a user's practice of how they observe things and what they do with those things."

Santosh explains the importance of storytelling in a project that aimed to bring together generations that have grown apart over time. They did a study with people who had retired in Florida with young relatives living in Chicago. They realized the project could also facilitate discovery of the city, "it was really good to start something small but then realize so many things can happen."

Santosh describes some common challenges in UCD work, but also within the context of the workshop. For instance, designing a user experience can include several hiccups and you have to go back because some concepts don't work. Sometimes there is a lack of empathy for the user, and then distributing that empathy among all of the stakeholders can be a challenge. In discussions that have already taken place during the workshop, Santosh has seen linguists feeling sorry for archivists and archivists feeling sorry for language community members and language community members feeling sorry for the technical people.

[11:15]
"Everybody has to have that empathy and we are trying- as a designer your main challenge is to bring everyone onto the same (3 syllables) bridges… Usually design gets a (reputation) for new ideas and fresh ideas, but actually what we usually do is bring people on the same page and then you do something with it. By the time they all come onto the same page they bring their own ideas and all the ideas start compiling on top of that and say, what will really make

sense. So it's developing empathy and then having the empathy distributed among all the stakeholders is very important."

He also cautions that user research can get "bogged down" because it's a lot of research and not everything is insightful. A lesson he has learned from working on multiple projects is to allow the data to accumulate, slow and steady, to find actionable insights, "then you know how to fix it."

Santosh shares a joke he usually tells his students. It's a comic that provides a series of pictures, each representing slight variations in a tire swing. The first picture asks "what the user asks for" and the following picture shows additions made by the programmer, and then there's a picture with additions by other team members, and at the end there's a picture with what the user really wanted (a simple tire swing), next to a picture of what the user actually received (a very complicated tire swing).

Coming into this project, Santosh says he is a "clean slate" because he knows nothing about language archives. This is often the case when assisting in the design of new technologies. He shares a story about how he donned scrubs to observe surgical procedures with the end goal of designing for medical operations. He terms this activity "absorption, where you get into a situation, absorb what's going on and then get the [person's] input." Sometime absorption can reveal previously hidden aspects of a design problem not considered by the client. In the case of the medical operations, it was the nurse in the room who was tasked with a great amount of responsibilities and who needed help to make her job run more smoothly. "It's always a learning process."

A current project Santosh is working on is a bike-sharing system in Chicago. There are bikes stationed on the south-side of Chicago but no one is using them, so they are trying to figure out why by immersing themselves there. One thing they have learned is that it's "extremely uncool to ride a bike" on the south-side of Chicago. In regards to the language archives project Santosh says "it's a new challenge, we welcome it."

Representatives of Funding Agencies

[18:17] – **Shobhana Chelliah**

Shobhana was at the National Science Foundation serving as the program officer for the Documenting Endangered Languages Program from 2012-2015. The process for receiving proposals begins in September when they are received, and then they send them out for review. Once they come back a panel goes over the reviews and offers their own review of the proposals. Shobhana noticed in 2012 that many of the proposals coming in asked for a re-up of funds because they had finished Phase I and wanted to go forward with Phase II. Reviewers and panelists asked, "how do we know what's happened with Phase I?" She says the point of the DEL program is for people to create this resource for science and humanitarian purposes. Therefore, when people ask for more money, the look closely at whether or not that mission has been fulfilled. In many cases materials were archived at DoBeS or ELAR, but they couldn't find the materials. As reviewers they would ask, "where are they?" and often they had to go through hoops to find the material. "It's not possible for reviewers to do that kind of exploring."

"After I realized that materials weren't really being deposited, I started talking to these folks." Shobhana is pointing at Susan from AILLA and Gary from ANLA. "We really didn't know what was being archived, how much was being archived, and what quality those deposits were at." Therefore, the DEL program put in place a stated requirement, "anybody who got money from DEL also would provide a plan for putting their stuff in an archive. And they also had to budget time and money to accomplish those tasks." This is part of the specific Data Management Plan (DMP) that DEL requires – which is more extensive than what is required by the NSF. Researchers have to contact an archivist and get clearance from that archivist for the amount of information they're going to be putting in there and that letter must be provided to reviewers and the account.

Loriene: "so that's why Jennifer gets all those-"

Shobhana says "yes!" She says there are intended and unintended/unknown consequences of the DMP. She says the intended consequence is people will be forced to realize that they are only contributing to science and humanity if materials are shared and put into a place that is useful. The unintended consequences can include the stressors on smaller archives, "even three proposals is too many," and then they are forced to choose among proposals. Or on universities (like her own) that can take in the material but does not yet have an infrastructure.

Shobhana envisions some solutions to these challenges. For instance, NSF people should come and do an outreach event to inform smaller universities and archives to explain how much NSF can contribute to their costs. For instance, Susan Kung has started to ask for 8% in direct costs from people who want to put their collections in AILLA. "So that there would be some sort of payback to the archives and perhaps helping with infrastructure building through NSF."

[24:43]
Susan: "yeah, but I think there's not an archivist here who has seen any returns from all of our work yet."

Mandana: "yes."

Susan: "you have gotten money from someone already?"

Mandana: "yes."

Susan: "because I haven't and I've written more DMP support letters than anybody in the room."

Shobhana: "did people get funded?"

Susan: "well how am I supposed to know if they don't write to me and tell me."

Shobhana: "okay, well we should add this on then."

Susan: "and I'm specifying in my letter that they're supposed to tell me."

Shobhana: "and if they don't get funded they should probably tell you, too. Which is something you can put into your letter. Put that into your requirement that they need to tell you, as soon as they hear, yay or nay. And then you can maybe set up a schedule for getting your money. Because you could be waiting for five years to get the money, and then the grant is over, and you're still waiting for the money because they're not archiving it."

Mandana: it's important to connect the two (those applying for funds and the funders) because stuff gets lost. She is concerned about the lack of a central archive in the US to deposit materials that are funded by the NSF.

Shobhana brings up the sustainability of archives in regards to funding. Currently the US congress is funding projects in relation to national security, health, and economics. "We have to be very careful to try and get money in a way that says, 'we're doing something that's useful'…that feeds into their notion of what is national security."

She offers a few ideas with the disclaimer that they may sound "cold" because they ignore the humanitarian dimension; they build on what the NSF funds, which is "science."
[27:00] At a general level, you could say you are developing ideas on how language works, and that contributes to our understanding of all languages, including languages of interest to national security

Specific to language archives:
- You can say you will build computationally sophisticated, annotated archives that facilitate cross-language comparisons. There are several NSF programs interested in funding that kind of archive
- You can go to the Computer Science and Engineering Directorate and say you will develop computational tools for making comparisons across languages that can then be used for other kinds of computer-assisted language analyses
- There is a special program within the Social, Behavioral and Economic Science Directorate called RIDIR, Resource implementations for Data Intensive Research. It's used by political scientists and sociologists, cause those guys have huge data sets, but linguists have been encouraged to participate. Shobhana would love to develop something for South Asia, where they would include data from both larger and smaller languages.

Mandana: "the problem is that's not sustainable, this is not something that will save the archive. This is something that gives you money for a certain amount of time." She references the soft money that funded DoBeS over the past 15 years. When that is gone you lose the sustainability of the structure. They build this structure on a particular kind of funding model.

Shobhana says "so we want to pull it out from just that soft money to something a little more sustainable and so I want to suggest that maybe tying in an educational component…let's talk about Native American communities and the educational component and what kinds of monies are out there for that and tie that into archiving as well." At UNT they are trying to create a person who can serve as a content specialist who can do archiving but also understands technical aspects. She posits the idea that training such a "new breed of student" would be possible if funding could be obtained from the educational side of NSF.

Shobhana outlines a few other opportunities that intersect with the education part: TCUP Program at NSF: Tribal Colleges and Universities Program. The Director of that program is very interested in language and linguistics. Shobhana says we should approach her and talk about the hubs and spokes model. The Director is also interested in funding the education of tribal university students. Shobhana thinks it would be nice to bring these students into non-tribal universities with archives for training in methods to then send them back out into their communities and develop small archives.

Loriene makes a comment how one of her classes right now is working on using a suite of resources for tribal college students by delivering them to tribal college librarians. They were developed in accordance with an Indigenous model of cultural fulfillment.

End Video 4 [33:18]

Start Audio 3 – start at [33:10] to catch the rest of Shobhana's talk

Jennifer: says there are also a few other programs that are working with tribal colleges. "There are some thing already in place that if you wanted to do that you can work off of."

Shobhana says she doesn't know if she would be the right person to do this. "I guess what I'm saying is I understand that having soft money fund-"

Mandana: "it's nice! But it's very dangerous." She says that language archives, at one point, were popping up all over the place and "since 2012, 2013, all of them are dying down."

Shobhana: "they're reforming in different ways."

Gary: "I also think we have to be careful when we think about NSF and soft money. It is soft money, but within the US, at least, this is the way we fund science…this *is* the model." He says even though it's soft money it is sustainable.

Mandana: from her outsider's perspective she doesn't understand why the Smithsonian doesn't have a central archive for all materials, specifically to help small communities. She says it's "a no brainer."

Jennifer: she says it's great the NSF is requiring those who receive funds to indicate where and how they will deposit materials, but "many of these tribal communities do not have that" and if they're requiring it "they need to come up with a solution." There is nothing central and these communities just don't have the funds.

Shobhana: "we can now say to NSF, we really want to follow through on this DMP, but we need you then to help us support, give us the training, give us the students…"

Jennifer: she says all of her colleagues in the Northwest are dealing with the same issue. Communities are coming to their organizations (embedded within universities) for assistance. She says they have to "fight" to get assistance, because it's not the highest priority for their organization.

Shobhana: says she would like to work with Jennifer on that.

Gary: provides an apology on behalf of Shobhana's successor, Colleen Fitzgerald, "who wanted to be at this meeting, but wanted to be able to come." She will be very interested in these discussions and this will go back to her.

Christina: she is also concerned about money and says this is part of the group conversation, "we user-centered design folks are aware that user-centered design is expensive, right? It involves more resources; it involves more people. You end up with a better design but it's not free because people are putting time into it, so this has been one of my questions, too…I want

us to think about funding sources. So I think the funding question is, in fact, completely relevant."

Jennifer: says it's like dealing with two separate issues. A funding issue for both preservation/storage of materials and also creating a user-centered design for access.

Susan: comments that "nowadays everything for NSF has to be innovative." Preservation and sustainability "falls on us" (the archives and institutions), to figure out how to sustain and preserve. "When we talk about the innovation, the idea you mentioned with annotation, that can only work for a subset of materials in the archive. Then that becomes a separate project within the archive that also has to be sustainable." She says they received money to do it, but not to sustain it. "The sustainability is the number one issue that is facing all digital scholarship." Others agree that it affects not just language archives.

Shobhana's presentation ends at [40:30] on Audio 3