INDEX

PORTFOLIO

BAILEY E BERRY
MLIS, Spring 2021
UCLA, Graduate School of Education and Information Studies
ADVISOR: Anne Gilliland, Associate Dean of Information Studies

ABOUT

This pdf contains the contents of my MLIS portfolio, a reflection of work completed and my professional development during the UCLA Masters of Library and Information Studies program. You can use the links in the Index to navigate to different sections of the portfolio. A link back to the Index is available at the bottom of each page.

The Professional Development section gives context for my decision to enroll in the MLIS program at UCLA and a description of how my professional skills have sharpened, interests have developed, and goals have become more focused. The Issue Paper explores a contemporary issue in the Information Studies field that is relevant to my interests and career goals. Included in the portfolio are examples of coursework. The Major Paper was written for the Archival Description and Access elective course, which is a course on the archives track taught by a ladder faculty member and accounted for 40% of my final grade in the course. I have included two projects completed as Elective Work and one example of Core Work completed during the Systems and Infrastructures core course. Also included in the portfolio is a full list of Courses Taken, a record of my Advising History, a CV, and an Accessibility Statement. An accessible html version is available online at https://baillizard.github.io/baileyeberry/

GO TO INDEX

PROFESSIONAL DEVELOPMENT

BACKGROUND AND OVERVIEW

When I decided to major in art history and minor in Studio Art, I did not know what I wanted my career to be. Cataloging and processing collections after graduation revealed an opportunity to channel my love of art history into a meaningful career. My first exposure to archives was as an undergraduate student intern with the Patton Archive Project, where I entered metadata about a series of personal diaries into the archive's online database. After graduation, I was the Principal Cataloger of 17th Century artifacts collected at the Smithsonian Environmental Archeology Lab in Edgewater, Maryland. After moving to Los Angeles, I helped manage collections at Royale Projects Contemporary Art Gallery as the Lead Intern. It was my experience inventorying and rehousing a private family collection belonging to the Franklin K. Lane Estate that convinced me to pursue a graduate degree in Library and Information Studies so that I could refine my practice and expand my skillset.

My MLIS program at UCLA has focused my goals even further. I want help lead the movement to provide better access to collections. I have spent time investigating how metadata and its presentation can reveal and communicate complex narratives about cultural heritage in collections. I am developing skills in data management and digital tools because I am interested in how access to collection data enables digital humanities projects, how humanities researchers' data can contribute to collections, and the digital projects that emerge out of the interaction between the two. I explore this idea further in my Major paper. Click here to go to that project in my portfolio.

METADATA AND ACCESS

When my summer 2020 internship with Sequoia and Kings Canyon National Park became a remote position, I needed to quickly shift the focus of my work. Using the final assignment for the spring Metadata course, I designed a metadata enhancement project for SEKI digital collections being aggregated with Calisphere and the California Consortium of Herbaria. Click here to go to that project in my portfolio. During my internship, I facilitated the harvesting of three digital collections of historic images onto Calisphere. Using OpenRefine, I cleaned legacy herbology data collected by scientists at SEKI to prepare them for publication on the California Consortium of Herbaria website. I collaborated with Ward Eldredge, the curator of collections, SEKI herbology scientists and data repository staff, and the liaison from CCH to write suggestions for implementing the new system into SEKI data workflows.

This project reinforced the importance of metadata in providing access to online collections. I continued to consider this connection during an inter-disciplinary course discussing African Objects in Museums. Building off questions regarding equitable access to museums and transparency in museum practice raised in a Museums in the Digital Age course during fall quarter, I decided to create a digital exhibit using ArcGIS Story Maps which clarified provenance metadata for a Kongo power figure in the Wellcome Collection at the Fowler Museum. Click here to go to that project in my portfolio.

GO TO INDEX

HUMANITIES DATA MANAGEMENT

During a Digital Preservation course taken winter 2019, I was part of a group project that worked with a PhD researcher in the Cesar E Chavez Department of Chicana/o Central American Studies who needed to manage a growing collection of digital and digitized Latinx Punk materials. In researching guidelines about managing humanities research data, I found that there were relatively few resources compared to the scientific community. An interview with Maureen Russell, the archivist at the UCLA Ethnomusicology Archive, brought my attention to ethnomusicology data as a case study. I continued to investigate the advantages and challenges of sharing and reusing humanities research data during a spring 2020 Data Curation and Policy course and an independent research project in fall 2020 where I used the Data Curation Profile interview as a loose template to interview ethnomusicology researchers. Click here to go to that project in my portfolio. The sensitive nature of cultural data collected by these researchers reinforced the importance of ensuring its preservation, providing responsible access, and collecting careful metadata about cultural guidelines for reuse. It also reinforced the growing necessity for better guidance on managing humanities research data from the beginning of its lifecycle.

ARCHIVAL APPRAISAL AND ACCESSION

Spring 2021, I will be appraising a private family collection to determine items for donation. I plan to complete an appraisal report under the guidance of Susan Anderson, the History Curator at the California African American Museum and in conjunction with the Archival Appraisal course taught by Anne Gilliland. I am also assisting Los Angeles Contemporary Archive develop an Accession form for donations received.

In these projects, I will be considering how archival processes can prioritize the needs and voices of communities and the creators of collections. For LACA's vision, an accession form must prioritize recording the voice of the creator at the time of donation and be written in clear and accessible language. In appraising the private collection, I hope to understand, honor, and record the value the family places in its records.

DIGITAL AND PHYSICAL PRESERVATION

I enrolled in Environmental Protection of Collections and Digital Preservation to gain knowledge of collection preservation. During Environmental Protection of Collections, I prepared a report for the Music Library with recommendations for a Collaborative Preservation Policy. The final project for the course consisted of regularly monitoring temperature and humidity, UV exposure, Insect presence, and acidity in one section of the Music Library and preparing recommendations for improving collection storage conditions. I participated in the Digital Preservation course winter quarter, 2020 and this spring, I am enrolled in the Digital Asset Management course, where I hope to learn more about best practices for managing digital collections.

COMMUNICATION AND MENTORSHIP

Before beginning the MLIS program, I had managed the collections of historical figures, archeologists, contemporary artists, and chefs. During my MLIS, I have worked with

GO TO INDEX

ethnomusicologists, herbology researchers, art historians, and anthropologists. Learning to translate the conventions of Library and Information Studies into different disciplinary languages has also prepared me to work with community members on archival projects. As an independent archivist, I have worked with two different families hoping to organize and donate personal collections. I have learned how to listen carefully to the needs and concerns of community members and to ensure they are honored in every aspect of my archival work.

I believe mentorship is important and have sought opportunities to learn and grow through relationships with supervisors and more senior professionals. I have also gained experience as a mentor both through leading a team of interns at Royale Projects and employment as a TA and Reader at UCLA. In these experiences, I have learned how to teach skills in a sustainable way and to be attentive to the needs of students beyond the internship or course. In my career, I will continue to seek out mentorship. I also plan to help the next generation of professionals in the same way that others have helped me.

FUTURE DEVELOPMENT

There is an opportunity for leadership in shaping how emerging humanities data is managed and preserved for the future. To better aid researchers collecting digital materials and data and to be a part of the growing conversation, I have sought out opportunities to improve my data literacy. I attended Library Carpentry workshops on Tidy Data, OpenRefine, Unix Shell, Python, and using Python for API's. I assisted with an Institute for Society and Genetics, Data in and of a Pandemic course, taught by Michael Scroggins, who is also a member of the UCLA Center for Knowledge Infrastructures group. The course investigated issues in data collection and communication as they relate to public health. I plan to continue learning about data visualization and analysis tools uniquely suited for humanities data. I have become a member of the Research Data Access and Preservation Association and look forward to attending their meeting and free training events.

Museums and archives are beginning to consider how better access to their collections can foster knowledge creation about broader movements in cultural history and enable better transparency about their practice. Digital tools and metadata are allowing overlooked histories to come to light. Through continuing to build my knowledge of metadata standards in the digital age and best practices in managing, sharing, and visualizing data, I hope to help lead the archival and museum professions in taking part in this new era in a way that is innovative, ethical, and sustainable. As a member of both the Society of California Archivists and Society of American Archivists, I hope to stay informed about how this change is affecting the practice of my colleagues and to learn about emerging tools, standards, and ideas.

GO TO INDEX

ISSUE PAPER

SUMMARY

I am interested in how the call for open access and increasing volumes of digital content affect the way that humanities researchers collect material associated with their research, share it, and deposit it into a collection. My issue paper focuses specifically on ethnomusicology data, which has a long history of being shared via archives. I explore how archives and researchers can develop ways for ethnomusicologists to take advantage of new concepts and technologies without over-simplifying the complex issues of intellectual property and access.

ETHNOMUSICOLOGY RESEARCH DATA

With the growing call for open data, open access, and data management plans in research there has been a parallel call for discipline-specific management and sharing practices. Even as more institutions are advocating for open data and funding agencies like the NIH and NSF require data management plans and data publishing, others are citing the need for a more careful analysis of what role open data and open access will play in different disciplines depending on existing methods of data sharing and reuse and the data management support needed to facilitate this new data network.[1] While best practices and principles and data publishing infrastructures have been robust in quantitative disciplines, qualitative researchers have often been left out of these solutions. At the same time, humanities and social science researchers amass large collections of digital objects during research that they are then tasked with curating and stewarding. Data management plans developed for quantitative scientific data offer little guidance for these researchers and lack of management training can mean this valuable data is often lost. It is important that we carefully consider the nature of data collected in the humanities, how it is curated, shared, and reused to build solutions that ensure the long-term availability of these data sets. Ethnomusicology, a discipline that collects sensitive cultural heritage data and that has a tradition of depositing that data into archives for reuse, poses an important case study in the importance of domain-specific considerations.

Ethnomusicologists seem to disagree on exactly how to define their discipline. In his Oxford Bibliography, Bruno Nettl offers the frequently used definition "The study of music in relationship to the rest of culture, and the study of the musics of the world from a comparative perspective."[2] The term ethnomusicology places the discipline at the intersection of ethnography and musicology. It is by nature an inter-disciplinary field that often questions its own identity in relation to and apart from those disciplines. Margaret Sarkissian and Ted Solis recently conducted interviews with

[1] For example, Samuelle Carlson and Ben Anderson. "What *Are* Data? The Many Kinds of Data and Their Implications for Data Re-Use" *Journal of Computer-Mediated Communication* 12, no. 2 (2007); Ixchel M. Faniel and Elizabeth Yakel. "Practices Do Not Make Perfect: Disciplinary Implications for Repository Data Curation." *Curating Research Data* 1 (2017); Atici, Levent et al. "Other People's Data: A Demonstration of the Imperitave of Publishing Primary Data". *Journal of Archaeological Method and Theory* 1, no.3 (2012)
[2] Bruno Nettl. "Ethnomusicology" *Oxford Bibliographies* November 29, 2017. DOI: 10.1093/OBO/9780199757824-0224.

GO TO INDEX

ethnomusicology researchers about how they would position themselves in the field and recorded a similar range of responses. While some researchers criticized the centrality of Western musical theory traditional to musicology, others questioned ethnographic traditions that risk exoticizing other cultures' music. Still others considered themselves musicians who study a diverse range of music making. One aspect that is common in many of these responses goes back to Nettl's definition- researchers are interested not only in music theory, but in the cultural contexts that helped form that music.[3].

In 2009, Janet Topp Fargion published an article advocating for better management of ethnomusicology field recordings. She identified the ways that recordings were useful both for reuse by ethnomusicologists who need to understand the historical context of music tradition and for source communities who can use these recordings to reestablish cultural traditions. At the same time, she highlighted the persistent practice of researchers creating recordings for their own research purposes. This meant they did not consider the metadata and formats which would enable long-term preservation and access. Her article called for more active care of these recordings throughout their lifecycle and better training in best practices so that researchers can take on this work in the field. This can be understood within the context of a growing movement for digital curation and open data in social science and the humanities. It was published three years after the American Council of Learned Societies released a report titled "Our Cultural Commonwealth" calling for an investment in cyberinfrastructure for the humanities which "will benefit the public and the specialist alike by providing access to the breadth and depth of the cultural record."[4] Also in 2006, the Open Context repository debuted to encourage archeologists to publish their field notes, maps, and other data on the web. Eric Kansa, who helped establish Open Context, later argued in a publication with Sarah Whitcher that in archeology, "lack of information sharing not only inhibits scholarship, but also represents a tragic loss of irreplaceable cultural and historical knowledge."[5] The same year Topp Fargion's article was published, Purdue University Libraries shared their data curation profiles project, which asks the question "Which researchers are willing to share data, when, with whom, and under what conditions?"[6]. In 2010, the DIPIR project sought to understand more about how social scientists, archeologists, and zoologists used contextual information to determine how they reused data created by colleagues in their fields.[7]

Like archeology, zoology, and social science, ethnomusicology has a history with the reuse and preservation of its data. Like other ethnographic research methods, however, this history is also tied to complicated colonial legacies. The field began in the late nineteenth century as comparative musicology- the invention of the phonogram allowed Western European archivists and researchers to travel to other cultures and record their music, hoping to trace the evolution of the

---

[3] Margaret Sarkissian and Ted Solis "Self-Positioning in and Reflections on the Field" *Living Ethnomusicology: Paths and Practices* University of Illinois Press, 2019. DOI: 10.5406/j.ctvhrd1nb.

[4] "Our Cultural Commonwealth". *American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* (American Council of Learned Societies, 2006), 2.

[5]Eric C. Kansa and Sarah Whitcher Kansa. 2013. "We all Know That 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication." *Journal of Eastern Mediterranean Archeology and Heritage Studies* 1, no. 1 (2013), 90.

[6] "Overview" Data Curation Profiles*,* 2009, http://datacurationprofiles.org/overview.php.

[7] "Dissemination Information Packages for Information Reuse (DIPIR)". *OCLC* https://www.oclc.org/research/areas/user-studies/dipir.html, Accessed December 15, 2020.

GO TO INDEX

human expression of musical culture from its more "primitive" forms.[8] Central to this study was the field recording, which acted as the raw data from which ethnomusicologists formed their theories. Ethnomusicologists have acknowledged their field's relationship with colonial ideologies, and many are considering how they can actively work against them in their practice. As Topp Fargion points out, this can translate into a hesitation to deposit or share field recordings, "Depositing recordings made in Ghana, for example, in an institution in the UK, could be construed as a colonial or imperialist act…The claim that recordings were being made for one's own research purposes and were not being deposited in archives was a way of avoiding such labelling."[9] Ethnomusicology researchers who do engage with open initiatives, therefore, do so with caveats. As Muriel Swijghuisen Reigersberg states, "The importance of sharing ethically and perhaps, therefore, selectively is not always fully understood by other open access supporters, some of whom lobby for the open sharing of all academic content, including data…"[10] It is not simply a lack of data management training that prevents the curation and sharing of ethnomusicology data, but concerns around how access to data relates to a broader reform of research practice.

It is still common practice for ethnomusicology researchers to deposit their research data in archives, but many archives are still faced with deposits that were collected without care for metadata and preservation-quality formats. These collections have only grown more complicated with the introduction of digital formats and digital data. It is now about ten years after the publication of Janet Topp Fargion's article and there is still uncertainty around data management in ethnomusicology. What needs to be explored is how debates around open access, open data, and increasing volumes of digital content affect the way that ethnomusicologists collect material associated with their research, share it, and deposit it into a collection. As a start, I have conducted semi-structured interviews with ethnomusicologists and professionals working with ethnomusicology data inspired by the basic format of Purdue University's data curation profile interview. Recruitment was through an email sent out to members of the Society for Ethnomusicology and to department heads at ethnomusicology programs in the United States. Interviews were conducted with seven participants including senior ethnomusicologists, graduate researchers in ethnomusicology, and professionals managing ethnomusicology data both in archives and at research institutions. They represented seven different institutions in the United States, Germany, and the UK. Data Curation Profiles are time-intensive interviews that ask researchers to describe their data management process, their data lifecycle, and their attitudes towards sharing and preserving their data. However, to avoid overwhelming participants with a long-format structured interview and to better address my research questions at this stage, my shortened interviews used lines of inquiry from the profiles to focus on the three following areas of inquiry:

---

[8] Eric Ames "The Sound of Evolution" *Modernism/ modernity* 10, 2 (April 2003): 297-325. DOI: https://doi.org/10.1353/mod.2003.0030

[9] Janet Topp Fargion, "For My Own Research Purposes? Examining Ethnomusicology Field Methods for a Sustainable Music." *The World of Music* 51, no. 1 (2009): 81.

[10] Muriel E. Swijghuisen Reigersberg "Ethical Scholarly Publishing Practices Copyright and Open Access: A view from ethnomusicology and anthropology." *Whose Book is it Anyway? A View from Elsewhere on Publishing, Copyright and Creativity*, edited Janis Jeffries and Sarah Kember. Cambridge, UK: OpenBook Publishers, 2019, 40.

GO TO INDEX

What is Ethnomusicology research data?

- Are recordings still the primary form of data being created and collected?
- How has the nature of data changed as the field has adopted to new research methodologies and digital information networks

How is it being shared or discovered?

- What motivates researchers to make their data available for discovery?
- Who is using ethnomusicology research data?

How is it managed?

- What best practices and training do researchers use?
- How do researchers manage their large collections of data as they are collecting it?
- What support do researchers want from repositories?

What is Ethnomusicology data?

Even as the field has changed, Ethnomusicology largely remains concerned with recording different processes of making music through ethnographic fieldwork and collaborating with musicians to capture the context within which the music is created. The field recording remains the dominant data produced during research. In 2009, when Janet Topp Fargion wrote her article about research data management practices in ethnomusicology, she stated "Making recordings in the field is, seemingly, a methodology taken for granted."[11] This was confirmed during my research, with almost all participants listing field recordings as their primary form of data. The contents of these recordings vary somewhat between participants. Of the seven individuals interviewed, all referenced interviews as an important aspect of their work. One researcher stated, "I think most of our most of our data is actually discussions about music. And then maybe some recordings of examples and things, but we're much more concerned with how people talk about and think about the music that they're doing." As this researcher notes, recording performance is also a prominent form of data. Some researchers mentioned leaving their recording devices running during performances, cultural festivals, and music rehearsals so that they could have a record of the event to reference later. One more senior researcher mentioned that creating good-quality recordings of performances he felt were especially historically significant was a central part of his research practice. Over half of the participants also mentioned that they had private reference collections of commercial recordings, some of which are considered rare or important.

Reinforcing the importance of context, researcher data includes a large volume of photographs and video recordings. Field notes further record details of context, information that would not otherwise appear in a recording or photograph. Along with capturing external contextual information, ethnomusicology researchers are also concerned with the internal experience they have as participants in musical culture. Field notes act as ways of organizing data and assigning metadata at the end of a day of fieldwork, but many of the participants also emphasized the importance of recording their own affective experience of the performances or traditions they witnessed. One researcher mentioned extensive field notes reflecting on his experience, he

[11] Janet Topp Fargion. "For My Own Research Purposes", 77.

characterized his field note method as “I’d start taking notes about everything I remember that happened there, from my point of view, because I'm in my own body.” Another researcher, who also helps manage ethnomusicology data, mentioned the importance of tacit knowledge in some research methods. She frequently has researchers who are unsure how to describe their emotional experience in terms of “data”, “if they say We're playing music myself with learning an instrument myself. Then the next question is difficult to answer, like what is the data then?”

Interviews and in initial research revealed some emerging changes in research methodology and a corresponding shift in ethnomusicology data. In some ways, the traditional method of traveling to a community and recording their performances is changing. One ethnomusicology researcher, who now works at an ethnomusicology archive, stated that when he was a graduate student “there were a lot of hesitation on the parts of graduate students back then, and I shared the idea, of going out and being the expert in the field and making the recordings and then coming back and deciphering the music culture and publishing about it and being the expert.” This shift is related to the emergence of Applied Ethnomusicology in the 1990’s as a subfield of the discipline. Jeffrey Titon, among others, called for a shift away from a researcher as passive observer of a community to active collaborator. Applied Ethnomusicology involves a field method that directly benefited communities through helping to change policies that would benefit the artists they work with, advocating for cultural traditions, and assisting in educating community members to preserve those traditions.[12] This methodology can be understood within a growing movement in qualitative research methods away from the researcher as expert, to a practice that acknowledges subjectivity and invites collaboration. Anthropologist Eric Lasseter called for collaborative ethnography, which he defines as “The collaboration of researchers and subjects in the production of ethnographic texts”[13] Building on feminist principles and the concepts of applied and public anthropology, Lassiter explains the importance of recognizing interlocutors as collaborators, which could mean co-authorship in published texts. Researchers are thus more aware of communities as stakeholders in research products, meaning that their needs and concerns should be factored into how they are curated. It also means that any published version of those products may need to recognize multiple co-authors, thus complicating Intellectual Property and copyright concerns.

Digital networks have, in many ways, made this collaborative research practice more feasible. As opposed to limiting interaction with communities to time spent in the field, researchers can interact with communities throughout their research process through online tools. One PhD researcher noted less travel and more interaction and collaboration with community members online, including accessing recordings they had made themselves, “many of the other people you're studying have ways to record and document themselves. And that information, sometimes it's even more interesting than what you can get yourself.” It is in many ways fitting that research took place during the Coronavirus pandemic. Travel restrictions have forced many researchers to consider how they can continue fieldwork through digital environments in lieu of traditional ethnographic methods. Beyond collaboration with interlocutors, researchers are finding a wealth of contextual information about musical cultures online. This includes websites, social media posts, and tiktok

[12] Jeff Titon and Svanibor Pettan. “An Introduction to Applied Ethnomusicology”. *The Oxford Handbook of Applied Ethnomusicology.* Oxford University Press, 2015. DOI: 10.1093/oxfordhb/9780199351701.013.2 .
[13] Luke Eric Lassiter. “Collaborative Ethnography and Public Anthropology.” *Current Anthropology* 46, No. 1 (2005): 83.

GO TO INDEX

videos. Researchers mentioned printing out websites, taking screenshots of facebook comments or Instagram posts, and collecting website URL's. Social media platforms also enable long-distance research collaborations between different disciplinary researchers and communities. For example, a research project presented at the 2017 International and Inter Disciplinary Conference on Arts Creation and Studies involved conservationists, ethnomusicologists, and performers connecting over social media to research, conserve, and revive interest in the bundengan instrument from Indonesia.[14] This type of ethnography has been labeled "Hybrid ethnography" by Liz Przybylski who has an upcoming publication with that title.[15] Born-digital data and collaborative research practices continue to complicate an already complicated data management process in ethnographic methods. Researchers may consider websites, social media platforms, texting or team collaboration threads as data, but preserving data created using proprietary software is notoriously difficult. User agreements specify different levels of ownership over material created on these sites, meaning researchers may not have total ownership over their data. Companies tend to limit API capabilities and options to download content from websites, often meaning some aspect of the look or feel of content is compromised and the data is in a format that is difficult to preserve. Comments and websites may be taken down by platforms after a period and even non-proprietary websites need to be carefully maintained. Finally, researchers using online recordings created by interlocutors or collaborators from communities must be even more cautious to obtain permissions and to record creator, rights, and ownership metadata. While these recordings may be a vital part of research, ownership of this content legally belongs to its creator- or, more often- to the platform on which it was created. With increasingly digital ethnographic and collaborative research methods come new concerns for data management practices.

While ethnomusicologists do not often reuse the work of their colleagues, revisiting historic collections marks another recent shift in methodology. Most researchers interviewed did not initially use other researchers' recordings or visit archives. One participant stated, "I mean we do have a lot of historical traditional stuff. But I guess we're interested in how it's being performed now, or very recently." Another researcher, when asked if he used other researchers' data or recordings in an archive, said it hadn't been a part of his research practice, explaining "…it's also very much a solitary endeavor. So there's a lot of expectation that you have done all your own field work, that you have written the whole thing yourself. It's all by you." Carolyn Landau and Janet Topp Fargion noted in 2012, "Ethnomusicologists made their own recordings in the field and often considered recordings made by others to be of less importance or interest."[16] In this publication, however, they support a growing interest in researching existing collections and making them accessible. This is important for a revival and reinvention of comparative musicology, but also for the partnership with source communities advocated by applied ethnomusicology. The second researcher quoted above also mentioned that he has transitioned away from working in the field and is currently working on researching, describing, and making accessible archived folkson musical scores with the hope that

---

[14]Rosie H. Cook, et al. "Bundengan: Social Media as a Space for Collaboration in the Conservation and Revival of an Endangered Instrument." *Proceedings The 2nd Annual of International and Interdisciplinary Conference on Arts Creation and Studies* Institut Seni Indonesia, 2017.

[15] Liz Przybylski. "Hybrid Ethnography, Online, Offline, and In Between". *Qualitative Research Methods* 58. Sage Publications, 2020.

[16] Carolyn Landau. "We're All Archivists Now: Towards a more Equitable Ethnomusicology." *Ethnomusicology Forum* 21, 2 (2012). DOI: 10.1080/17411912.2012.690188

GO TO INDEX

it will revive interest in the tradition he researches. There is also growing interest in the kinds of data analytics that could be applied to digitized or born digital ethnomusicology data. One researcher I interviewed is conducting most of her research through using available online datasets and performing social scientific analysis on them. In a 2014 research paper titled "Big Data for Musicology", researchers from the University of London listed ways in which widely available, aggregated recordings could be analyzed to explore how performance changes over time and between different cultures, questions central to the fields of comparative musicology and ethnomusicology.[17] As previously noted, however, historical lack of data management in ethnomusicology and resulting messy collections mean there is much data cleaning labor to be done to enable this work.

How is it being shared and discovered?

Even as more researchers are attempting to work with existing ethnomusicology data and to make collections accessible, many still hesitate when it comes to sharing their data. In many ways, this is due to precedents in ethnomusicology research that do not support sharing and using field recordings, as one researcher noted "the nature of the ethnomusicology discipline, researchers tend to be very individualistic, you know, it's like they collect their data, but there's not usually a desire to share in particular." A researcher who has published his field recordings as LPs noted that this type of access is unusual, "they (ethnomusicologists) might write about them and they might make printed musical transcriptions of some of them, but you never get to hear the stuff. Nobody else does until it finally goes into an archive and a lot of them are frankly not very well-recorded." Another researcher currently working on a dissertation mentioned that he would eventually deposit all of his material into an archive to be accessed, but not for 30-40 years when the material is already considered old. He explained that this was the existing model for him to follow. This practice was echoed by more senior researchers that I interviewed, who confirmed that most of their field recordings had not been shared, that they had been waiting until retirement from research to determine which archive they would donate material to. Beyond a lack of precedence, some participants noted a fear that sharing too early could harm their careers. As one researcher stated, "… it's kind of like the scientists sharing their data, that's their work, they have to publish stuff that's unique in order to get credit, same thing goes for data in our field… I'm not going to release a lot of my materials until I've written about it and publish what I can from it". Another participant had worked with younger ethnomusicology researchers with rich private archives who had no intention of sharing because they have no professional reason to, "They have splendid private archives that nobody knows about, and they don't talk about it because they think they're not getting credit for what they're doing."

Beyond personal, professional barriers to sharing is the growing emphasis on honoring partnerships with performers and communities and concerns about how access to recordings could impact those partnerships. This is particularly sensitive for some researchers who work with indigenous groups, as with one of the researchers interviewed who stated, "ethical concerns are not so much around making money and who owns that money, but around cultural identity property and the good or in some cases damage you could do by sharing things that may or may not be

[17] Tillman Weyde et al. "Big Data for Musicology" *DLfM '14: Proceedings of the 1st International Workshop on Digital Libraries for Musicology* (2014), 1.

shareable with particular communities". For one researcher, remuneration was an aspect of concern when it came to allowing access to his recordings. He created commercial recordings from his field work and had built networks of trust with his community by giving performers a share of royalties. He explained, "I wanted to share the material, the problem is that of course there's an awful lot of unauthorized use of material these days…things that are in print just get reused without attribution, photographs are put on the internet, recordings wind up on YouTube, and no royalties are paid or no credit given, it's just a really a bad situation." Beyond concerns with intellectual property, researchers were worried about sensitive content in their recordings that may cause harm. The nature of some researchers' field methods, of simply leaving recording instruments on during informal jamming sessions, meant that participants were recorded saying things they would not otherwise share "so you can imagine... Five or six guys sitting around in a circle in chairs, playing music and taking long breaks between songs, talking about all kinds of stuff and gossiping." Another researcher confirmed what he referred to as "inter-community issues" that could be caused by what some community members divulged during interviews or informal conversations. Ethnomusicologists also think about the complex history with community access to recordings. One participant with experience in archives mentioned that open access may only be a painfully ironic reminder of a long history of barring community members from using their own material, "…in the past, people couldn't come and listen to a recording of their grandmother because they weren't a researcher… am I really going to be the person to.. say, now we want to put them online for free for everyone to listen to".

For some researchers, however, dedication to these partnerships is what motivates them to provide better access to their data. As researchers have noted, many ethnomusicologists take issue with the practice of recording cultural heritage without any systems of reciprocity built into field work. Not only does this mean giving copies of their recordings to communities or community archives, but it involves assisting in making existing archival collections more accessible. One researcher notes her frustration that while the archive she partners with is able to provide some access to the source community, it's limited "there is this problem of how do people get access to this. The communities I'm working with have contacted the archive, but if people really want to see my whole collection, they have to travel to (the archive)… that's a problem." Like other humanities scholars, many of the ethnomusicologists I interviewed believed access to their work was also beneficial to the general public, "You know, I don't publish just for my professional colleagues or students and I don't write just for them, I write for anyone that's interested in the music." Another researcher stated, "I think a lot of us look at our work as being useful not just for our own academic endeavors, but potentially also useful for the communities that we work with, or also the general public." Researchers mentioned the belief that preserving recordings was important because it preserved a vital aspect of human history, "I just feel that way about my own data, but also about my colleagues, because otherwise, why are we doing this if it's not... If we don't think it's important, and we're trying to preserve a snapshot in time of musical culture." This language echoes some of the advocacy around the open access movement such as the 2006 "Our Cultural Commonwealth" report, which advocated for ways in which digital access to cultural heritage would benefit both scholars and the general public.[18] Jeffrey Titon uses the language of commonwealth on his

[18] "Our Cultural Commonwealth". *American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* (American Council of Learned Societies, 2006), 2.

sustainable music blog, advocating for expressive culture to become more of a shared resource, "What folklorists can contribute to the discussion of a cultural commons, then, is based in part on this longstanding concern, where the advantages for a community of shared resources are plain: acceptance by, and accessibility to, anyone and everyone."[19]

While several participants reinforced the individualistic nature of ethnomusicology research, they also mentioned networks of collaboration with researchers that they had already participated in or would like to. While this rarely involved sharing raw data, many researchers explained that it was important to share some resources with other scholars working in their area or to be in contact with them. One participant explained,

> "There's two other researchers that I have kind of shared with but .. it's very limited about we actually shared with each other…what we shared was older recordings that are just harder to find, and digitized kind of more like archival material… we went to different repositories and different people and shared everything that we were allowed to share with each other."

Another researcher advocated for the benefits of more collaboration between researchers, "Like if I'm going to study a community in China, for example. Oh, it will be great to know that some people already went to that community in China… I already can have data from those experiences, data that can be shared and public and then that can give me a foundation in order for me to conduct my own research and guide my own research." Most participants stated concerns about sharing recordings too early, or sharing certain types of recordings, but one researcher relayed a positive example of how sharing her recordings enabled new scholarship, "I gave all my audiovisual recordings to a bandleader in… the village where I work. And he published a blog, but he didn't publish my material because people also have some ethics, they make something new out of it." Interestingly, some researchers mentioned that they wanted to share their interviews in service of transparency and trust, "I think that people reading my work, they should be able to access the original files to see how much I divert or how much things could be different." Another researcher stated, "When I do deposit it in an archive… people do have access to the recording itself. And so they can go back to that section and actually hear the person speak for themselves and they hear what they say on how to get permissions."

A few of the researchers actively supported open access. For these participants, one of the main reasons they believed ethnomusicology should engage in this movement was a fear that the discipline could be left behind when important infrastructures and policies are created. They stated that better support for managing digital data was going to disciplines that more readily participated in open access, "Our needs, in terms of support, are less-understood than the sciences, because they have been running the show as far as data management- they were the first to prioritize open access, so there is still a lot of catching up that we need to do in the arts and humanities." Another researcher believed that not only does participating actively in open access enable better support, it allows researchers to determine themselves what aspect of their data to make open, "if you want a good management of your data. You also need to be open to publish at least not open to publish everything or to make everything open data, but to make a decision on what could become open

[19] Jeffrey Titon. "The Commonwealth of Culture". Sustainable Music, 2013. https://sustainablemusic.blogspot.com/2013/

GO TO INDEX

data, even if it's just the metadata." The researcher who works with aggregated ethnomusicology data and is excited about more collaborative work expressed her frustration that fewer humanities disciplines were working to create their own data repositories,

> "numbers are the core for quantitative research, obviously we're gonna find way more repositories that provide numerical data than ones that are mostly based on observations or qualitative data. So, that is why, sadly, you know, so many sciences are ahead of the humanities. I believe we have not done enough because it's not that we cannot do more… but we still have a lot of work to do to close that one gap in the ways you can create repositories."

For these researchers, open access represents not only the potential to expand ethnomusicology research, but an increasingly expected and accepted aspect of scholarship. Avoiding engagement with this practice may deny researchers access to vital support and resources, not only in terms of data management, but in government funding as more agencies require depositing data.

How is it managed?

One of the open access advocates made the argument that while the recordings or field notes may not be open, metadata about these collections should be discoverable, "Metadata, then, is more important than raw data. If people know where metadata is, they then know where to find the data and can decide for themselves whether they need access to the raw data." Metadata was not a part of the field methods training most participants received. Many admitted that they neglected organizing and labeling their data until it was too late and the contents were even indecipherable to the researcher, "I didn't take field notes back then, but I would still want to provide some metadata... Yeah, on some of the recordings where I don't have any field notes, I go back and listen. Who was that guy? I recognize his voice, but... Who's that again?" or, as another researcher recalls, "And it was just kind of like… this mishmash of stuff. And I was like, I guess I can only use these interviews really and sort of my own, like, practice recordings, because I don't know where everything is." Many researchers did not realize how important it was to curate their data until they were attempting to deposit it and their lack of metadata or their data format created a barrier. In addition to organization and data formats, ethnomusicologists must be careful about obtaining permissions from individuals in their recordings and they must make it clear how these recordings will be used and where they will be stored. As one ethnomusicologist and data manager explains, "In order to manage my data well I should have thought of it at the start of my project not just in terms of who should have access to it, but also what technological format should I be recording stuff in, who should be ethically made aware that I'm collecting this information and tell them where I'm storing it so that if they want to access it, they can know where to look in order to ask."

Interestingly, a few researchers mentioned the individual nature of ethnomusicology research as an explanation for a lack of preparation for organizing and describing their field data. One researcher expressed frustration that when he sought examples of systems of tracking and organizing data and metadata, he was told "It was always like do something that works for you and actually frustrating because I felt there was lack of examples of what that meant." When asked her thoughts about the discoverability of ethnomusicology data and metadata, another researcher explained, "They have like their own system individual systems that they use for data management and collection… we have to get the community to come together in order to define parameters for

that as well." Topp Fargion pointed to this in her own publication about ethnomusicology data management. She explains that many ethnomusicologists only receive very general training, the assumption being that most field conditions will be unique and thus specific practices could not apply to all of them.[20] Among the researchers I interviewed, the amount of training they received varied. Some had been trained on recording equipment and had some knowledge of permission forms and documentation, but most were left to figure this out on their own by reaching out to colleagues and mentors or learning what not to do in the field through experience.

Many researchers mentioned a desire for more training in ethnomusicology data management. Knowing that I have a background in Library and Information Science, a few participants stated that they wanted more guidance from the archival community on how they should be preparing their collections for donation. Two researchers mentioned wanting archives to specify what data they wanted to collect, what metadata was important to include with that data, and what formats metadata should be in in order to easily deposit data, "…(how) archivists could also do a service for these academic fields, by yeah, giving them a little guidance is that in terms of metadata, here would be an excellent format that we could just import right into... Translate to Marc format or whatever the library and used to do their work, so that when you did go to deposit it in an archive will already be pretty much ready to go.." Another researcher, when asked whether he wanted more training in data management answered, "how would be like an acceptable way and easiest way for people to Intake documents and things and just kind of like a guideline." At the same time, other participants who had experience in archives cautioned that it was important for archives and policy makers to account for the specific needs of ethnomusicologists as data management policies are being formed. When asked what support ethnomusicologists would want from a repository, one participant answered, "Ideally, staff who are prepared to have that conversation with us about what we need on a project-specific basis and not to assume that we are either like all ethnomusicologists out there and not just like the sciences – some rules that apply to the sciences do not apply to us, and that we may not want to deposit our work in the University archive but could potentially deposit elsewhere."

Archives and ethnomusicology have a long history. Ethnomusicology archives are almost as old as the discipline itself- the phonogram archive in Berlin was founded in 1900 and was used by the founding members of comparative musicology.[21] For many ethnomusicologists, and for all of the researchers I spoke to, depositing their research materials into an archive is an expected part of their research process. Even still, most researchers in ethnomusicology do not interact with archives until the end of their career, when they are ready to deposit their work. As indicated above, ethnomusicology researchers seem unsure of the archival process: what they should include in their collection, what metadata needs to accompany it, and what formats and quality the archive will accept. This leaves archives to deal with historically messy researcher collections to process and a huge back log of work. One participant, a current ethnomusicology archivist, shared that "I'm also hesitant to take very much right now because I already have such a huge backlog, it's again kind of unethical to keep taking collections and just letting them sit.." When asked what ethnomusicologists prioritized when considering a repository to donate their collections to, several brought up the issue of accessibility again, "they'll (archives) just prioritize whatever people want

[20] Janet Topp Fargion "For My Own Research Purposes"?, 84.
[21] Eric Ames "The Sound of Evolution", 299.

GO TO INDEX

and they'll just digitize those things and then other things will be left because you know there's not enough time and money… Yeah, so just accessibility. I guess I would want to check up on that." However, with the evident lack of curation practice in the field, processing deposited ethnomusicology collections to prepare them for broader access is often a labor-intensive task. While ethnomusicologists seem willing to learn more about managing data and using metadata, archives and repositories seem cautious about prescribing metadata formats and practices that are too technical and may not apply to the needs of researchers, "people Have come to the library and ask for help with a digital project have just been inundated with metadata standards and sometimes I think you know what they don't need that." And while it seems that more archives are prioritizing access through digitization, they share a concern about harm that could be caused by unrestricted access to collections, "There are a lot of recordings that I currently oversee that I shouldn't even listen to them and I'm supposed to take care of them." What both archives and researchers seem to agree on is that there needs to be more communication between them in the service of better access to collections. This means working together to determine how to provide better data management training and communication at every stage of the data lifecycle.

Discussion:

There is increasing interest in fostering collaboration between ethnomusicology researchers, renewing tradition through community use of recordings, and treating ethnomusicology collections as a commonwealth for the public through broader accessibility to data. Several online projects and repositories have already begun to explore how to enable this accessibility. As a part of his exploration of the benefits of open access in musicology as a PhD student, Darren Mueller cofounded Provoke! Digital Sound Studies, an online repository for experimental scholarship centered around sound.[22] "Transforming Musicology" is a digital platform started in collaboration between several Universities in the UK that aims to digitize and provide access to musicology data in order to sponsor new projects in the study of music.[23] EVIA Digital Archive is a repository for ethnographic field videos, which makes them available for use by scholars and instructors.[24] PARADISEC is a digital archive for recordings of endangered languages and associated materials.[25] This reflects, perhaps, a broader trend in ethnomusicology away from the traditional model of collecting recordings and content from communities and depositing them in Western archives after a long embargo. Many researchers are more interested in working collaboratively with communities and each other, reevaluating and digitizing existing collections, and broadening the ways they collect and analyze data using new digital tools. These researchers are willing to consider ways to enable better discovery of ethnomusicology collections through metadata and linked networks. At the same time, there is more awareness about the potential harm of open access being universally applied to cultural heritage collections. Researchers who have long been working to foster trust with communities are wary of the call for sharing data that concerns their traditions. If ethnomusicologists want to join the open access movement, there must

[22] "Provoke! Digital Sound Studies" Duke University, https://soundboxproject.com/index.html . Accessed November 4, 2020.
[23] "Transforming Musicology", an AHRC Digital Transformations project" Arts and Humanities Research Council, https://tm.web.ox.ac.uk/ Accessed November 4, 2020.
[24] "EVIA Digital Archive" Indiana University, http://eviada.webhost.iu.edu/Scripts/default.cfm . Accessed November 4, 2020.
[25] "PARADISEC" PARADISEC, https://www.paradisec.org.au/ . Accessed November 4, 2020.

GO TO INDEX

be changes to the existing model for depositing data, more and better training on data management throughout the research data lifecycle, and collaboration on best practices between researchers and archives and within ethnomusicology research networks.

When Janet Topp Fargion wrote her article advocating for better management of ethnomusicology recordings, it was in the context of a more sustainable music that would preserve these pieces of intangible cultural heritage for use by others. She stated the importance of being intentional about the purpose and lifecycle of these recordings from the start of a project, "To do this we have to have specific, carefully planned reasons and methodologies for making or using recordings".[26] If researchers want to participate in open access on their own terms, they must be clear about their intentions for sharing from the start of their project. They must carry this intentionality into how they manage their data. As one researcher stated, "how does my knowledge of research data management affect the way I do ethnography. It's interesting. Every step…the way you interact with People like you need to give them permission sheets and everything and from the beginning, you know that you need good quality data." This is especially important when applied and hybrid ethnographies mean complicated digital data formats with various co-authors and copyright concerns. Researchers must be especially careful about choosing what platforms to use for communication, how they plan to preserve websites and avoid broken urls, and whether their formats are open and non-proprietary. They must also be careful about recognizing collaborators and recording vital rights metadata specifying how this data should be accessed and whether it can be reused. It also means critically considering where this data should be deposited as it is being collected, conversations with community collaborators and archives about how to enable responsible access. This will make field work more difficult and researchers will need more support navigating the data management process. As Janet Topp Fargion argued in her publication, data management training must be a part of field methods courses in ethnomusicology. Library and archival professionals have been and could continue to be a source of this support. However, support and guidance will also have to come from within the profession. One participant stated, "I think people who are supervisors- people who, I guess, are early to mid-career researchers and have that relationship with their doctoral students, should be required to teach data management or to at least flag it at the time that researchers are doing work." As ethnomusicologists are beginning to support more collaboration in the research process, researchers should be more collaborative in sharing their data management strategies and building best practices from within the field.

Making ethnomusicology data more open will mean a reevaluation of the relationship between ethnomusicology researchers and archives. Archivists have experience with creating metadata for and providing access to collections. They could offer useful insight about what researchers can do at the beginning of and throughout their research process to facilitate better access once they are deposited. Researchers will have to be advised to reach out to archives about standards and expectations throughout their project instead of waiting until they are ready to deposit. In many ways, collaborative ethnomusicology requires researchers to think more about the archival process, which one participant pointed out to me "I think they will kind of and maybe they already are beginning to realize how important (archival processes) are. To the field in general… especially in this kind of giving back to communities. And making connections with other fans and researchers who are interested in this music." Researchers are considering how their recordings

[26] "For My Own Research Purposes?", 86.

can be useful beyond their own scholarship, this requires a more careful curation of their collections. While library and archives staff can provide guidance on best practice for managing collections, researchers should advocate to staff members for how broad policies translate into their specific research process. They can provide insight about how ethnomusicology research is currently being done and their expertise can inform the development of policies about privacy, sharing, access, and preservation. These partnerships can ensure that broader access is done in a way that considers the unique nature of ethnomusicology data.

This attention to archival practice extends beyond preparing collections for traditional storage in an archive, but to a more proactive engagement with what depositing data could mean in an applied and open practice. As some participants have noted, more funding agencies are requiring researchers to create data management plans and deposit their research data. Discipline-specific and academic repositories have emerged in response to this change. While more common in quantitative sciences, repositories such as the archeology data repository "Open Context" represent domain solutions to needs in cultural heritage disciplines. Ethnomusicology data that is sensitive and has multiple co-authors may not be compatible with repositories that require an open or CC-By copyright policy. Heterogenous data sets that consist of audio-visual file formats require considerable storage space and specified preservation solutions. Ethnomusicologists stand to gain from advocating for the creation of their own domain repositories that address these needs. Even if depositing raw data into digital repositories is unrealistic for ethnomusicologists, the practice of making existing archival collections more discoverable should extend beyond digitizing collections to critical conversations about how collections should be accessed and how digital tools can aid in this work. Digitized collections do not have to mean uninhibited access. Library and Information scholars have developed tools specifically for navigating digitized cultural heritage collections in a respectful way. The Mukurtu Collections Management System, for example, was created with non-Western cultural knowledge sharing protocols in mind and its mission is reflected in its metadata standards and software capabilities.[27] There are other solutions to providing better discovery of data. The SHARE model represents a tool that is not a repository, but an inventory of existing data in the sciences. Researchers record metadata about "what the data is; who created it and their affiliations; what organization and what program or grant funded its creation or capture (if any); where it is currently stored; who is funding the management of the data and how long that funding is guaranteed; and some notes on any access or use restrictions (e.g., embargoes, human subject constraints) that may apply to the data."[28] Investment in similar infrastructure for digitized ethnomusicology collections could allow broader discoverability while allowing researchers to determine what data they would like to deposit, where, and under what access conditions.

Conclusion

The current model for ethnomusicology data management largely leaves the organization and description of data up to the individual researcher. Due to the sensitive nature of recordings and the amount of work it takes to prepare them to be shared, these researchers wait to deposit their data in an archive until the end of their career. Once they are deposited, the bulk of

[27] Mukurtu.org, Accessed December 15, 2020.
[28] Clifford Lynch. "The Need for Research Data Inventories and the Vision for SHARE." *Information Standards Quarterly* 26, no. 2 (2014), 29.

responsibility for managing that collection and making it accessible falls to archivists. Open Access will mean that ethnomusicologists fundamentally change this model. Researchers will have to be conscious of curation at the beginning of their projects. They will have to consider who should have access to their work, the benefit of sharing material, what data is too sensitive to share, and how this access will affect the way they manage their data. This will make what is already complicated field work more difficult. Ethnomusicologists need to decide whether they feel the benefits are worth the challenges. While some researchers were interested in talking about data management and open access, other potential participants responded that it did not apply to their practice. One participant expressed frustration with increasing amounts of paperwork and policies already being added to the fieldwork process. There are compelling arguments for the benefits of ethnomusicology taking a more active role in data curation. There exist precedents for domain solutions to data sharing and repositories that allow for a more complicated model of data sharing. At this moment, it is up to ethnomusicology researchers to decide whether these benefits are worth fundamental changes to the existing model. They have an opportunity to engage in the open access dialog, make their specific needs known. They do not have to take part in this work alone. Archival and data management specialists have much to offer as partners in taking on these changes. Working with ethnomusicologists earlier and at more stages in the data lifecycle will also mean more work for professionals dealing with existing backlogs. At the same time, forging new and specific practices for data management in ethnomusicology will not only aid a shared dedication to access to ethnomusicology collections, but will facilitate faster and easier processing of collections in the future.

GO TO INDEX

Works Cited

Ames, Eric "The Sound of Evolution" *Modernism/ modernity* 10, 2 (April 2003): 297-325. DOI: https://doi.org/10.1353/mod.2003.0030

Atici, Levent et al. "Other People's Data: A Demonstration of the Imperitave of Publishing Primary Data". *Journal of Archaeological Method and Theory* 1, no.3 (2012): 663-681.

Carlson, Samuelle and Ben Anderson. "What *Are* Data? The Many Kinds of Data and Their Implications for Data Re-Use" *Journal of Computer-Mediated Communication* 12, no. 2 (2007): 635-651.

Cook, Rosie H., et al. "Bundengan: Social Media as a Space for Collaboration in the Conservation and Revival of an Endangered Instrument." *Proceedings The 2nd Annual of International and Interdisciplinary Conference on Arts Creation and Studies* Institut Seni Indonesia, 2017.

"Dissemination Information Packages for Information Reuse (DIPIR)". *OCLC* https://www.oclc.org/research/areas/user-studies/dipir.html, Accessed December 15, 2020.

"EVIA Digital Archive" Indiana University, http://eviada.webhost.iu.edu/Scripts/default.cfm . Accessed November 4, 2020.

Faniel, Ixchel M. and Elizabeth Yakel. "Practices Do Not Make Perfect: Disciplinary Implications for Repository Data Curation." *Curating Research Data* 1 (2017): 103-126.

Kansa, Eric C. and Sarah Whitcher Kansa. 2013. "We all Know That 14 Is a Sheep: Data Publication and Professionalism in Archaeological Communication." *Journal of Eastern Mediterranean Archeology and Heritage Studies* 1, no. 1 (2013): 88-97.

Landau, Carolyn. "We're All Archivists Now: Towards a more Equitable Ethnomusicology." *Ethnomusicology Forum* 21, 2 (2012). DOI: 10.1080/17411912.2012.690188

Lassiter, Luke Eric. "Collaborative Ethnography and Public Anthropology." *Current Anthropology* 46, No. 1 (2005): 83-106.

Lynch, Clifford. "The Need for Research Data Inventories and the Vision for SHARE." *Information Standards Quarterly* 26, no. 2 (2014): 29-31.

Nettl, Bruno. "Ethnomusicology" *Oxford Bibliographies* November 29, 2017. DOI: 10.1093/OBO/9780199757824-0224.

"Our Cultural Commonwealth". *American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* American Council of Learned Societies, (2006).

"Overview" Data Curation Profiles, (2009) http://datacurationprofiles.org/overview.php.

"PARADISEC" PARADISEC, https://www.paradisec.org.au/ . Accessed November 4, 2020.

"Provoke! Digital Sound Studies" Duke University, https://soundboxproject.com/index.html . Accessed November 4, 2020.

GO TO INDEX

Przybylski, Liz. “Hybrid Ethnography, Online, Offline, and In Between”. *Qualitative Research Methods* 58. Sage Publications, 2020.

Sarkissian, Margaret and Ted Solis “Self-Positioning in and Reflections on the Field” *Living Ethnomusicology: Paths and Practices* University of Illinois Press, 2019. DOI: 10.5406/j.ctvhrd1nb.

Swijghuisen Reigersberg, M.E. “Ethical Scholarly Publishing Practices Copyright and Open Access: A view from ethnomusicology and anthropology.” *Whose Book is it Anyway? A View from Elsewhere on Publishing, Copyright and Creativity*, edited Janis Jeffries and Sarah Kember. Cambridge, UK: OpenBook Publishers, 2019. 309-345.

Titon, Jeffrey. “The Commonwealth of Culture”. Sustainable Music, 2013. https://sustainablemusic.blogspot.com/2013/

Titon, Jeff and Svanibor Pettan. “An Introduction to Applied Ethnomusicology”. *The Oxford Handbook of Applied Ethnomusicology.* Oxford University Press, 2015. DOI: 10.1093/oxfordhb/9780199351701.013.2

Topp Fargion, Janet. “For My Own Research Purposes”?: Examining Ethnomusicology Field Methods for a Sustainable Music.” *The World of Music* 51, no. 1 (2009): 75-93.

“Transforming Musicology”, an AHRC Digital Transformations project” Arts and Humanities Research Council, https://tm.web.ox.ac.uk/ Accessed November 4, 2020.

Weyde, Tillman et al. “Big Data for Musicology” *DLfM ’14: Proceedings of the 1st International Workshop on Digital Libraries for Musicology* (2014).

GO TO INDEX

MAJOR PAPER

ART MUSEUM PROVENANCE AS DATA

Winter, 2021
IS 438B Archival Description and Access
Professor Jonathan Furner

Prompt: For this assignment, you will write a final paper on an issue in archival informatics. The topic of your paper should be some issue or research question that is currently attracting interest in the field and/or profession

ABSTRACT

Art museums have recently been making not only their collections digital, but their provenance data available as datasets. Behind this change is the intersection of several different movements including a more critical museology, the rise in provenance research, the open access movement, and the development of digital art history. Most of the museums who have actively taken part in this work are major institutions with large staffs and impressive budgets. While the decision of whether to publish provenance data is a political one, it is also a practical one: the labor necessary to prepare machine-readable datasets from archival records is intimidating and not all museums may have the resources necessary to take this on. Museums are at a moment of change, and in many ways the practice of publishing collections data represents an opportunity to fully embrace this change. This paper explains why some feel this work is necessary, the political factors museums face in deciding to participate, and the practical barriers that make this work so challenging. While the call is for all museums to publish collections data more broadly, this paper focuses on art museums and provenance data specifically. Digital art history is a growing field that relies on the publication of art museum collection data, and most often provenance data. As a discipline, it seeks to address some of the same challenges that museums are facing in trying to build a practice that reflects the issues and priorities of the twenty first Century. The paper proposes that the answer to many of the challenges museums face in publishing provenance data and meeting changes in their role can be addressed through active partnerships with digital art historians as stakeholders in this work.

TEXT

Several art museums have recently made their collection's provenance available as datasets online. Behind this is the intersection of a few different movements including a more critical museology, the rise in provenance research, the open access movement, and the development of digital art history. Museums, more broadly, are at a moment of change and publishing collections data is evidence of this change and an opportunity to fully embrace it. The decision of whether to do so is thus a political one, but it is also a practical one: the labor necessary to prepare machine-readable datasets from archival records is intimidating and not all museums may have the resources necessary to take this on. It is important to explore why some feel this work is important, the political factors museums face in deciding to participate, and the practical barriers that make

GO TO INDEX

this work so challenging. While the call is for all cultural heritage institutions to publish collections data more broadly, this paper focuses on art museums and provenance data specifically. Digital art history is a growing field that relies on the publication of art museum collection data, and most often provenance data. As a discipline, it seeks to address some of the same challenges that museums are facing in trying to build a practice that reflects the issues and priorities of the twenty-first century. Through examining art museum provenance data, digital art history, and the shared challenges museums and art historians face in collecting and using this data, this paper points out the advantages of an active partnership between the two.

Museums are literally in the process of redefining themselves. In 2017, under the leadership of Jette Sandahl, members of the International Council of Museums (ICOM) began meeting to rewrite the existing ICOM museum definition, which lays out the purpose and functions of museums internationally. In an interview in 2017, Sandahl explained that she felt the existing museum definition did not speak the language of the twenty-first century.[29] The new definition reframes museums as "democratizing, inclusive and polyphonic spaces for critical dialog…" directly references the need to address contemporary issues and asserts "They are participatory and transparent, and work in active partnership with and for diverse communities to collect, preserve, research, interpret, exhibit, and enhance understandings of the world…"[30] Museums have faced growing controversy at the turn of the twentieth and into the twenty-first centuries. Concepts like the "Exhibitionary Complex"[31] and "Museum as Contact Zone"[32], which critically examined the history, heritage, and motives of museums ushered in a new era of questioning these institutions. Around the same time that these articles were coming out, there were protests in front of the British Museum by Greek students who felt that its prized Elgin marbles were taken through illegal looting and should be returned to Greece. More museums were faced with a reckoning over the provenance of their collections, a lack of transparency about their practices, exclusion of more diverse voices on staff and among visitors, and a lack of effort to reach out to and interact with communities. Growing movements around decolonization have directly addressed colonial museums and their collections of indigenous and non-Western heritage. The recent Black Lives Matter movement has brought more attention to the ways in which museums and related institutions have historically excluded Black individuals in its collecting policies, hiring, and outreach. Clearly, the twenty-first century has marked a major and contentious shift in the way the public interacts with and thinks about museums.

The beginning of the twenty-first century also marked the growing Open Data movement. This movement gained traction, initially, in scientific disciplines where scholarship has relied on principles of validity and reproducibility, researchers' individual data often belongs and contributes to the work of a network of colleagues, and reusing existing and published datasets has become a more encouraged form of scholarship. Perhaps best exemplified by the Human Genome Project, which began in 1990 and was completed by 2003, sharing open scientific data has more recently been formalized as a broad policy. The NIH released a mandate in 2003 that requires any research

---

[29] "The Challenge of Revising the Museum Definition" *International Council of Museums*, https://icom.museum/en/news/the-challenge-of-revising-the-museum-definition/ , Accessed February 6, 2021.

[30] "ICOM announces the alternative museum definition that will be subject to a vote" *International Council of Museums*, https://icom.museum/en/news/icom-announces-the-alternative-museum-definition-that-will-be-subject-to-a-vote/ , Accessed February 5, 2021.

[31] Tony Bennett. "The Exhibitionary Complex" *New Formations*, 4 (1988): 73-102.

[32] James Clifford. "Museums as Contact Zones" *Routes, Travel, and Translation in the Twentieth Century* (London: Harvard University Press, 1999).

project funded by its grants to make their research available to the public.[33] In 2004, the international Organization for Economic Co-Operation and Development presented a declaration calling for publicly funded scientific data to be made open.[34]

The humanities have had their own trajectory towards a more open sharing policy. Just as the internet has enabled new policies in sharing scientific data, it has brought new challenges and initiatives for sharing cultural heritage. Early efforts for open access include Project Gutenberg, the American Memory Project, and the Universal Digital Library. Creative Commons was established in 2002 to encourage the publishing and reuse of creative works online through more flexible copyright policies.[35] That same year, Google began its project to digitize the world's books.[36] In 2006, the American Council of Learned Societies released a report titled "Our Cultural Commonwealth" calling for an investment in cyberinfrastructure for the humanities which "will benefit the public and the specialist alike by providing access to the breadth and depth of the cultural record."[37] The digitization of texts and archival collections brought with it new possibilities to apply computational tools to literature and historical analysis. The burgeoning field of Digital Humanities considers both how information technology can be applied to humanities research, and more broadly how humanities research methodologies can shape and be shaped by these information infrastructures.[38] As attention to and funding for Digital Humanities has grown along with the amount of digitized and open cultural heritage, more humanities disciplines are considering what role information technologies will play in their own work and have begun defining these resources as humanities data.

Much of the record of human cultural heritage is held within museum collections, so museums are implicated in calls for better access to cultural heritage and associated data. While many museums are working to digitize their collections, even this simple act comes with a lot of historical baggage around the history, identity, and definition of museums. Broadening access to collections online gives up some of the control museums have typically had over narratives about them and allows the public to interact with collections outside of a traditional museum context. Carl Hogsden and Emma K. Poulter, who organized the "Talking Objects" project at the British Museum, point out the ways in which digital collections can renew and reimagine the museum as contact zone by enabling multiple interpretations and understandings, even multiple surrogates of digital collections to exist simultaneously in connected networks.[39] They emphasize that the digital realm represents a new way to actively invite meaningful participation from diverse communities into the

---

[33] "Final NIH Statement on Sharing Research Data." *National Institutes of Health*, February 26, 2003. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html

[34] "OECD Principles and Guidelines for Access to Research Data from Public Funding" *Organization for Economic Co-operation and Development*, 2007. https://www.oecd.org/sti/inno/38500813.pdf

[35] "1.1 The Story of Creative Commons" *Creative Commons*, https://certificates.creativecommons.org/cccertedu/chapter/1-1-the-story-of-creative-commons/ , Accessed February 11, 2021.

[36] "Google Books History" *Google Books*, Archived on the Wayback Machine, https://web.archive.org/web/20160206043510/http://books.google.com/googlebooks/about/history.html , Accessed February 11, 2021.

[37] "Our Cultural Commonwealth". *American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* (American Council of Learned Societies, 2006), 2.

[38] Patrik Svensson and David Theo Goldberg "The Field of Digital Humanities", *Between Humanities and the Digital.* (MIT Press, 2015): p. 9-16.

[39] Carl Hogsden and Emma K Poulter, "The Real Other? Museum Objects in Digital Contact Networks" *Journal of Material Culture*, 17, no. 3 (2012): 265-286.

GO TO INDEX

museum. The movement around digital repatriation asks how digital museum collections can return important cultural knowledge, and control over its dispersal and representation, to original communities.[40] Repatriation advocates have cited the importance of digitizing collections to allow for more transparency about what heritage is currently within museum collections and to enable colonized nations to find that heritage.[41] Many will view the digitization of collections as a positive and necessary change, and for museums it represents a move towards the contested definition for museums in the twenty-first century.

A movement in digital humanities pushing museums further towards this new definition is the "Collections as Data" Initiative. In 2017, the IMLS funded Collections as Data National Forum released its "Santa Barbara Statement" outlining the ten principles of this initiative. The first principle states "Collections as data development aims to encourage computational use of digitized and born digital collections."[42] Cultural Heritage institutions are encouraged not only to digitize collections, but to do so in a way that enables them to be processed as data. In many ways this represents the intersection of the changing role of museums in the public sphere, the open data movement, and the digital humanities. When museums publish collections as data, they enable new digital humanities scholarship, invite an unprecedented level of engagement with their objects, and become more transparent institutions. While digitizing collections still allows some control over how these objects are presented and characterized, allowing for the large-scale analysis of these collections as data means truly inviting the public to interpret the collections in whatever way they see fit. This unprecedented transparency also enables a growing number of projects examining museum practices. This means, for example, articles pointing out the percentage of white male artists in museum collections.[43] At a time when museums are embroiled in controversies over their relationship to the public, the provenance of their collections, and their broader role in society, opening their collections to analysis invites further scrutiny.

Despite pushback from museums about their changing definition, major institutions such as the Getty, the Metropolitan Museum, and MoMA have made their collection data open through a combination of digital collections, API's, and published datasets on Github. In February of 2020, The Smithsonian premiered "Smithsonian Open Access", which released copyright restrictions from 2.8 million digital collection images and "nearly two centuries of data."[44] This new model takes the idea of public participation and dialog to a new level. Beyond presenting exhibits or educational products that inspire dialog, it invites the public to do what they will with the information without the interference of the museum. As the Smithsonian states in their press release, "With Smithsonian

---

[40] Joshua A. Bell, Kimberly Christen, and Mark Turin, "After the Return, Digital Repatriation and the Circulation of Indigenous Knowledge Workshop Report" *Museum Worlds*, 1, no. 1 (2013): 195-203. DOI: https://doi.org/10.3167/armw.2013.010112

[41] See Susan Douglas and Melanie Haynes, "Giving Diligence its Due: Accessing Digital Images in Indigenous Repatriation Efforts." *Heritage*, 2 (2019): 1260-1273. doi:10.3390/heritage2020081 and "Repatriation of Museum Objects" Webinar Recording, Cornell University Center for International Studies, October 23, 2020. https://einaudi.cornell.edu/discover/news/watch-repatriation-museum-objects-webinar-recording

[42] "The Santa Barbara Statement on Collections as Data" *Always Already Computational – Collections as Data*, https://collectionsasdata.github.io/statement/ , Accessed February 11, 2021.

[43] See Julia Halperin and Charlotte Burns, "Museums Claim They're Paying More Attention to Female Artists. That's an Illusion" *Artnet News*, September 19, 2019. https://news.artnet.com/womens-place-in-the-art-world/womens-place-art-world-museums-1654714; and Chad M. Topaz et al. "Diversity of artists in major U.S. Museums" *Plos One*, (March 20, 2019). https://doi.org/10.1371/journal.pone.0212852

[44] "Smithsonian Releases 2.8 Million Free Images for Broader Public Use." *The Smithsonian*, February 25, 2020. https://www.si.edu/newsdesk/releases/smithsonian-releases-28-million-free-images-broader-public-use

GO TO INDEX

Open Access, we're inviting people everywhere to make that knowledge their own—to share and build on our digital collections for everything from creative works, to education and scholarly research, to bold innovations we have yet to imagine."[45] Museums have increasingly factored the public into how they are run and what they exhibit. Now they are removing traditional barriers to collections access and are inviting the public to examine, research, and make sense of those collections on their own terms.

Art museums and their collections present an interesting case study for the advantages and challenges of cultural heritage collections as data. Their collections are closely tied to the developing Digital Humanities turn in art history and the rise of Provenance research. The Getty Provenance Index, which was a project that the curator of European paintings, Burton Fredericksen, began in the 1980's, is an early attempt to make provenance and collection data more available to a wider research community. In 2016, the Getty Provenance Index began a "remodel" to republish the index as Linked Open Data sets and began distributing completed sets on Github. This ambitious project reflected and continues to reflect a growing emphasis on provenance research in art history. This is partially due to demands for better transparency from museums around their collection practices. In fact, the Getty Museum has been involved in several controversies around Italian and Greek antiquities collected between 1977-1996.[46] Provenance research is associated, as well, with the growing trend of the Object Biography, a theory attributed to Igor Kopytoff in 1986 which argues for the importance of researching and publishing information about the lifecycle of objects through creation, exchange, and consumption. Art history, museum studies, and anthropology have embraced the process of studying the biography of an object through museum collection records and metadata. Object biographies and museum metadata have increasingly been associated with a movement to confront the colonial legacies of many universal and ethnographic Museums. For example, in 2001 Susan M. Vogel created the film "Fang: An Epic Journey", which traces the fictional biography of an African sculpture as it is removed from Africa and recontextualized through the art market and displayed in an art museum. The film is based on similar life histories of African objects in the Metropolitan Museum, where Vogel was a curator of African art.[47] The recently published book *The Brutish Museums: The Benin Bronzes, Colonial Violence, and Cultural Restitution*, by Dan Hicks, calls for the study and publication of "Object Necrographies", which draw attention to the silences in the 'lives' of objects, traced through museum provenance records, and the colonial legacies they represent.[48]

Digitized collection data, and specifically provenance data, has become the centerpiece of an emerging digital humanities shift in art history. In 2010, before the remodel began, Maximillian Schich and a team of colleagues proposed applying network science to the Getty Provenance Index data. The result was several maps of British, French, Belgian, and Dutch sales data from 1801-1820 presented at conferences at the National Gallery, London and the International Conference for Network Science in Berkeley in 2013.[49] In 2009, Beatrice Joyeaux-Prunel and

---

[45] Ibid.
[46] Hugh Eakin. "An Odyssey in Antiquities Ends in Questions at the Getty Museum." *The New York Times*, October 15, 2005. https://www.nytimes.com/2005/10/15/arts/design/an-odyssey-in-antiquities-ends-in-questions-at-the-getty-museum.html
[47] Susan M. Vogel "Fang: An Epic Journey" *Rand African Art*, (2001). https://www.randafricanart.com/Fang_an_epic_journey.html
[48] Dan Hicks. *The Brutish Museums: The Benin Bronzes, Colonial Violence, and Cultural Restitution.* (Pluto Press, 2020).
[49] Maximillian Schich, et al. "Network Dimensions in the Getty Provenance Index". *ArXiv: Physics and Society* (June, 2017).

GO TO INDEX

colleagues at the Ecole Normale Superior in Paris began the Artl@s project, which supports a spatial art history. The project advocates visualizing digital art historical and provenance data in mapping software to construct a more global and pluralistic network of artistic exchange.[50] Artl@s is an ongoing digital project that continues to compile art catalogs from the 19th and 20th Century in order to trace the internationalization of the art market during that period.[51] In 2015, IMLS funded the "Art Tracks" project by the Carnegie Museum of Art that developed a suite of open-source tools for turning art museum provenance records into data. They also developed the Digital Provenance Standard which provides guidance for making museum provenance data machine-readable.[52] In May of 2020, the first edition of the *Routledge Companion to Digital Humanities and Art History* was published with an entire section devoted to "Archives, Networks, and Maps" with several examples of projects that have used provenance data in art history to create maps and network diagrams.[53]

Practitioners of digital art history projects point out that the more collections and collections provenance data are digitized, the more it opens the profession up to a distance viewing research practice. Aligned with movements to reform the museum, these scholars are interested in broadening the scope of art history and shifting its center away from traditional Western points of focus. Instead of focusing on individual objects or actors of note, art historians using provenance data analysis can examine the broader movements, shifts, and networks of art history. This can mean that traditionally overlooked artists or movements are more likely to be found- as Harold Klinke states, Digital Art History brings the 'Great Unseen' to scholarly attention.[54] Joyeux-Prunel elaborates on this concept in discussing her own work, stating that with new trends in art historical mapping and digital art history, "Research has become more global and includes the 'peripheries' as well as the traditional 'centers'; more social, insofar as it takes into account groups and populations as opposed to canonical artists; and more art historical in its tracing of the circulation of artifacts and images in diverse contexts."[55] This trend in art history, which relies on the partnership of museums who make their collections and metadata open and available, builds on the larger movement for a focus on the contexts behind objects in museums, transparency about networks of collection and trade, and a shift towards more inclusive and complete narratives in the cultural information being conveyed by these institutions. Digital Art History projects also have the potential to visualize and clarify problematic provenance data, contributing to better transparency about museum collections and the art market. Jodi Cranston, who has built websites mapping the journeys of European artworks, states "Shameful pasts are easier to conceal in a list of text. Gaps in provenance are far more obvious when noted in a map where there is no place for a pin to land or where a painting essentially disappears from history because of the incompleteness of archival information that is typically relied upon to guarantee authenticity and originality."[56]

---

[50]Beatrice Joyeaux-Prunel. "Spatial Digital Art History: A Total Art History? – The Artl@s Project". *Visual Resources*, 29: 1-2 (2013): 47-58.
[51] "Data" *Artl@s*. https://artlas.huma-num.fr/map/#/ , Accessed February 1, 2021.
[52] "Overview" *Art Tracks*. Carnegie Museum of Art, http://www.museumprovenance.org/ , Accessed February 1, 2021.
[53] *The Routledge Companion to Digital Humanities and Art History.* Edited by Katheryn Brown (New York: Routledge, 2020). https://doi.org/10.4324/9780429505188
[54] Harold Klinke, "The Digital Transformation of Art History", *The Routledge Companion to Digital Humanities and Art History*, (New York: Routledge, 2020): 35.
[55] Beatrice Joyeux-Prunel, "Digital Humanities For a Spacial, Global History of Art", 88.
[56] Jodi Cranston, "Mapping Paintings, or How to Breathe Life into Provenance", *The Routledge Companion to Digital Humanities and Art History*, (New York: Routledge, 2020): 117.

GO TO INDEX

These projects center around the process of turning provenance records into computational data. This process is complicated, however, in the same way that the idea of "humanities data" is complicated. As Miriam Posner has stated, calling something data means that "its meaningful qualities can be enumerated in a finite list; that someone else performing the same operations on the same data will come up with the same results."[57] The truth is that something like provenance "data" is messy and the act of compiling it or collecting it requires a complex series of interpretive, subjective decisions. Eric Hormell, who currently works on the Getty Provenance Index project, described the complications of converting even one data value, sales prices, into linked open data. He explains that prices can be in different forms, written by hand and difficult to read, or reflect estimations instead of an exact price. Translating this into fixed data is a difficult and time-consuming task.[58] It also requires making decisions about what is included, how it is presented, and how missing information is conveyed (or whether it is conveyed) that inevitably represent an institutional perspective or worldview. Differences in how these decisions are made between institutions can make aggregating data from several sources difficult. Communities of Practice, such as OpenGLAM and LODLAM have emerged to shape policies to address and standardize these decisions in support of broader sharing of and access to collections data. The Carnegie Museum of Art provenance standard, which builds on the American Association of Museums' provenance standard, is designed to aid in resolving data ambiguities in provenance specifically. Acknowledging the need for new data models that better support event-based models for provenance, more closely aligned with the computational-friendly CIDOC-CRM ontology, the CMOA provenance standard offers a way to work with existing document-based provenance models. Its solutions are deceptively simple: for example, standardizing how to express period uncertainty (with the inclusion of the word "possibly"), where to begin provenance records (with "Artist as first period"), or providing a list of standardized terms for methods of acquisition.[59]

The difficult decisions involved in applying even these simple guidelines are especially evident when considering non-Western collections. The more institutions are asked to make complex data conform to a standardized, computational form, the more their perspective and worldview becomes expressed in that data. This can mean that, instead of revealing marginalized and diverse narratives, publishing and interpreting collections data risks further centering a Western, European worldview. In the "Mapping Senufo" project, art historians point out that the metadata in archival records is often socially constructed. They explore how the art historical label "Senufo", referring to a particular tribe in Africa, was in fact a label given to an artwork for the convenience of European art markets looking to credit a "creator". [60] Thus, an instruction as straightforward as listing the "artist as first period" can become complicated when applied to non-Western art forms. Mary Nooter Roberts points out that most African artworks taken by colonial officials were done so without the collection of any accompanying metadata. Even if art historians do attribute, after painstaking archival research, an artist- it is often not a traditional name, but a constructed title created by an historian or curator, like the "Warua Master".[61] Nooter goes on to

[57] Miriam Posner "Humanities Data: A Necessary Contradiction", *Miriam Posner's Blog*, June 25, 2015. https://miriamposner.com/blog/humanities-data-a-necessary-contradiction/
[58] Eric Hormell, "The Ripple Effect in Provenance Data Standardization", *Iris*, The Getty, August 21, 2018. https://blogs.getty.edu/iris/the-ripple-effect-in-provenance-data-standardization/
[59] "The CMOA Digital Provenance Standard" *Art Tracks*, Carnegie Museum of Art, October 20, 2016. http://www.museumprovenance.org/reference/standard/
[60] Susan Elizabeth Gagliardi, "Mapping Senufo" *The Routledge Companion to Digital Humanities and Art History*, (New York: Routledge, 2020).
[61] Mary Nooter Roberts. "The Naming Game, Ideologies of Luba Artistic Identity" *African Arts*, 31, no. 4, 56-73, 90-92.

GO TO INDEX

point out that, in the case of the Luba people of West Central Africa, the same weight is not given to identifying an artist because the importance of an object has always been its use, and not its creation.[62] These issues only become amplified with concepts of linked data, which rely on authority records. Determining what is considered an authoritative version of an object's title, its creator, or its culture is fraught with complicated power dynamics. Standardizing terms for "method of acquisition" is also especially loaded when applied to collections in colonial museums. Many museums will have to think carefully about how they would like to characterize the ways in which objects were collected by colonial officers.

Art museums hoping to publish their provenance data face an unenviably difficult task. Using archival documents and related materials to create datasets takes time and a lot of labor. Building these datasets means facing complex decisions about how to convey uncertain, contradictory, or missing information. Publishing these datasets means, often, exposing some skeletons in these museums' collection practices. Anne Luther outlines the five stars of making provenance data truly open access. At the lowest level, museums can simply digitize records, such as catalogs. To truly be part of the movement and achieve five stars, however, means putting in the labor to convert this material to a structured data format, linking data to other sources, and adding context.[63] Smaller museums who have a more limited budget may be hesitant to participate. Even larger museums may be tempted to wait until their provenance data is polished enough to publish. However, for institutions who prioritize connecting diverse communities with their collections to allow for meaning making to occur, it is a question of quantity and not quality: publishing a digitized version of catalog records is better than not publishing anything at all.

Museums do not have to take on this responsibility alone: digital art history scholars are an important and eager partner in this work. In defining the Collections as Data Imperative, Thomas Padilla challenges cultural institutions not to think of "target audiences", but to think about partnerships."[64] By making imperfect or messy provenance data sets open to digital art historians, museums can share some of the responsibility of clarifying them. For digital art history scholars, investigating, filling in, and problematizing missing data is central to their practice. The digital art history projects that have already taken place often center precisely around, as Cranston states, investigating and "filling in" gaps in the archival record and revealing where they represent a bias or worldview. While museums may be hesitant to share provenance data with large gaps, "As an academic, I see those gaps not as barriers but as invitations for interpretation and as symptoms of some related, yet distinct, historical event or moment."[65] Opening their imperfect provenance data to scholarly partners offers an opportunity for museums and scholars to build richer, more nuanced data sets over time.

While complicated, developing and adopting new metadata standards and ontologies that support the creation and exchange of collections provenance data, is a necessary step. The CMOA provenance standard represents an important development in this direction. However, critical work must take place to ensure that there is flexibility and diversity within these standards. Museums and museum professionals will have to consider and consult more diverse voices, definitions, languages, and uses in their development. Here, again, partnerships between museums and art history scholars offer rich opportunities for growth. One example of where this partnership has been effective is the Numisma Linked Data project. Scholars from the coins of the Roman Empire

[62] Ibid, 66.
[63] Anne Luther, "Digital Provenance, Open Access, and Data-Driven Art History", 450.
[64] Thomas Padilla, "On A Collections as Data Imperative" UC Santa Barbara (2017): 2.
[65] Jodi Cranston, "Mapping Paintings", 117.

community developed a standardized metadata set with discipline-specific terms to address the need for linked open data sets of digitized coins. For this scholarly community, tracing the dispersal of these coins and their uses within different contexts was vital for research- creating standardized metadata for publishing roman coin collections allowed scholars to trace the movement of coins internationally.[66] Scholars continue to suggest how these vocabularies should be expanded, which is essential when considering such an international network of owners and contexts. Through its website, Numisma actively invites feedback on its standards and new contributions to its vocabularies. Of course, there is a precedent for partnership with scholarly communities in developing standards. The Getty Vocabularies, a major source of authority records, welcomes the contributions of independent scholars as an important part of its development. As art historians make data analysis a more essential aspect of their work, they will become even more vital resources for identifying gaps in existing standards and vocabularies and offering a broader perspective for enabling discoverability on an international scale.

Part of a more open access practice, and embracing collections as data, is also considering how information about these collections can be disseminated online and in the museum. Digital art history projects have shown the potential for better methods of communicating complex metadata in a way that invites further questions and open dialog. These methods extend beyond the traditional exhibit or museum catalog. Digital art historians are considering how new tools can breathe life into art historical data. In describing a project that digitally projected a recreation of a Studiolo based on archival research, Stefania De Vincentis and Luca Nicolo Vascon write, "By initiating a dialogue between the academic study of the history of the artwork and current museum practices, DAH [Digital Art History] can focus on the purposeful dissemination of research results and configure a novel space of interaction between the artwork, scholars, and visitors".[67] The Mapping Senufo project aims to build a relational database and layered map to reveal and complicate the positionality of assigning Senufo as a marker of identity in art historical metadata. The team plans to publish findings as a digital publication, enabling interactivity. Readers of the publication will be able to "isolate and investigate particular objects, images, people, places, and events integral to the ongoing definition of a single category of art." In doing so, "Mapping Senufo will thus embody complexity and contingency at the core of identities, style labels, and knowledge construction."[68] At the "Art History in Digital Dimensions" conference sponsored by the Getty Research Institute and the Samuel H. Kress Foundation, participants also voiced the hope that Digital Art History could help build a more accessible Public Art History that considers and includes a broader audience.[69] Digital art historians as partners offer new ways to present collections and data about collections. Opening collections to these projects can expand the possibility for knowledge distribution and engagement with the public.

For both museums and art historians, this new type of scholarship means learning the conventions and best practices of the larger data community. Museums and humanities scholars who want to build and share datasets must know how to manage, document, and preserve that data. Increasingly, as more scholars take on the work of building provenance datasets that represent global networks of cultural exchange, they will have to think about how to allow their colleagues to use and build off their data. This will require an even more active participation in the

[66] Robert Wellington, "Metadata, Material Culture, and Global Art History", 330.
[67] Stefania De Vincentis and Luca Nicolo Vascon, "Digital Languages for Art History, Audience Engagement, Virtual and Augmented Reality", 276.
[68] Susan Elizabeth Gagliardi, "Mapping Sefufo", 137.
[69] Stephen Bury, et al. "Art History in Digital Dimensions, The White Paper" *Digital Art History Lab Committee*, Frick Art Reference Library, February, 2017, 6.

open data movement and critical engagement with traditionally scientific practices, such as FAIR (Findeable, Accessible, Interoperable, Reusable) data principles. Thomas Padilla points to issues in lack of documentation for library collections data that prevent this open exchange. He cites the vital importance of making it clear how the data was cleaned and processed, what schema were used, the purpose behind choices for which collections are digitized, and data quality overall.[70] For museum provenance data, which is the product of a myriad of choices and often includes imperfect and incomplete datasets, this documentation is especially important. While museums may not have to wait until their data is perfect and "complete" to publish it, they do have to ensure that every choice made in preparing that data, the stakeholders and authors involved in the dataset, any associated schema or standard consulted, and potential imperfections in the data are well documented and readily available to a potential user. Scholars and museums will also have to turn a critical eye to how data is collected and visualized. As Harold Klinke points out, they must develop "data literacy", or broader critiques of how data was gathered and potential biases, the algorithms used to gather that data, and the visualizations used to represent that data.[71] Johanna Drucker is often cited when discussing the dangers of simply applying visualization tools created for quantitative data to humanities data. She explains that these visualizations often carry assumptions about data as "given", a realist perspective "fundamentally at odds with approaches to humanities scholarship premised on constructivist principles."[72] Scholars and museums who visualize provenance data must be critical of themselves and their colleagues, pushing the presentation of this data to reflect the more constructivist, humanistic lens traditional to art history. They must adopt scientific methods and question them, moving back and forth between qualitative and quantitative analysis to "make research of a particular subject more systematic or precise, while at the same time complicate the very process of classification and quantification."[73]

For museums and for art historians, to engage in the collections as data movement requires potentially major shifts in their current practices. Museums will have to make difficult decisions about allocating staff and budgets to undertake this work. It may mean a fundamental reordering in institutional priorities for education and exhibition. For art historians to be active participants in this work, they will have to expand the definition of art history research. Researchers will need to learn about data mining, analysis, and visualization rather than or alongside more traditional uses of archival databases. Both will have to learn about, and engage in, data literacy. Museums and scholars will have to learn best practices for ensuring their data is legible, findable, and reusable. They will have to be active in creating, editing, and implementing metadata standards.

This does not mean that the risks or challenges outweigh the potential rewards of this work. Publishing collections as data presents an opportunity for museums to embrace the twenty-first century role advocated by the new museum definition. By taking on the challenge to make not only their collections available for wide public use, but data about those collections, museums are creating space for a more participatory and collaborative public. Allowing the scrutiny and complicated politics that may result from this transparency will not be comfortable but could foster a meaningful interaction with contemporary issues in social justice, decolonization, and equity. For art historians, engaging in this work could facilitate the presentation of more diverse, overlooked narratives in art history. It could expand the field and pave the way for new research questions. The

[70] Thomas Padilla, "On a Collections as Data Imperative", *Scholarly Communications*, 79, no. 6 (2018): 3.
[71] Harold Klinke, "The Digital Transformation of Art History", 38.
[72] Johanna Drucker, "Humanities Approaches to Graphical Display", *Digital Humanities Quarterly*, 5, no. 1 (2011). http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html
[73] Miriam Kienle, "Between Nodes and Edges: Possibilities and Limits of Network Analysis in Art History." *Artl@s Bulletin*, 6, no. 3 (2017): 6.

GO TO INDEX

partnership between the two could enable a new level of access to museum collections and a critical engagement with the worldviews and biases implicit within them. As more museums develop their digital collections and more data becomes available to art historians, the question may become not “whether”, but “how soon” will these changes occur. At this point an active partnership will be important not only in carrying out the work, but doing so in a way that is careful, critical, and inclusive.

GO TO INDEX

Works Cited

Bell, Joshua A., et. al. "After the Return, Digital Repatriation and the Circulation of Indigenous Knowledge Workshop Report" *Museum Worlds*, 1, no. 1 (2013): 195-203. DOI: https://doi.org/10.3167/armw.2013.010112

Bennett, Tony. "The Exhibitionary Complex" *New Formations*, 4 (1988): 73-102.

Bury, Stephen et al. "Art History in Digital Dimensions, The White Paper" *Digital Art History Lab Committee*, Frick Art Reference Library, February, 2017.

"The Challenge of Revising the Museum Definition" *International Council of Museums*, https://icom.museum/en/news/the-challenge-of-revising-the-museum-definition/ , Accessed February 6, 2021.

Clifford, James. "Museums as Contact Zones." *Routes, Travel, and Translation in the Twentieth Century.* London: Harvard University Press, 1999.

Cranston, Jodi. "Mapping Paintings, or How to Breathe Life into Provenance", *The Routledge Companion to Digital Humanities and Art History*. New York: Routledge, 2020.

Douglas, Susan and Melanie Haynes, "Giving Diligence its Due: Accessing Digital Images in Indigenous Repatriation Efforts." *Heritage*, 2 (2019): 1260-1273. doi:10.3390/heritage2020081

Drucker, Johanna. "Humanities Approaches to Graphical Display", *Digital Humanities Quarterly*, 5, no. 1 (2011). http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html

Eakin, Hugh. "An Odyssey in Antiquities Ends in Questions at the Getty Museum." *The New York Times*, October 15, 2005. https://www.nytimes.com/2005/10/15/arts/design/an-odyssey-in-antiquities-ends-in-questions-at-the-getty-museum.html

Gagliardi, Susan Elizabeth. "Mapping Senufo." *The Routledge Companion to Digital Humanities and Art History*. New York: Routledge, 2020.

Halperin, Julia and Charlotte Burns, "Museums Claim They're Paying More Attention to Female Artists. That's an Illusion" *Artnet News*, September 19, 2019. https://news.artnet.com/womens-place-in-the-art-world/womens-place-art-world-museums-1654714

Hicks, Dan. *The Brutish Museums: The Benin Bronzes, Colonial Violence, and Cultural Restitution.* Pluto Press, 2020.

Hogsden, Carl and Emma K Poulter, "The Real Other? Museum Objects in Digital Contact Networks" *Journal of Material Culture*, 17, no. 3 (2012): 265-286.

Hormell, Eric, "The Ripple Effect in Provenance Data Standardization", *Iris*, The Getty, August 21, 2018. https://blogs.getty.edu/iris/the-ripple-effect-in-provenance-data-standardization/

GO TO INDEX

"ICOM announces the alternative museum definition that will be subject to a vote" *International Council of Museums*, https://icom.museum/en/news/icom-announces-the-alternative-museum-definition-that-will-be-subject-to-a-vote/ , Accessed February 5, 2021.

Joyeaux-Prunel, Beatrice. "Spatial Digital Art History: A Total Art History? – The Artl@s Project". *Visual Resources*, 29: 1-2 (2013): 47-58.

Kienle, Miriam. "Between Nodes and Edges: Possibilities and Limits of Network Analysis in Art History." *Artl@s Bulletin*, 6, no. 3 (2017).

Klinke, Harold. "The Digital Transformation of Art History." *The Routledge Companion to Digital Humanities and Art* History. New York: Routledge, 2020.

"OECD Principles and Guidelines for Access to Research Data from Public Funding" *Organization for Economic Co-operation and Development*, 2007. https://www.oecd.org/sti/inno/38500813.pdf

"Our Cultural Commonwealth". *American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences.* American Council of Learned Societies, 2006.

Padilla, Thomas. "On a Collections as Data Imperative", *Scholarly Communications*, 79, no. 6 (2018).

Posner, Miriam. "Humanities Data: A Necessary Contradiction." *Miriam Posner's Blog*, June 25, 2015. https://miriamposner.com/blog/humanities-data-a-necessary-contradiction/

"Repatriation of Museum Objects" Webinar Recording, Cornell University Center for International Studies, October 23, 2020. https://einaudi.cornell.edu/discover/news/watch-repatriation-museum-objects-webinar-recording

Schich, Maximillian, et al. "Network Dimensions in the Getty Provenance Index". *ArXiv: Physics and Society* (June, 2017).

"Smithsonian Releases 2.8 Million Free Images for Broader Public Use." *The Smithsonian*, February 25, 2020. https://www.si.edu/newsdesk/releases/smithsonian-releases-28-million-free-images-broader-public-use

Svensson, Patrick and David Theo Goldberg "The Field of Digital Humanities", *Between Humanities and the Digital.* MIT Press, 2015: p. 9-16.

"The Santa Barbara Statement on Collections as Data" *Always Already Computational – Collections as Data*, https://collectionsasdata.github.io/statement/ , Accessed February 11, 2021.

Topaz, Chad M. et al. "Diversity of artists in major U.S. Museums" *Plos One*, (March 20, 2019). https://doi.org/10.1371/journal.pone.0212852

Vogel, Susan M. "Fang: An Epic Journey" *Rand African Art*, (2001). https://www.randafricanart.com/Fang_an_epic_journey.html

GO TO INDEX

ELECTIVE WORK

SUGGESTIONS FOR METADATA CLEANING FOR SEQUOIA AND KINGS CANYON ROBERTS AND HISTORIC COLLECTIONS

Spring 2020
IS 464 Metadata
Professor Melissa Gil

Prompt: For your final assignment, create a comprehensive metadata strategy for a digital resource with a focus on serving your end-users that maps and remediates metadata from at least two collections or repositories.

Your assignment must address the following: Collections and repositories, intended users and provide an example use case, a simple conceptual data model, an example of a resource description issue, identify controlled vocabularies to be used, metadata schema to be used, how your resources will be discoverable, rights metadata issues (if relevant), whether or not you will use crowdsourcing or machine learning to enhance meatadata, search and browse functionalities, workflow and project management

ABSTRACT

Sequoia and Kings Canyon National Park (referred to here as SEKI) has several digitized and digital collections of images, which it has stored and made available through the NP Gallery portal. NP Gallery is a digital asset management system that was designed to aggregate digital and digitized collections from the National Parks. SEKI has four digitized collections which it has made available or will make available through NP Gallery: The Roberts Collection, Historic Images, Resource Management Slides, and Civilian Conservation Corps. For the purposes of this project, I will be focusing on the Roberts Collection and Historic Images. Initial metadata creation for these collections on NP Gallery was done by a small staff with limited time, thus metadata for these collections succumbs to "Lag": The dissemination of metadata is not necessarily synchronized with the dissemination of the object to which it applies. NP Gallery uses a local schema for metadata terms and arrangement. SEKI's use of these terms is not consistent and should be clarified and standardized across collections. Description of objects within these elements is often incomplete, with vital information about the contents of photographs or their date of creation absent. SEKI would like to increase the discoverability of its collections by allowing California Digital Library to harvest it and include it in Calisphere. Before this process can be completed, metadata must be cleaned and enhanced both to help with the harvesting process and to increase access and discoverability both in NP Gallery and on Calisphere.

GO TO INDEX

TEXT

Collections Background:

Sequoia and Kings Canyon National Park (referred to here as SEKI) has several digitized and digital collections of images, which it has stored and made available through the NP Gallery portal. NP Gallery is a digital asset management system that was designed to aggregate digital and digitized collections from the National Parks. SEKI has four digitized collections which it has made available or will make available through NP Gallery: The Roberts Collection, Historic Images, Resource Management Slides, and Civilian Conservation Corps. For the purposes of this project, I will be focusing on the Roberts Collection and Historic Images.

The Roberts Collection consists of materials belonging to Henry E. Roberts, who was photo concessioner for General Grant National Park between 1900-1930. Included in the collection are glass plates, prints, stereographs, slides, and correspondence relating to the history of SEKI ranging from 1888-1960. The materials digitized and available on NP Gallery are the Slides and Prints Series in the collection.

Historic Images represent a collection of images from the early history of the park taken by a variety of photographers, some of whom are unknown. They represent subjects related to park operations such as buildings, events and museum collections.

Project Overview:

Initial metadata creation for these collections on NP Gallery was done by a small staff with limited time, thus metadata for these collections succumbs to what describe as "Lag"- "The dissemination of metadata is not necessarily synchronized with the dissemination of the object to which it applies".[74] NP Gallery uses a local schema for metadata terms and arrangement. SEKI's use of these terms is not consistent and should be clarified and standardized across collections. Description of objects within these elements is often incomplete, with vital information about the contents of photographs or their date of creation absent. SEKI would like to increase the discoverability of its collections by allowing California Digital Library to harvest it and include it in Calisphere. Before this process can be completed, metadata must be cleaned and enhanced both to help with the harvesting process and to increase access and discoverability both in NP Gallery and on Calisphere.

Recommendations for enhancing this metadata have been made based on access to the front-end user experience of NP Gallery and available metadata. They will be amended once more information is available.

Intended Audience/ Use Case:

The Roberts and Historical Images Collections have three principle uses.

- NP Gallery is useful for potential or past visitors to the park who want to learn more about it to plan their visit or who have a connection with the park and would like to understand more about its history. For example, someone who had hiked in a part of SEKI may want to look through the Historic Images or Roberts Collection files to understand how related landmarks have changed over time.

[74] Bruce, Thomas R. and Diane I. Hillman. "The Continuum of Metadata Quality: Defining, Expressing, Exploiting". *Metadata in Practice* American Library Association (Chicago, 2004), 238-256.

GO TO INDEX

- The National Parks collect material and have a large following from enthusiasts for nature and nature photography. Both collections offer beautiful examples of historical park photography.
- Both collections also document elements of California's history, especially the early history of logging in California.

For these three use cases, it is important that the associated metadata include relevant information about where the photographs were taken, any contextual information about events, landmarks and people featured in photographs, the dates of creation and metadata about the original format of the digitized file.

Existing Schema:

The metadata schema in NP Gallery has its own set of terms. The intended use for these terms is open-ended and both collections have decided to use some of them in different ways. For example, "Description" is used by the Roberts Collection for comments about processing the images, which I believe are in Henry E. Roberts' own words and relate to the margins and exposure time. Historical Images has used this section to define the collection, populating "Description" with the phrase "SEKI Historical Images". In another example, collections have chosen a different term to describe the image creator: The Roberts Collection has chosen the term "PhotoCredit" while Historical Images uses the term "Photographer". Other uses of terms seem inappropriate or like a poor fit for their content. For example, the Roberts Collection uses "AtlText" to provide information about the Collection and Series. Finally, some metadata is repeated across several terms. This is especially true in historical collections. In some instances, the same information included in the "title" is included again in "Comment", "Image Details" and "Locations". As Bruce and Hillman describe, this means the metadata lacks logical consistency both within each collection and across the two collections. "Users expect to be able to search collections of similar objects using similar criteria, and increasingly they expect search results and indicative indexes to have similar structures and appearance."[75]

Figure 1 is a data dictionary with the current terms used and the ways each collection is defining and populating them. In order to provide the user with a coherent and logically consistent set of metadata, Figure 2 is a proposal for how SEKI should define terms for its metadata across collections, how they would map to the Calisphere schema and whether terms are relevant for the Calisphere catalog. The proposed changes should allow for a more accurate set of metadata elements which more logically describe the collections. For example, the creator of the photograph is uniformly referred to as "PhotoCredit". "Comment" is used to retain the original language of the cataloger and of Henry E. Roberts in describing the images and "Description" contains a more accessible and informative piece of information about the contents of the image. "Source Information" is added to include information about the original type of material digitized (Negative, Print, Glass plate, etc.). The misleading "Camera Information" element is replaced with the more accurate "Scanner Information". "Alt Text" has been removed- series and subseries information may be helpful to include under an element such as "source location" if available. This information is only useful to a user if properly contextualized by its schema element. Series and subseries are also used to populate keywords, which helps users to locate materials related to images they are interested in.

Metadata Content/ Resource Description Issues:

[75] Bruce, Thomas R. and Diane Hillman

The Roberts and Historic Images collections were digitized and added to NP Gallery for discovery using a limited staff and with incomplete source metadata. Thus, metadata for images in both the Roberts Collection and the Historic Images Collection is often incomplete and inaccurate. While metadata for these collections will continue to be limited by a lack of source information, there are ways that the content of metadata elements for both collections can be enhanced to increase their usefulness and relevance.

There are some simple inconsistencies that can be easily corrected in the content metadata. The format of item titles is not consistent across collections. For example, in the Robert's Collection, specific waterfalls are given their full name ("Tornado Creek Falls"), while unknown waterfalls are labeled "falls", "Falls", or "Waterfall". Some animals are labeled "deer", "Deer 2" or qualified with an action "Feeding Deer" or "Woman feeding deer". Consistency in capitalizations and use of numbers could and should be corrected. Titles in Historic Images are grouped into general series ("Vehicles and Equipment", "Aerial Views", "Buildings and Utilities") followed by more specific information ("Vehicles and Equipment, Trucks, Ash Mountain Headquarters"). These titles have many inconsistencies in the way they are formatted – which words are capitalized, the use of commas vs. periods, whether the series information goes before or after the location of the photograph. These should also be corrected and unified. Assuming that these titles were devised by catalogers adding metadata to digitized images, titles should be edited to be more informative and uniform. Cataloging of Cultural Objects (CCO) offers some guidance on how to construct titles for objects that have no existing one. Many of these titles are long and descriptive, CCO emphasizes the need for more concise titles and recommends that descriptive information be included in descriptions or subjects, instead.[76] Titles from the Roberts Collection could be more descriptive. For example, "Falls" could be qualified with a type of fall, its location or name.

There should also be better conformance with metadata value standards within metadata records. Create Date in the Roberts Collection should conform to the standard date format used by the Historic Images collection (mm/dd/yyyy) instead of the form it is currently using (yyyymmdd). Time Submitted, Last Edited, Embedded Timestamp and Metadata date all use different date formats, as well, and should be unified. Geographic Locations should also be unified and clarified. The Roberts Collection gives the latitude and longitude for both Kings Canyon National Park and Sequoia and Kings Canyon National Park. Historic Collection does not give any coordinates, but does include what "locations" are present in the photograph along with the county information and state for Sequoia and Kings Canyon. Both collections must determine what location information to include and which order. In their Geographical and Temporal Guidelines for mapping, Digital Public Library of America recommends Country, State, City, County and Coordinates.[77]

Much of the material described in both collections contains names and terms specific to Sequoia and Kings Canyon National Park and is, therefore, not described by controlled vocabularies, authority files or thesauri. However, there needs to be control for how certain subjects or terms are used across these collections. For example, both "Sequoia & Kings Canyon" and "Sequoia and Kings Canyon" are used. The collections are titled the abbreviated form of the park name: "SEKI Historical Images" and "SEKI Roberts Collection" instead of using the full name, which can make browsing for these collections difficult: users may not know that SEKI is synonymous with Sequoia and Kings Canyon. I propose listing the names of these collections as "Sequoia and Kings

[76] "Rules for Title" *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images.* Edited by Murtha Baca, et al. American Library Association (Chicago, 2006): 58.
[77] "DPLA Geographic and Temporal Guidelines for MAP 3.1". *Digital Public Library of America.*(December, 2017). http://bit.ly/dpla-geo-styleguide-3_1

Canyon Roberts Collection" and "Sequoia and Kings Canyon Historic Images". Some cultural terms used are outdated and should be improved with a current preferred term. For example, the word "Indian" is used instead of "Native American" for the title "Indian Baskets". In this example, the use of the preferred term used in both the Library of Congress Subject Headings and in the Getty Art and Architecture Thesaurus would be "Native American Baskets". Overall, creating a local controlled vocabulary for SEKI names, terms and locations will help with consistency across metadata.

Finally, inaccuracies and lack of detail should be corrected in the metadata. "CreateDate" in the Roberts Collection is not only in a different format, but is an incorrect date from 2018. While the exact year of the creation of each negative may be unknown, it would be more helpful to include the date range of the collection (1888-1950) or simply, "Unknown". I would argue that in this case some information, even a wide date range, is better than no information. The descriptions used in each collection should be edited to include more useful information. The descriptions in the Roberts Collection include notes from Henry E. Roberts about how negatives should be processed or about the contents of an image. Descriptions in the Historic Images simply state "SEKI Historic Images". In order to be more clear and useful to a user, descriptions should be rewritten to include any known details about the image in clear language. Information can simply be taken from the Title, Comment, Original Description or from the Finding Aid associated with the collection. The information currently populating the "Description" element should then be moved to more appropriate elements in the metadata schema.

Discoverability:

The use of "Keywords" should help users browse for certain subjects within NP Gallery and within the collection. The current keywords used, however, are not necessarily useful. For example, the keywords "Water" and "Building" lack specificity. The keyword "H. E. Roberts" is used instead of "Henry E. Roberts" to describe the Roberts Collection, which is inconsistent with how his name is written throughout the rest of the catalog. "Box 1" may help link objects in a particular location in the archive, but are not helpful as a keyword for searching. Keywords are essential in NP Gallery. A user logging in is prompted to use a keyword search to find any item, which they can refine by "state" and "Park". It is important that keywords include the controlled versions of the object title, the creator, the collection, the series and any names or landmarks included in the description. They should contain determined preferred terms and common equivalents ("Henry E. Roberts", "H.E. Roberts", and "Roberts" or "National Park Service Rustic" and "Parkitecture").

A user logging in to NP Gallery who wants to find material only related to Sequoia and Kings Canyon can search by "Park". However, this brings up any image from the park in no particular order- not organized by collection, subject or time period. If a User wants to filter results, they can only do so by "Categories" (Scenic, Historic, Events, etc.), "NPS Units and Sites", which includes all National Parks referenced in the collection, or "File Type" (Image, Document or Album). Within NP Gallery, better use of Keywords and the Albums feature can help organize items from the same collection and series to help with discovery. If a user wants to search for a "Historic" image or document, an entry they find useful should have keywords that point to related images and should be part of an album that reflects its collection and series.

Aggregation with Calisphere is an exciting way for Sequoia and Kings Canyon to increase discovery of its collection, and this is why quality metadata is important. By clarifying subject, description, locations, create date and by creating more concise and descriptive titles, SEKI can

GO TO INDEX

maximize access to their digital and digitized contents both in Calisphere and for users directed back to NP Gallery.

Rights Metadata:

Calisphere and its associated organizations (The California Digital Library and the Digital Public Library of America) have helped develop Rightsstatements.org in order to unify the rights information being submitted by its partner institutions. According to their use of Rightsstatements.org, partners are encouraged to submit a URI indicating whether the work is under copyright restrictions or not and a Rights Statement or narrative that clarifies any specifics of this status.

Currently, most work in both collections are part of the Public Domain. This is currently listed under "Constraints Information". Each collection has its own Rights Statement under "Copyright", which – using different language – clarifies that "Restrictions apply on use and/or reproduction" or that "Permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this website. Digital assets without any copyright restrictions are public domain." This language is confusing, especially when constraints information for these works is listed as "Public Domain", which typically means that there are no copyright restrictions on the works listed. If there are restrictions on the use of these items even though they are in the Public Domain, these restrictions should be clear in the Copyright Statement. If these restrictions only apply to certain items in the collection which are copyrighted, this should also be clear to the user. Finally, users should be directed to where to find more information about restrictions, either through a link to a page with more details – as recommended by rightsstatements.org[78] or by directing the user to the contact information included in the metadata under "Contacts". I have included a sample statement in Figure 2 under "Constraints Information".

Workflow/ Project Management:

I believe that improving the quality of this metadata is important, especially as Sequoia and Kings Canyon move to make their collections more available through Calisphere. I also recognize that it requires a lot of work. There are some ways that can help streamline the process. Metadata can be taken from the front end and converted into a .csv format through Simple Text and updated to OpenRefine for bulk editing. Alternatively, metadata already compiled in a .csv file can easily be updated and edited and will streamline the process even further. Object IDs from NP Gallery should also be matched with any internal archive record ID's to avoid confusion when exporting and importing data. Once metadata has been cleaned through OpenRefine, my hope is that it can be updated in bulk. If not, OpenRefine will still make individual changes to records easier to manage and aggregating metadata will help identify more large-scale inconsistencies. Minor changes to formatting, such as unifying the date format, are easy to make and can help foster trust from the user in the accuracy of the metadata. More complicated changes- such as rewriting the titles and descriptions- can be done simply using existing metadata included in titles, comments, image details and keywords and expanded later through research, if necessary.

I understand that initial NP Gallery metadata was limited by the amount of information available in the physical archive. To continue improving metadata for these collections in the future, crowdsourcing could be used to provide more detailed descriptions of many of the landmarks included in the Roberts Collections. For example, visitors to the park may be able to help identify

[78] "Contractual Restrictions". *Rightsstatements.* Rightsstatements.org. https://rightsstatements.org/page/1.0/?language=en#collection-nc

GO TO INDEX

waterfalls or other landmarks featured in photographs, which would be time-consuming for an individual cataloger to research. Crowdsourcing could also help to improve coordinates of landmarks included in object records, which would be valuable information for many users who would like to locate these areas of the park during their visit.

GO TO INDEX

Label	Description	Example	Appears in Both?
Asset ID	The unique digital hash of the photograph	e773efdbb4424875a2e1300dc90daaec	Yes
Title	Title of digital object, describes contents of photograph	Falls (Roberts Collection); Buildings and Utilities, tent frame seasonal housing (Historical Images)	Yes
AltText	Information about archive collection and series related to image	SEKI Roberts Collection Series: 2 Negatives; 2 Falls	No, only Roberts Collection
Description	Not standardized. Varies from information about type of resource ([print]) to technical information about photograph (Roberts Collection); labels the collection (Historic Images)	North Grove, 13x17 - contrast PMC, Expose full neg. to 11/2 min, shade big tree 1/2 min longer Then shade center ol g - ?, 0+ trees 1 1/2 min -, Exposure 3 1/4 min altogether, 2 - 6 1/2 x 81/2 will borders, 2 - 5x7 withought " " (Roberts Collection); SEKI Historical Image (Historical Images)	Yes
Comment	Notes from cataloger	Unknown Photographer, Lodgepole, SNP, Buildings and Utilities, tent frame seasonal housing. 490100	No, only Historical Images
Image Details	Information about archive series related to image (Roberts Collection); Alternative version of image title (Historic Images)	Series: 2 Negatives; 2 Falls (Roberts Collection); Buildings and Utilities, tent frame seasonal housing (Historical Images).	Yes
Publisher	Claims ownership of National Park Service	U. S. National Park Service	No, only Roberts Collection
Copyright	Rights Statement	Restrictions apply on use and/or reproduction (Roberts Collection); Permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this website. Digital assets without any copyright restrictions are public domain (Historical Images)	Yes
PhotoCredit	Image creator	Henry E. Roberts	No, only Roberts Collection
Photographer	Image creator	Unknown Photographer	No, only Historical Images
Constraints Information	Rights Statement	Public Domain	Yes

GO TO INDEX

Keywords	Hyperlinks to image collections that contain related keywords, describe archival location (Box 1), contents of photograph and types of resource (negative)	Falls , H. E. Roberts , Kings Canyon National Park , Negatives , Roberts Collection , Sequoia National Park , Water	Yes
Resource Type	Describes the category of object the resource falls into	Image	Yes
NPS Units	Park location of collection and abbreviation	Sequoia and Kings Canyon National Parks, Code: SEKI Kings Canyon National Park, Code: KICA	No, only Roberts Collection
Locations	The geographic Coordinates of the image (Roberts Collection); More specific information about landmarks or locations in the image (Historic image)	Kings Canyon National Park, California Latitude: 36.8744888305664, Longitude: -118.593803405762 Sequoia and Kings Canyon National Parks, California Latitude: 36.7172584533691, Longitude: -118.537902832031 (Roberts Collection); Lodgepole ; tent frame seasonal housing, Sequoia National Park, Tulare County, California (Historical Images)	Yes
Related Collections	Name of Collection as hyperlink to collection in NP Gallery	SEKI Roberts Collection; SEKI Historic Images	Yes
Create Date	Date original (not digital) image was created- Roberts Collection has non-standardized number format	20180221 (Roberts Collection); 01/01/1949 (Historical Images)	Yes
Contacts	Contact information for the curator of the collection	Person: Curator Organization: US National Park Service Position: SEKI Resource Management Address: Sequoia and Kings Canyon National Park, Three Rivers CA Email: seki/superintendent@nps.gov	No, only Roberts Collection
Intended Audience	Refers to a cataloging category that is unclear	Option 3	No, only Historical Images
Time Submitted	Date digital copy was added to repository	Wednesday, January 11, 2017 11:22:50 AM	Yes

GO TO INDEX

Last Edited	Date metadata of digital copy was last changed	Thursday, June 6, 2019 10:37:17 AM	Yes
Original File Name	Digital image file name	00075.tiff	Yes
Camera Information	Information about camera used to create image	EPSON EPSON, Perfection V700/V750	No, only Historical Images
Embedded Timestamp	Date digital image was created? (all timestamp dates have the same date and format)	Monday, January 1, 0001 12:00:00 AM	No, only Historical Images
Resource Format	Type of File	Tiff or Jpeg	Yes
File Size	Expressed in bytes	35.9 MB	Yes
Metadata Date	Date original metadata was created?	2015-04-16T11:33:51-07:00	No, only Historical Images
Rating	? Based on a five star system	★★★☆☆	Yes
Categories	Describes the types of content in the image	Scenic, Historic	Yes
Related Portals	Links to Sequioa and Kings Canyon NP Gallery portal	Sequoia & Kings Canyon National Parks	Yes

Figure 1

Label	Calisphere Term	Description	Example	Appears in Calisphere?
Asset ID	Identifier	The unique digital hash of the photograph	e773efdbb4424875a2e1300dc90daaec	Yes
Title	Title	Title of digital object, describes contents of photograph	Falls (Roberts Collection); Buildings and Utilities, tent frame seasonal housing (Historical Images)	Yes
Description	Description	Additional information about contents of the image, stated concisely and clearly	Image of waterfall from the early development of Sequoia and Kings Canyon (Roberts Collection) Lodgepole Campground, tent frame seasonal housing (Historic Images)	Yes
Comment	N/A	Notes from original catalog entry	North Grove, 13x17 – contrast PMC, Expose full neg. to 11/2 min, shade big tree ½ min longer Then shade center ol	No, Description has enough information

GO TO INDEX

			g - ?, 0+ trees 1 ½ min -, Exposure 3 ¼ min altogether, 2 – 6 ½ x 81/2 will borders, 2 – 5x7 withought " " (Roberts Collection); Unknown Photographer, Lodgepole, SNP, Buildings and Utilities, tent frame seasonal housing. 490100 (Historic Images)	
Publisher	Publisher	Claims ownership of National Park Service	U. S. National Park Service	Yes
Copyright	Rights	Whether the item is under copyright restrictions	Public Domain	Yes
PhotoCredit	Creator	Image creator	Henry E. Roberts	Yes
Constraints Information	Rights	Rights Statement	Digital assets without any copyright restrictions are public domain and can be used freely. Permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this website. Please see "Copyright" and "Contacts" for more information.	Yes
Keywords	Subject	Hyperlinks to image collections that contain related keywords created by crowdsourcing or from original metadata	Negatives, Falls, Henry E. Roberts, Roberts, Waterfall (Roberts Collection); Buildings and Utilities, Tent Frame Seasonal Housing, Buildings, Housing, Tent Frame, National Park Service Rustic, Parkitecture (Historic Images)	Yes, for browsing/ discovery
Resource Type	Type	Describes the category of object the resource falls into	Image	Yes
NPS Units	N/A	Park location of collection and abbreviation	Sequoia and Kings Canyon National Parks, Code: SEKI Kings Canyon National Park, Code: KICA	No, only relevant for internal Park knowledge
Locations	Place	More specific information about landmarks, Vegetation Zones, Trails or buildings in the image	Kings Canyon National Park, California Latitude: 36.8744888305664, Longitude: -118.593803405762 Sequoia and Kings Canyon National	Yes

GO TO INDEX

			Parks, California Latitude: 36.7172584533691, Longitude: -118.537902832031 (Roberts Collection); Lodgepole ; Tent frame seasonal housing, Sequoia National Park, Tulare County, California (Historical Images)	
Related Collections	Collection and Series Information	Name of Collection as hyperlink to keyword search in NP Gallery	Sequoia and Kings Canyon Roberts Collection (Roberts Collection); Sequoia and Kings Canyon Historical Image Series (Historical Images)	Yes
Create Date	Date Created	Date original (not digital) image was created- Roberts Collection dates are unknown, but date range is helpful	1888-1950 (Roberts Collection); 01/01/1949 (Historical Images)	Yes
Source Information	Source	Information about original type of resource digitized: composition and creation of image	Negative (Roberts Collection); Negative (Historic Images)	Yes
Contacts	N/A	Contact information for the curator of the collection	Person: Curator Organization: US National Park Service Position: SEKI Resource Management Address: Sequoia and Kings Canyon National Park, Three Rivers CA Email: seki/superintendent@nps.gov	No, Publisher information is enough (with option to see original resource in NP Gallery)
Intended Audience	N/A	Refers to a cataloging category that is unclear	Option 3	No, only relevant for internal park knowledge
Time Submitted	N/A	Date digital copy was added to repository	Wednesday, January 11, 2017 11:22:50 AM	No, only relevant for internal park knowledge
Last Edited	N/A	Date metadata of digital copy was last changed	Thursday, June 6, 2019 10:37:17 AM	No, only relevant for internal park knowledge
Original File Name	N/A	Digital image file name	00075.tiff	No, only relevant for internal park knowledge
Scanner Information	N/A	Information about scanner used to create image	EPSON EPSON, Perfection V700/V750	No, only relevant for internal park knowledge

GO TO INDEX

Embedded Timestamp	N/A	Date digital image was created? (all timestamp dates have the same date and format)	Monday, January 1, 0001 12:00:00 AM	No, only relevant for internal park knowledge
Resource Format	Format	Type of File	Tiff or Jpeg	Yes
File Size	N/A	Expressed in bytes	35.9 MB	No, only necessary for image download- which is accessed through NP Gallery
Metadata Date	N/A	Date original metadata was created?	2015-04-16T11:33:51-07:00	No, only relevant for internal park knowledge
Rating	N/A	? Based on a five star system	★★★☆☆	No, only relevant on NP Gallery
Categories	N/A	Describes the types of content in the image	Scenic, Historic	No, only relevant for NP Gallery
Related Portals	N/A	Links to Sequioa and Kings Canyon NP Gallery portal	Sequoia & Kings Canyon National Parks	No, only relevant for NP Gallery
Albums in Which This Asset Appears	N/A	Links to related albums	Negatives; Negatives, Falls (Roberts Collection); Buildings and Utilities; Buildings and Utilities, Tent Frame Seasonal Housing (Historic Images)	No, only relevant for NP Gallery

Figure 2

GO TO INDEX

Works Cited

Bruce, Thomas R. and Diane I. Hillman. “The Continuum of Metadata Quality: Defining, Expressing, Exploiting”. *Metadata in Practice.* American Library Association (Chicago, 2004), 238-256.

Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images. Edited by Murtha Baca, et al. American Library Association (Chicago, 2006).

“Contractual Restrictions”. *Rightsstatements.* Rightsstatements.org. https://rightsstatements.org/page/1.0/?language=en#collection-nc

“DPLA Geographic and Temporal Guidelines for MAP 3.1”. *Digital Public Library of America.*(December, 2017). http://bit.ly/dpla-geo-styleguide-3_1

GO TO INDEX

CLARIFYING PROVENANCE

Winter, 2021
IS 596 African Objects in Museums
Professor Ellen Pearlstein, Professor Andrew Apter, Professor Allen Roberts, Professor Glenn Wharton

Prompt: All over the world, museums are reflecting on how to make collections of African objects more relevant to their audiences. An interdisciplinary cohort of graduate students from Anthropology, Conservation, History, Information Studies, and World Arts and Cultures joined together in fall 2020 to each select an object from digital images provided from the Henry Wellcome Collection of African objects from the Fowler Museum. Students submitted an independent research proposal at the end of fall 2020. Students submitted a revised research proposal at the end of winter 2021.

ABSTRACT

My project was inspired by digitized museum collections and the ways they enable new methods for investigating those collections, providing access to them, and presenting information about them. I am also interested in how provenance research connects to methods to decolonize museums. Along with making their collections more accessible through digitization, publishing information about the provenance of these collections is an opportunity for museums to engage with their association with colonial systems and motives for collecting and displaying non-Western objects. I wanted to use the StoryMaps tool to present my research because it is one of the many digital platforms that allow museums and scholars to visualize and communicate complicated collections data. I also wanted to center my project around the "map" as a way to visualize this figure's provenance "data" in a way that is more immediate: seeing its various locations as points on a map better emphasizes the global network surrounding the trade and display of African objects.

A video tour of the online Story Map exhibit can be accessed at: https://vimeo.com/528978240

REVISED RESEARCH STATEMENT

My project was inspired by digitized museum collections and the ways they enable new methods for investigating those collections, providing access to them, and presenting information about them. I am also interested in how provenance research connects to methods to decolonize museums. Along with making their collections more accessible through digitization, publishing information about the provenance of these collections is an opportunity for museums to engage with their association with colonial systems and motives for collecting and displaying non-Western objects. I wanted to use the StoryMaps tool to present my research because it is one of the many digital platforms that allow museums and scholars to visualize and communicate complicated collections data. I also wanted to center my project around the "map" as a way to visualize this figure's

GO TO INDEX

provenance "data" in a way that is more immediate: seeing its various locations as points on a map better emphasizes the global network surrounding the trade and display of African objects.

In thinking about this object, I was first struck by the comparatively little information we do know with certainty verses the information we cannot know. We know that it was owned, in 1926, by Jeanne Walschot- we have the address of her shop and we have her own words describing the figure. We know that it was then owned by Sir Henry Wellcome, we also have his employee's description of the figure. We know that it was donated to the Fowler Museum somewhere between 1965-66 and an account of its role in the museum since then. This all centers around European actors. I do think that the way these actors have characterized, traded, and displayed this object (or similar ones) tells an important history about changing Western perceptions of African culture that parallels and entangles the histories of both Africa and the West. However, I did not want to imply that the importance of this or any other African object lies simply in its appearance in a European collection. I wanted to begin, first, by explaining its importance and identity outside of and before any European context, which is often left out of a record of provenance.

The problem I ran into here, though, was that the information we have about this context largely comes from European traders, missionaries, and colonial officers. Even the Kongo individuals who assisted Karl Lamman did so from a newly converted Christian perspective. This was further complicated by our discussions of secrecy and the idea that much religious knowledge in Africa is earned rather than given. I wanted to convey that this figure came out of a specific religious practice and set of visual conventions, and to point some of them out. I used Wyatt MacGaffey's scholarship and other essays in *Astonishment and Power* to point to some of the conventions that may explain the choices of the artist. I also wanted to be accurate about the limitations of taking an art historical approach in examining this object: making educated guesses about the intentions of the artist, already a complicated task, is made more complicated by the secondary nature of sources of information about Kongo art and the secrecy that acts as a barrier to knowing more information than we are meant to.

I also ran into some difficulty when deciding what to call the object. At first, I referred to it as "Nkisi", but after reading MacGaffey's perspectives about when an nkisi is active and when it is not, and after our discussions about life verses death histories, I decided that it was no longer accurate to refer to this object as nkisi since that is no longer the purpose it serves. I thought about calling it a sculpture, but was concerned that this insisted that the object be interpreted as art, when really the aesthetics of the object are only one part of its whole. Object feels a bit too inanimate when referring to a sculpture that was once active. After seeing the word "figure" in many digital collections (that is how the Fowler refers to this object now), I decided that this is perhaps the best and most neutral label to give it. I wanted to give some indication of this struggle through the "African object as" sections of my story map- to explain the various roles that make up the current identity of the object.

My exploration of provenance, a Western museum concept, inevitably became about Western museums. My research began with looking at the various actors who would have been responsible for taking objects out of Africa and back to Europe, especially in the 19th Century when much of this collecting took place. Here I relied on the essays of Linda Heywood, John Thornton, Hein Vanhee, and Jelmer Vos in *Kongo Across the Waters.* I also read the works of Raymond Corbey: *Tribal Art*

Traffic and "Ethnographic Showcases, 1870-1930". These introduced me to the historical contexts of trade with West Central Africa, missionaries collecting and documenting culture, and colonial officers taking objects due to national mandates and personal interest. I then focused on the relationship between Belgium and Congo, relying heavily on the work of Deborah Silverman who does an excellent job explaining the role that collecting played in nation-building in Belgium. I also read Maarten Couttenier's essay, "'One speaks softly, as if in a sacred place': collecting, studying and exhibiting Congolese artefacts as African art in Belgium," which discusses the early art/ artefact tension in the European presentation of African objects. Boris Wastau's article on Jeanne Walschot presented information about her motivations for collecting and how they fit into the broader Belgian relationship with Congo. I then, of course, turned to Frances Larson's *An Infinity of Things* to understand Wellcome's motivations for collecting non-Western artifacts. The book does a thorough job of positioning this, as well, within the broader framework of the developing field of anthropology and collecting as scholarship in the UK.

I was stuck on the idea of whether African objects are presented as art objects or ethnographic ones because I think it is at the heart of understanding how the West has justified and come to terms with its perceptions of Africa. There is plenty of literature available about the influence of African objects on European modernism and its display as art in the mid-twentieth century- I relied on essays published by the Met and Yaelle Brio's "African Art, New York, and the Avant-Garde" especially. I also found Christa Clark's chapter, "From Theory to Practice: Exhibiting African Art in the Twenty-First Century" in *Art and Its Publics*, which gave a thorough historical framework. I was grateful to the Brooklyn Museum and MoMA for publishing their exhibition catalogs so I could further investigate the *Primitive Negro Art, Chiefly from the Belgian Congo* and *African Negro Art* exhibits. It seems no coincidence that a growing insistence on the aesthetics of traditional African objects in the United States coincided with movements defending African American rights and celebrating African American culture. Not all American curators exhibiting African objects as art were concerned with how it affected African American rights. If anything, from what I have read, emphasizing its perceived immediacy and simplicity was related to concerns about Western excess (especially after WWI). However, it did represent a significant shift away from using these objects as evidence for Western superiority. As this shift was occurring, African objects as art were a source of inspiration and empowerment for a growing African American art movement. The essay on "Kongo in Contemporary Art" in *Kongo Across the* Waters, scholarship discussing Christa Clarke's *African Art in the Barnes Foundation,* and (fittingly) an online exhibit on the Harlem Renaissance from the Smithsonian African American Museum via Google (https://artsandculture.google.com/exhibit/african-american-art-harlem-renaissance-civil-rights-era-and-beyond-smithsonian-american-art-museum/fALyo3o0fCnXJg?hl=en) helped introduce me to the connection between African art exhibits and the Harlem Renaissance. Carlee Forbes (Mellon Curator at the Fowler) has also provided a useful bibliography for continuing to learn about the connection between African objects as art and the Harlem Renaissance!

In the 1960's, the Fowler's exhibition of Wellcome objects as art objects was in the context of a growing involvement with the Civil Rights movement at UCLA and the establishment of cultural studies departments. Marla C. Berns notes this connection in her book about the Fowler, *World arts, local lives: the collections of the Fowler Museum at UCLA.* In the 80's Susan B. Vogel talked

about the importance of presenting African objects at the Met as art on par with anything else being exhibited there. The implication seems to have been that for African culture to be respected in the West, it needed to be from an aesthetic perspective first. It was through acknowledging the artistic abilities of African cultures that the West could move away from insisting on their own superiority. Through describing some more recent museum exhibits and situating the mission of the Fowler within the broader reflexive trend in museums exhibiting African art, I hoped to draw some attention to this shift. I also read Vogel's reflections about mounting *ART/Artifact* and some reviews of the exhibit as well as Christa Clark's chapter and, of course, the *Exhibition-ism* publication to learn more about this shift in museums exhibiting African art.

I think the narrative that many museums in the West began collecting African objects to prove Western superiority is an important one to acknowledge out loud to museum audiences. At the same time, it continues to be important to acknowledge the agency and power of African culture apart from and independent to the West, which the Fowler and many other Western museums have done through exhibits and associated literature. As Mary Nooter Roberts points out in her essay in *Exhibitionism*, "Yet even if the ownership and display of foreign objects always ultimately reflects the exhibitors as much as or more than the producers, it is nevertheless important to consider how African peoples experience their relationships to objects and to the empirical world." (39). Following the example of exhibits like *Astonishment and Power* and *Kongo Across the Waters*, I wanted to book end the presentation with an acknowledgement of the ongoing evolution of African culture and art and the ways in which traditional objects continue to influence artists in Africa and the diaspora. The final reflections of the Story Map try to situate this project within the broader mission of the Fowler and provenance research:

> "Ongoing research at the Fowler and at other museums with African collections seeks to go beyond recording the context of the creation of an object, or highlighting its visual power. The challenge now is to research and present how these collections are a result of and represent networks of global cultural exchange. The provenance of objects like this one is much more than a list of locations on a map. Their existence in Western museums is a testament to centuries of trade, conflict, and diplomacy between Africa and the West. They are also a testament to the lasting power and influence of African visual cultural throughout the world."

I realize that my work and research, over the course of ten weeks, has only scratched the surface of the complexities surrounding African objects in museums, but my hope was to build a framework for a digital exhibit that better contextualizes collection provenance information. What I will continue to be interested in is how these complexities can be translated to the digital and what tools can best accomplish this. When visitors go to a museum, their interaction with collections is contextualized through an exhibit. When visitors see a collection online, it is sometimes just a series of images with minimal metadata. Digitizing collections provides exciting opportunities for access to these collections, but does a lack of accompanying metadata obscure these objects for non-scholarly audiences? Further, do museums have an obligation not only to provide contextual and provenance metadata, but to contextualize this provenance metadata? To be explicit about the history of museums collecting African and other non-Western objects? Considering that digital collections are the first entry point to museum collections for many, and the only entry point to "visitors" from other parts of the world, should more attention be paid to how they are presented? And how does

metadata and its presentation online avoid perpetuating a Western perspective about these objects and burying an African voice?

I do think it is important to record the history of the exhibition of African objects in Western museums and to make this history accessible to museum audiences. However, I recognize that the scale of this project is impossible to apply to every museum object. It might even be impossible for most museums to apply it to their general digital collections without having to significantly shift staff and budget models. I will be interested to see whether more museums do decide to dedicate resources to enriching metadata in their digital collections or to investing in databases that look more like the ones at the British Museum or the Met, with plenty of associated essays, timelines, and online exhibits. And to see what additional tools they decide to use.

I should state outright that I hope that more museums dedicate resources to contextualizing digital collections and improving access to them. I also hope to see more museums taking advantage of the uniquely flexible nature of digital collections in the context of reevaluating their role in the diaspora of African objects. Digital interfaces enable emphasis on the aggregate, to take a broad look at-for example- the dispersal of minkisi from Mayombe throughout various museum collections. They also allow opportunities to begin linking related collections and interpretations of collections. The Sierra Leone Heritage project (http://www.sierraleoneheritage.org/) is an example of linking Western and African collections together through object-centered databases. It also connects the Western interpretation at the British Museum with videos of the objects being used in real time in Sierra Leone. While this type of polyphonic presentation is difficult in the physical exhibit, the flexibility of the digital realm makes these grand ambitions more attainable.

Through future research, I hope to continue thinking about how museums and archives can collaborate with scholars and communities to build projects that contextualize and link collections. I hope that more communities can not only access collections, but research and metadata about the history of those collections. And that more collaborative work can be done to begin tracing and visualizing the complex networks of exchange centered around the diasporas of African, or other non-Western collections.

GO TO INDEX

Bibliography

Provenance: Creation

Hersak, Dunja. "Beyond the Naked Eye, Ethnography Screened through the Scientific Lens." *Critical Interventions*, 7:1 (2013).

MacGaffey, Wyatt:

"The Eyes of Understanding: Kongo Minkisi" *Astonishment and Power* (Washington, DC: Smithsonian Institution, 1993).

"Meaning and Aesthetics in Kongo Art" Kongo Across the Waters (Gainesville: University Press of Florida, 2014).

"Complexity, Astonishment, and Power: The Visual Vocabulary of Kongo Minkisi". *Journal of Southern African Studies,* 14: 2 (1988).

Provenance: Assumption of Ownership

Bay, Edna G. "The Scramble for Art in Central Africa." *American Anthropologist*, 101, no. 2 (1999).

Corbey, Raymond:

"Ethnographic Showcases, 1870-1930." *Cultural Anthropology*, 8, no. 3 (1993): 338-369.

Tribal art traffic: a chronicle of taste, trade and desire in colonial and post-colonial times. Amsterdam: Royal Tropical Institute, 2000.

Couttenier, Maarten. "'One Speaks Softly, Like in a Sacred Place'. Collecting, Studying and Exhibiting Congolese Artefacts as African Art in Belgium" *Journal of Art Historiography,* 12, no. 40 (July 1, 2015): 1-40.

Heywood, Linda and John Thornton. "The Kongo Kingdom and European Diplomacy." *Kongo Across the Waters*, ed. Susan Cooksey, Robin Poynor, and Hein Vanhee. Gainesville: University Press of Florida, 2014.

Luscombe, Stephen. "Africa and the British Empire" The British Empire, Accessed March 1, 2021, https://www.britishempire.co.uk/maproom/africa.htm

Silverman, Deborah:

"Art Nouveau, Art of Darkness: African Lineages of Belgian Modernism, Part III" *West 86th: A Journal of Decorative Arts, Design History, and Material Culture*, 20, no. 1 (2013)

“Diasporas of Art: History, the Tervuren Royal Museum for Central Africa and the Politics of Memory in Belgium, 1885-2014,” Journal of Modern History, 87, 3 (Sept. 2015): 615-667

Vanhee, Hein and Jelmer Vos. "Kongo in the Age of Empire" *Kongo Across the Waters*, ed. Susan Cooksey, Robin Poynor, and Hein Vanhee. Gainesville: University Press of Florida, 2014.

Wastiau, Boris:

GO TO INDEX

"The Legacy of Collecting: Colonial Collecting in the Belgian Congo and the Duty of Unveiling Provenance." *The Oxford Handbook of Public History* (2017) DOI: 10.1093/oxfordhb/9780199766024.013.25

"The Universe of Miss Jeanne Walschot, Belgian Collector and dealer of 'Congoliana', 1896-1977." *Collectors: individuals and instituions* (2001): 223 - 238.

Provenance: Private Sale

Larson, Frances. *An Infinity of Things, How Sir Henry Wellcome Collected the World*. New York: Oxford University Press, 2009.

Russell, Georgina. "The Wellcome Historical Medical Museum's Dispersal of Non-Medical Material, 1936 to 1983." *Newsletter (Museum Ethnographers Group)*, no. 20 (February, 1987).

Provenance: Donation

"African American Art: Harlem Renaissance, Civil Rights Era, and Beyond" Smithsonian American Art Museum. Accessed February 28, 2021, https://artsandculture.google.com/exhibit/african-american-art-harlem-renaissance-civil-rights-era-and-beyond/fALyo3o0fCnXJg

Altman, Ralph C. *Masterpieces from the Sir Henry Wellcome Collection at UCLA*. Los Angeles: UCLA, 1966.

Berns, Marla C. *World arts, local lives: the collections of the Fowler Museum at UCLA*. Los Angeles: Fowler Museum at UCLA, 2014.

Clark, Christa. "From Theory to Practice: Exhibiting African Art in the Twenty-First Century." *Art and its publics : museum studies at the millennium : New interventions in art history, ed. Andrew McClellan*. Malden: Blackwell Publishing, 2003.

Cooksey, Susan and Robin Poynor. "Kongo Inspiration in Contemporary World Art" *Kongo Across the Waters*. Gainesville: University of Florida Press, 2014.

Culin, Stewart. *Primitive Negro Art chiefly from the Belgian Congo*. New York: Brooklyn Museum, 1923.

Danto, Arthur C. *ART/artifact: African art in anthropology collections*. New York: Center for African Art, 1989.

Harris, Michael D. "Resonance, Transformation, and Rhyme: The Art of Renee Stout", *Astonishment and Power* (Washington DC: Smithsonian Institution Press, 1993).

Murrell, Denise. "African Influences in Modern Art." *Heilbrunn Timeline of Art History*. The Met, April, 2008. https://www.metmuseum.org/toah/hd/aima/hd_aima.html

Roberts, Mary Nooter et al. *Exhibition-ism: Museums and African Art*. New York: Museum for African Art, 1994.

Sweeney, James Johnson. *African Negro Art*. New York: Museum of Modern Art, 1935.

Vogel, Susan. "Bringing African Art to the Metropolitan Museum," *African Arts* 15, no. 2 (1982).

GO TO INDEX

CORE WORK

DROPBOX TECHNOLOGY REPORT

Winter 2021
IS 270 Systems and Infrastructures
Professor Miriam Posner

Prompt: Because of the rapid pace of evolution of information technologies, it is important to identify ways in which you can keep your skills fresh. This course will help you to identify, access and use resources for keeping up-to-date with the field of information technology, through writing a 4-page policy brief. The brief will be addressed to campus leaders, e.g., members of the UCLA Information Technology Planning Board, or University Librarian Virginia Steele, or Faculty Senate Chair Joseph Bristow, regarding technologies that might be adopted on campus or, more broadly, have an impact on the future of higher education, or on the university as a place where student, staff, and faculty spend considerable time.

ABSTRACT

I would like to use the Final Project as an opportunity to explore cloud based storage services. I believe that this is a technology that is relevant both personally and professionally. Most of the public understands the importance of storing digital information in a secure location, accessible by multiple people and across multiple devices, which is why many have turned to services like Dropbox for storing important records. Cloud based storage also answers the call for redundancy- ensuring the life of data is not determined by the durability of one hardware. In considering the longevity of data, accessibility and open source capabilities and the privacy and security of information, I want to explore what the trade offs are for storing data through a cloud system. Many businesses and individuals use these services for their convenience, assuming that their data is secure and that they will continue to have long-term access to it. I want to understand the extent to which this is true and examine whether there are better, more secure options available.

I am also including four shorter assignments written for the course investigating the infrastructure, architecture, standards, and issues of privacy and equity involved in the operation of Dropbox.

TECHNOLOGY REPORT

Joseph Bristow,

Cloud systems have become increasingly popular for use by business professionals and scholars. There is a growing call for open access to information and collaboration between institutions as well as flexible storage solutions and cloud providers have stepped in to answer this call. Multiple parties can view and edit content as well as comment on each other's work through the cloud. In addition to streamlining teamwork, cloud systems offer simple and secure storage solutions for growing amounts of born-digital content. By storing digital information in series of blocks, which are then copied and stored at various data centers and accessed through connected networks, these born-digital objects can be viewed by anyone with an internet connection. Beyond its obvious use in businesses that require multiple team members to work on one project, often remotely, libraries and

GO TO INDEX

archives are increasingly holding up cloud storage to answer the call for dynamic interaction between institutions and between users and repositories.

For professionals and individuals whose growing collection of digital materials poses a storage issue not only because of the amount of space it takes up on a hard drive, but because of the catastrophic loss that could occur if that hard drive failed, cloud storage is a more secure way of ensuring long term access to those materials.

Dropbox was one of the first businesses to take advantage of cloud technology. At first, Dropbox was simply a way to store information and allow access to this information by multiple users. Following the example of other cloud-based businesses, such as Google Drive, Dropbox now allows users to create and comment on documents on its server. It has recently shifted its brand to cater to independent, creative businesses. Dropbox continues to be one of the most widely used cloud solutions for individuals and businesses. Its interface is simple and easy to use and it offers competitive pricing for impressive storage capabilities.

In many ways, however, Dropbox is falling behind the trend in Cloud Computing. As more and more businesses invest in this technology, competitors are developing increasingly complex and customizable applications. An increasing awareness about security concerns inherent in third party hosting of sensitive data has caused businesses to consider private storage options. This isn't possible for users who don't have unlimited local IT support and a considerable technology budget for building stacked networks onsite. Thus, cloud computing providers are allowing businesses to build hybrid systems, maintaining some information in the private cloud and relying on third party services for a majority of their IT and network support. Dropbox competitors are allowing users to have their cake and eat it, too.

Dropbox is tempting as a sleek and easy to use storage and collaboration tool. Please see the attached briefing for an overview of its technologies, key issues surrounding cloud computing and the ways in which this technology is developing.

Sincerely,

Bailey Berry

GO TO INDEX

Dropbox: Cloud Computing Solution for Business

As more and more industries are relying on cloud computing to provide secure storage to their digital files and access to content across departments and institutions, Dropbox has marketed itself as a Infrastructure as a Service cloud computing solution for creative businesses.

- Dropbox offers Encrypted Block technology for the secure transfer and storage of files
- New features allow the creation of and collaboration on files on their platform
- Dropbox faces growing concerns of the security of its technology and its privacy standards for user data
- There is fierce competition from other cloud computing services, which offer more diverse applications and customized API's and computing capabilities
- With growing concern of privacy in public cloud providers, more businesses are turning to Private, hybrid and community cloud solutions

Background

When explaining why he created Dropbox, Drew Houston describes becoming frustrated with using a USB drive to transfer files between his work computer, home computer and laptops- which he frequently forgot to bring with him. At one point, one of his hard drives crashed and he lost the bulk of his content. He also struggled with attempting to share large amounts of data with colleagues through multiple email attachments[79]. Dropbox emerged to store and share large amounts of data using a system that was accessible across multiple devices and to multiple users.

Dropbox was launched in 2007 and had over 500 million users by the end of 2018.[80] It organizes its tiers of services offered into individual or team, each with different pricing plans. Differences in usage are divided on Dropbox's site into a few categories. "Dropbox core features" include cloud storage space, from 2 GB to "as much space as needed", access to a desktop app, access from any device able to connect to WIFI, and sharing capabilities secured with 256-bit and SSL/TLS encryption. "Content and accident protection" includes file recovery, remote device wipe, shared link controls and device approvals. "Productivity and sharing tools" refers to the tools Dropbox offers for sharing and working collaboratively on files stored in Dropbox. "Team management" is the tools offered for administrative control for larger companies such as billing, multi-team login, billing and audit logs with file event tracking. "Support" offers three levels of technical assistance: Priority email support, available to all users, Live chat support, available only to "professional" users and above and phone support, available to "standard" business users and above.[81]

[79] Ying, Jon "Meet the Team! (Part 1)"
[80] Trefis Team "Dropbox is Doing Well, But Looks Rich in the Face of Industry"
[81] "Choose the right Dropbox for you".

GO TO INDEX

Technology

Dropbox is Infrastructure as a Service, which means it offers users software, hardware, networks and IT support. Storing data on the cloud means that instead of data being stored locally in your own hard drive, it is sent to remote data centers through the internet. These data centers are made up of series of network servers clustered together in racks and connected by routers. These systems are hooked up to a backbone, which connects the whole system to the internet. When your computer is also connected to the internet, you can send your data in the form of packages to the backbone and then into the data center, where it is stored in several different places.

Dropbox has data centers across the United States in Washington, California, Texas, Illinois, New York and Virginia. It has expanded its data centers internationally to Germany, Australia and Japan where more secure storage is available for some Dropbox Business users.

Within a data center, Dropbox has developed what it calls a Quad-Plane, 3-Tier fabric. This system was developed to address issues with storage capabilities and scalability and address design issues with limitations of stacking rigid hardware. This new system unifies ASICs at every tier (Application-specific integrated circuit- the microchip which holds all of the potential logic gates, but created specifically to handle cloud computing/ data transfer operations), meaning that data can more easily communicate between different pieces of the stack. The fabric uses ECMP (Equal-cost multi-path routing), which allows for several "best" routes for data to take to get to one destination. Typically, routers have to compute which particular route data must take to get to the proper destination. By offering more than one route, data transfers should occur more quickly. Individual Racks are organized into series of sixteen, called a "Pod". Dropbox uses eBGP (external border gate protocol) to communicate between its clusters and iBGP (internal border gate protocol) to communicate transfer of information within its clusters. Connections within and between these building blocks in the fabric are made with Multimode fiber optic cables, the longest one being 140 feet.

Dropbox has something called a "Magic Pocket", which is the technology that manages the storage of data. Dropbox stores data using block technology. This means that files and the metadata about the files are split into separate blocks of bits or data, encrypted, duplicated and stored separately for more secure storage. This is all done ultimately using Solid State Drives. Incoming data is accepted by front end nodes, which determine where to store the data. The data is split into blocks, which are identified by a unique hash and assigned a checksum for ensuring the authenticity of data. Blocks are copied several times and duplicates are then stored in different "zones", which refer to regions of the United States where data centers are located. Blocks are grouped together by a Block Index into buckets with other blocks, which are then grouped into volumes. The Block Index stores these locations by associating the hash with a cell, bucket and checksum. These volumes are written onto physical discs where they are stored in different cells.

Key Issues and Future Trends

One of the most important issues with the use of Dropbox and the use of cloud technology in general is concern for privacy and security. While Dropbox is proud of its encryption services and block technology, which should ensure the secure transmission of information between networked systems, there are instances where data is unencrypted and thus vulnerable to security threat.

GO TO INDEX

Dropbox itself mentions in its privacy policy that there are some employees who may regularly access user data for legal purposes, these employees can potentially unencrypt and access your files at any time. There have been several data breaches at Dropbox, including in 2016 when 68 million Dropbox Account email addresses and passwords were published online.[82] Dropbox tests its infrastructure by using third-party services, such as "Hacker One", which invites internet users to find vulnerabilities and offers a bounty for any found and reported to Dropbox.[83]

Dropbox admits that it collects information regarding your account (name, email address, phone number, payment info, etc.); "Your Stuff"- the records and files you store in Dropbox along with "Related information" regarding your profile and "the size of the file, the time it was uploaded, collaborators, and usage activity"; Contacts; Usage Information- actions you take in your account; and Device Information (IP addresses, type of browser and device you use, webpage you visited before Dropbox and your location).[84] It also mentions that it uses Cookies and pixel tags and targets marketing towards its users and towards others associated with its users. Dropbox shares your information with "Trusted third parties": various incorporated businesses that offer support on Dropbox applications. They access your data "only..to perform tasks on Drobox's behalf". Added to this is concern about government interference. This is especially a concern in the United States where the Patriot Act allows the government to access any data stored in a provider's jurisdiction. According to Dropbox, you can request that they stop, limit use of or delete personal data- but only if "we have no lawful basis to keep using your data", defining "lawful basis" as: "to provide to Dropbox Services to you pursuant to our contract with you; in furtherance of its legitimate interests in operating our services and businesses".[85] Its recommendations for correcting or deleting your personal data are either changing your account information or deleting your Dropbox.

Beyond concerns about the security of files and collection of personal data, businesses who use Dropbox as a primary way of storing important information may also be concerned about legal human resources issues and compliance standards. Dropbox Business offers some assistance in this by allowing users to limit access to certain files, monitor changes to them and recover lost versions making Dropbox a digital and more sophisticated version of the locked filing cabinet. Dropbox also offers features such as "HIPAA Compliance" and Billing and audit logs. Still, there remains the possibility that Dropbox, a "trusted third party" or a government agent could access sensitive files and that this information could be leaked. Further, in a report for the Canadian Journal of Information and Library Science, Jessica Bushey et al. highlighted these concerns with maintaining reliability and authenticity of records in the cloud: : unauthorized access to information and records, privacy breaches, loss of access to and management of information and records, alteration of information in the cloud, lack of transparency regarding account management, server locations, data destruction and recovery.[86]

In response to these serious security concerns, many businesses are exchanging "public" cloud computing, such as a service like Dropbox, for "private" cloud computing, or a combination of both services, which is called "hybrid". Private cloud computing requires that companies provide

[82] Mendelsohn, Tom. "Dropbox hackers stole e-mail addresses, hashed passwords from 68M accounts"
[83] "Security"
[84] "Dropbox Privacy Policy".
[85] "The Dropbox Privacy Policy: Frequently asked questions."
[86] Bushey, Jessica et al. "Cloud Service Contracts: An Issue of Trust" p. 142

GO TO INDEX

their own networks for storage often located on site or close by. This option means that companies will have to take on a large portion of the costs for creating and maintaining stacked network technologies. Further, the growing quantity of digital information being produced by businesses often requires physically large and technologically sophisticated storage systems. To ease this cost, many businesses opt for "hybrid" cloud computing. For example, the University of Illinois uses Microsoft and Google suites for email and some team collaboration, but offers U of I Box – a private cloud – for sensitive files that require especially secure storage.[87]

U of I also uses Amazon (AWS) Suite for many of its cloud computing needs. Unlike Drobox, Amazon advertises hybrid cloud systems and offers a diverse set of cloud computing applications that allow businesses to pick and choose between simple Software as a Service models and customizable Platform as a Service applications.[88]Another of Dropbox's main competitors, "Box" offers a wide and diverse array of applications and API development capabilities meant to address the privacy needs of many different businesses, including medical and biological scientific labs, which increasingly rely on cloud systems for publishing and storing raw data.[89] IBM and Microsoft have also incorporated more flexible Platform as a Service models that allow businesses to decide how much of their network they would like to control and what they can rely on third party service providers to manage. Finally, libraries and smaller archives are also developing "community clouds", where several institutions share the maintenance and development of private data centers and networks, which not only distributes cost, but fosters greater collaboration between institutions in related fields.[90]

Cloud services that that enable companies to have some control over securing sensitive files, but relieve the burden of maintaining complex and dense hardware stacks, are taking over the market.[91] Many businesses are not only opting for these providers, but are choosing to combine applications from more than one provider. This has introduced a new service called "Cloud Brokers", which facilitate interoperability between different service providers. These brokers promise to stay informed about up-to-date cloud technologies so they can continue to advise clients about how to build the best hybrid cloud systems and ensure they operate smoothly as a complete system.[92] Some cloud computing services, such as AWS and IBM offer their own cloud brokerage services and promise interoperability between their applications and other third-party providers.

Compared to these increasingly complex competitors, Dropbox's services are fairly simple. While it is advertising itself as a business solution, it is not well-equipped to handle the sophisticated storage and software needs of many large companies. Dropbox offers a user-friendly interface and simple storage solutions at the cost of security and control over file access and network customizability.

---

[87] "Types of Cloud Computing: Private, Public and Hybrid Clouds"
[88] "Cloud Storage with AWS".
[89] "Box Platform, Extend the power of Box with APIs".
[90] Grant, Hurley. "Community Archives, Community Clouds: Enabling Digital Preservation for Small Archives."
[91] Drake, Nate and Brian Turner. "Best cloud computing services of 2020: for Digital Transformation."
[92] Yedlin, Debbie. "Pros and Cons of Using a Cloud Broker".

GO TO INDEX

Works Cited

"Box Platform, Extend the power of Box with APIs". *Box.com.* Box. https://www.box.com/platform. Accessed February 29, 2020.

Bushey, Jessica et al. "Cloud Service Contracts: An Issue of Trust". *Canadian Journal of Information and Library Science*. University of Toronto Press: June, 2015. Vol. 39:2, pp. 128-153.

"Choose the right Dropbox for you". *Dropbox*. https://www.dropbox.com/plans?trigger=nr. Accessed February 9, 2020.

"Cloud Storage with AWS". *Amazon.com.* Amazon. https://aws.amazon.com/products/storage/. Accessed February 29, 2020.

"Dropbox Privacy Policy". *Dropbox.com.* Dropbox: December 17, 2019. https://www.dropbox.com/terms#privacy, Accessed February 15, 2020.

Drake, Nate and Brian Turner. "Best cloud computing services of 2020: for Digital Transformation". *Techradar.pro.* December 20, 2019. https://www.techradar.com/best/best-cloud-computing-services. Accessed February 29, 2020.

Grant, Hurley. "Community Archives, Community Clouds: Enabling Digital Preservation for Small Archives." *Archivaria.* Vol. 81: 2016.

Mendelsohn, Tom. "Dropbox hackers stole e-mail addresses, hashed passwords from 68M accounts". *Ars Technica*. Condé Nast: August 31, 2016. https://arstechnica.com/information-technology/2016/08/dropbox-hackers-stole-email-addresses-hashed-passwords-68m-accounts/. Accessed February 15, 2020.

"Security". *Dropbox.* https://www.dropbox.com/security. Accessed February 15, 2020.

"The Dropbox Privacy Policy: Frequently asked questions." *Dropbox.com.* Dropbox. https://help.dropbox.com/accounts-billing/security/privacy-policy-faq, Accessed February 15, 2020.

Trefis Team "Dropbox is Doing Well, But Looks Rich in the Face of Industry". *Forbes.* May 21, 2018. https://www.forbes.com/sites/greatspeculations/2018/05/21/dropbox-is-doing-well-but-looks-rich-in-the-face-of-industry-headwinds/#7e8b2c5836ed

"Types of Cloud Computing: Private, Public and Hybrid Clouds". *Technology Services.* University of Illinois: 2020. https://cloud.illinois.edu/types-of-cloud-computing-private-public-and-hybrid-clouds/. Accessed February 29, 2020.

Yedlin, Debbie. "Pros and Cons of Using a Cloud Broker". *Technology and Business Integrators.* April 17, 2015. http://www.tbicentral.com/before-you-use-a-cloud-broker-pros-and-cons-of-using-one/. Accessed February 29, 2020.

Ying, Jon "Meet the Team! (Part 1)" *Work in Progress*. Dropbox. February, 5 2009. https://blog.dropbox.com/topics/company/meet-the-team-part-1. Accessed February 15, 2020.

GO TO INDEX

INFRASTRUCTURE

Dropbox was designed by Drew Houston and Arash Ferdowsi. When explaining why he created Dropbox, Houston describes becoming frustrated with having to use a USB drive to transfer files between his work computer, home computer and laptops- which he frequently forgot to bring with him. At one point, one of his hard drives crashed and he lost a huge bulk of his content. In addition to these problems, he struggled with attempting to share large amounts of data with colleagues through multiple email attachments[93]. Dropbox emerged to store and share large amounts of data using a system that was accessible across multiple devices and to multiple users.

Dropbox was launched in 2007 and had over 500 million users by the end of 2018.[94] It organizes its tiers of services offered into individual or team, each with different pricing plans. Differences in usage are divided on Dropbox's site into a few categories. "Dropbox core features" include cloud storage space, from 2 GB to "as much space as needed", access to a desktop app, access from any device able to connect to WIFI, and sharing capabilities secured with 256-bit and SSL/TLS encryption. "Content and accident protection" includes file recovery, remote device wipe, shared link controls and device approvals. "Productivity and sharing tools" refers to the tools Dropbox offers for sharing and working collaboratively on files stored in Dropbox. "Team management" is the tools offered for administrative control for larger companies such as billing, multi-team login, billing and audit logs with file event tracking. "Support" offers three levels of technical assistance: Priority email support, available to all users, Live chat support, available only to "professional" users and above and phone support, available to "standard" business users and above.[95]

Dropbox seems to be moving away from attracting individual users. It is competing with many other platforms offering to store and curate data and it has abandoned a couple of apps targeted towards individual use, like the photo app "Carousel". While Dropbox offers 2 GB of free storage to users with the option to upgrade to 2 TB of storage for $9.99/ month, icloud offers more tiers of storage for individuals starting at 5 GB of storage for free, 50 GB for $0.99/ month, 200 GB for $2.99/ month or 2 TB for $9.99/ month. Google Drive offers to store 15 GB of data for free, including photographs, documents and emails. These are only a few of the many other services which offer a more competitive platform for individuals storing personal digital files.

Dropbox now focuses on building its business capabilities. On their "About" page, they state "Our mission is to design a more enlightened way of working". Expanding on its original mission to allow file sharing, Dropbox is going up against Google Drive by including services that allow teams to collaborate on and build shared documents and projects. Dropbox recently unveiled a colorful redesign and the message "The world needs your creative energy"[96]. The primary audience for Dropbox is increasingly creative businesses.

---

[93] Ying, Jon "Meet the Team! (Part 1)" *Work in Progress*. Dropbox. February, 5 2009. https://blog.dropbox.com/topics/company/meet-the-team-part-1

[94] Trefis Team "Dropbox is Doing Well, But Looks Rich in the Face of Industry". *Forbes.* May 21, 2018. https://www.forbes.com/sites/greatspeculations/2018/05/21/dropbox-is-doing-well-but-looks-rich-in-the-face-of-industry-headwinds/#7e8b2c5836ed

[95] "Choose the right Dropbox for you". *Dropbox*. https://www.dropbox.com/plans?trigger=nr

[96] Richards, Katie. "Dropbox Wants to Unlock Creativity with an Unexpected Rebrand". *Adweek.* October 3, 2017. http://www.adweek.com/brand-marketing/dropbox-wants-to-unlock-creativity-with-an-unexpected-rebrand-that-features-a-trippy-video/

GO TO INDEX

Interestingly, as Dropbox is developing its business tools, it is also answering the call for the secure and responsible storage, organization and sharing of sensitive files. Businesses can limit access to certain files, monitor changes to them and recover lost versions making Dropbox a digital and more sophisticated version of the locked filing cabinet. It even offers a new premium feature that ensures that file storage policies comply with HIPAA.

One of the reasons that Dropbox is focusing on developing itself as a smart tool for businesses could be its desire to attract more customers willing to pay premium prices for their services. 2 GB of storage, offered for free, will hold about 200 jpegs, about 125 pdf files or about 300 docx files, which is not much storage in our digital age. In addition to more storage, paid plans offer security like viewer history and file wipe as well as recovery up to 180 days, larger file sharing capabilities and basic support, like the ability to speak to a representative over the phone. Dropbox has made its "basic" service essentially useless for most users. This presents a First Function Issue- if a business or individual would like to use Dropbox, they will most likely have to pay at least $9.99/ month. Dropbox's business plans are billed per user per month (with a minimum of three users), starting at $15/ user/ month (at least $45/ month) and charging $25/ user/ month for unlimited storage space and "sophisticated control and security features". If a business decides that it can absorb the cost of working with Dropbox for storing and collaborating on files, another First Order Issue presents itself in training employees on how to properly use the app. Dropbox's basic interface is simple- files are organized by folder and sub-folder allowing businesses to customize organization based on their individual needs. As new functions are added and businesses conduct more of their work through the Dropbox platform, organizations will not only be responsible for training their employees on how to organize files and where to find the files they need, but how to manipulate new software in order to complete their tasks.

In the face of so many other applications developed for allowing businesses to create and collaborate on files- from Google Drive to simply sharing files through email- a Second Function problem may present itself in convincing employees to restrict their communication and activity on files to Dropbox. If a business is going to commit to Dropbox as its primary platform, employees must be careful to make any changes to documents or files through Dropbox or to replace all shared files on Dropbox with the most recent version. A potential problem could exist with multiple different versions of the same file existing on several different employees' computers. Another problem could stem with employees' frustration at not having access to certain documents. Finally, employees may be frustrated at the prospect of learning Dropbox's tools in order to do the work they were able to accomplish through other services in the past. These problems will have to be addressed by managing the corporate culture and should hopefully be relieved by Dropbox's ease of use once employees become accustomed to the platform.

As businesses become more active in how their employees create, collaborate on, organize and archive digital content, it could be that Dropbox and all of the services it offers are already irrelevant. New platforms that more seamlessly and thoroughly encompass a business's needs at every stage of production could soon surpass Dropbox. Businesses may also determine that it is more efficient to instead incorporate a network of smaller platforms that perform each service more thoroughly. Before deciding to adopt Dropbox, businesses should determine what its needs are, from creation to archival storage of content, and investigate whether Dropbox truly meets all of these needs or whether there are other platforms available that do it better.

GO TO INDEX

ARCHITECTURE

Storing data on the cloud means basically that instead of data being stored locally in your own hard drive, it is sent to remote data centers through the internet. These data centers are made up of series of network servers clustered together in racks and connected by routers. These systems are hooked up to a backbone, which connects the whole system to the internet. When your computer is also connected to the internet, you can send your data in the form of packages to the backbone and then into the data center, where it is stored in several different places.

When your data is stored in Dropbox, it is packaged into data packages, which can be sent via servers to Dropbox's storage racks and retrieved from them when the data is needed again. This all works via a system of networks, connectors, routers and switches. All these devices connect together to tell data where it needs to go and provide pathways for it.

Dropbox has data centers across the United States in Washington, California, Texas, Illinois, New York and Virginia. It has expanded its data centers internationally to Germany, Australia and Japan where more secure storage is available for some Dropbox Business users.

Within a data center, Dropbox has developed what it calls a Quad-Plane, 3-Tier fabric. This system was developed to address issues with storage capabilities and scalability. Originally, Dropbox centers worked using a "four-post architecture" scheme featuring a cluster made up of CC (Cluster connectors) and CR (Cluster Routers). At each data center, there were several Clusters, which are tiered network systems. Networks are simply systems of communication for sharing data. The Cluster Routers handled the data packages traveling within racks in each cluster and data traveling out of the cluster. Data traveling between clusters at a data center is delivered by Cluster Connectors. The system used eBGP between Cluster Connector and Router and iBGP between a rack and cluster router, the different languages used to package and communicate what data is and where it should go.

GO TO INDEX

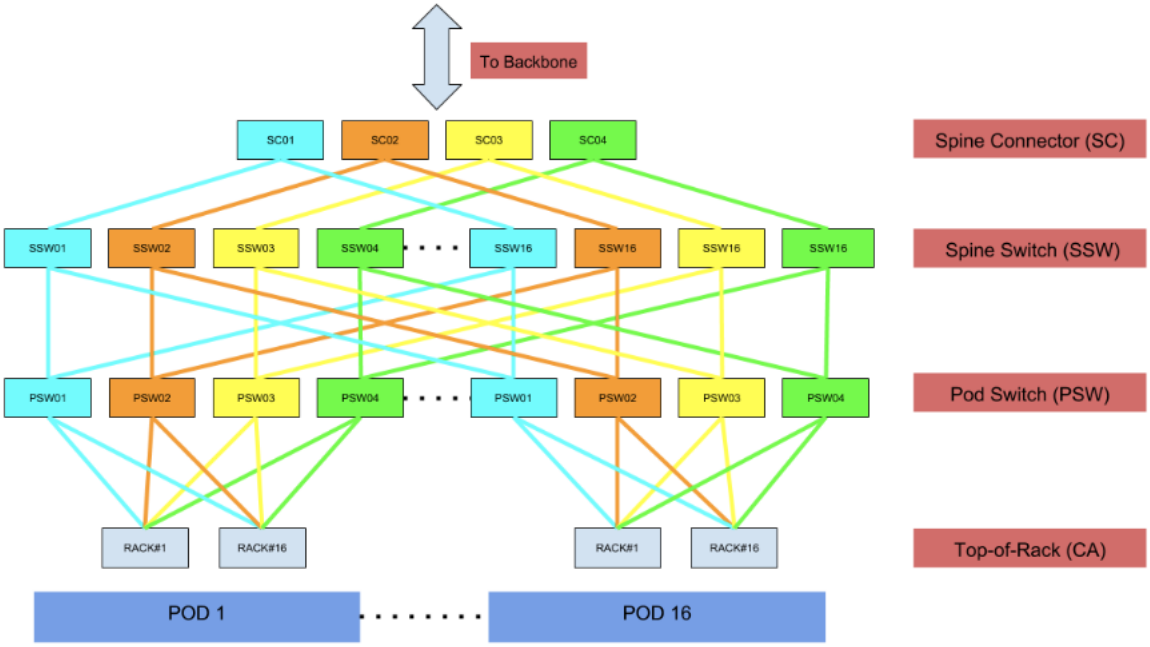


Figure 3: 256 rack fabric

[97]

Creating a "fabric" was meant to address design issues with limitations of stacking rigid hardware. Clusters were replaced with spine to reflect the more elegant horizontal connections between elements of the stack. Connecting hardware horizontally as opposed to vertically maximized the usage of ports on each device so that less ports went unused. This way, more racks could easily be gained to add capacity on demand. This new system unifies ASICs at every tier (Application-specific integrated circuit- the microchip which holds all of the potential logic gates, but created specifically to handle cloud computing/ data transfer operations), meaning that data can more easily communicate between different pieces of the stack. Uses ECMP (Equal-cost multi-path routing), which allows for several "best" routes for data to take to get to one destination. Typically, routers have to compute which particular route data must take to get to the proper destination. By offering more than one route, data transfers should occur more quickly. Individual Racks are organized into series of sixteen, called a "Pod". Instead of using both eBGP and iBGP, this system relies only on eBGP for ease of communication.

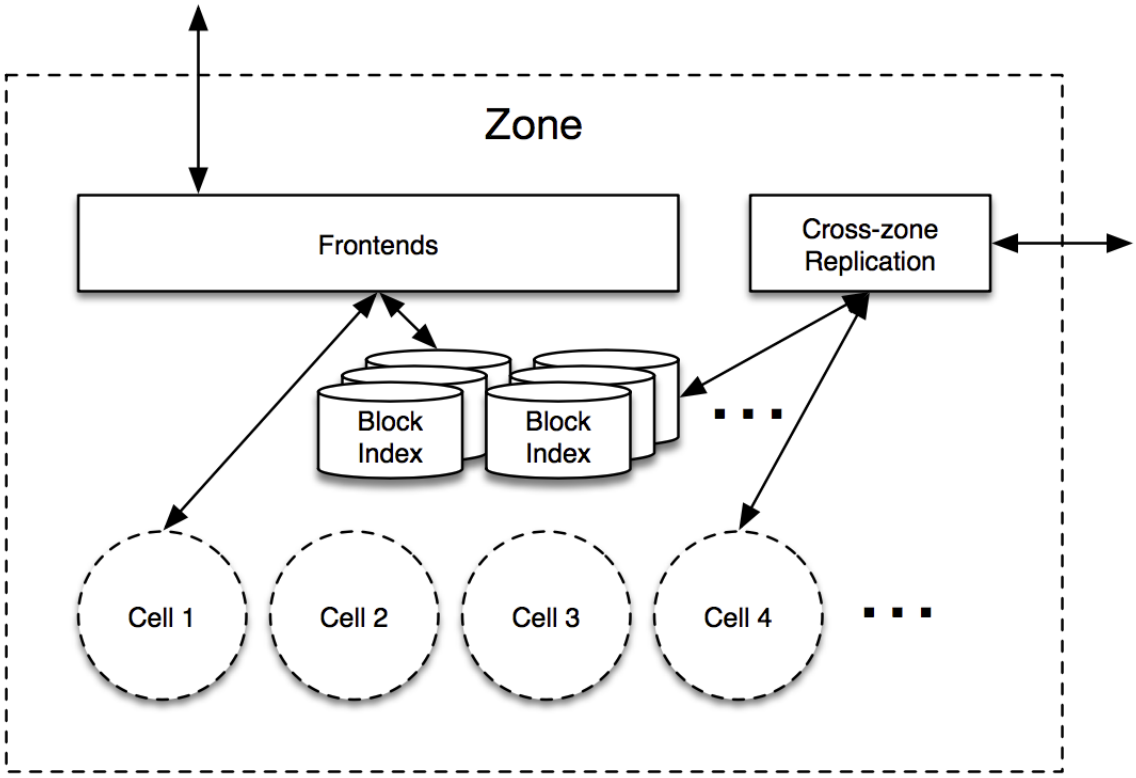
Connections within and between these building blocks in the fabric are made with Multimode fiber optic cables, the longest one being 140 feet.

Series of MDF (Main Distribution Frame), or cable rack, rows connect different pods to different power sources, meaning that in the event of any electrical anomaly, there are redundant electrical power distribution panels. Each MDF row "has diverse network pathways to each cabinet position".

[97] Sakpal, Vishal. "The scalable fabric behind our growing data center network". *Dropbox Blogs.* Dropbox: January 23, 2019. https://blogs.dropbox.com/tech/2019/01/the-scalable-fabric-behind-our-growing-data-center-network/ Accessed February 2, 2020.

GO TO INDEX

Dropbox has something called a “Magic Pocket”, which is the technology that manages the storage of data. Dropbox stores data using block technology. This means that files and the metadata about the files are split into separate blocks of bits or data, encrypted, duplicated and stored separately for more secure storage. This is all done ultimately using Solid State Drives. Incoming data is accepted by front end nodes, which determine where to store the data. The data is split into blocks, which are identified by a unique hash and assigned a checksum for ensuring the authenticity of data. Blocks are copied several times and duplicates are then stored in different “zones”, which refer to regions of the United States where data centers are located. Blocks are grouped together by a Block Index into buckets with other blocks, which are then grouped into volumes. The Block Index stores these locations by associating the hash with a cell, bucket and checksum. These volumes are written onto physical discs where they are stored in different cells.



[98]

Dropbox further categorizes data as “warm” storage or “cold” storage. It assumes that new data created will need to be accessed more often than older data. Dropbox’s blocks are “immutable”, so any changes to a piece of data will be stored in a separate system called “FileJournal”. Data that is “warm” is kept at a higher level in the stack so that it can easily be sent out through the network. Data that is “cold” is the data that is written onto solid state disc drives.[99]

---

[98] Cowling, James. “Inside the Magic Pocket”. *Dropbox Blog.* Dropbox: May 6, 2019. https://blogs.dropbox.com/tech/2016/05/inside-the-magic-pocket/. Accessed February 2, 2020.
[99] Le, Preslave. “How we optimized Magic Pocket for cold storage”. *Dropbox Blog.* Dropbox: May 6, 2019. https://blogs.dropbox.com/tech/2019/05/how-we-optimized-magic-pocket-for-cold-storage/. Accessed February 2, 2020.

GO TO INDEX

In order to use Dropbox, whether the website or the desktop app, your computer must be connected to the internet. This is the only way for your files and updates on your files to be sent to Drobox's data centers for storage. Dropbox does have options for offline access and even editing, but that information is only stored locally on your own computer until you are again connected to Wifi, meaning that it is vulnerable to any technological malfunctions that could destroy that information. Dropbox is very open about the ways in which it stores and transfers data. They seem to take pride in the methods they use to duplicate, spread and encrypt data to ensure secure storage and fast access.

STANDARDS

Dropbox is a cloud-based storage system, which means that it relies on standards primarily focused on transferring information through the internet. This means that every aspect of Dropbox's framework must focus on interconnectivity so that computing systems across multiple formats and hard wares can communicate with the Dropbox data storage system. This translates to standards associated with Transmission Control Protocol and Internet Protocol. These standards, set by the Internet Engineering Task Force, mandate how data is packaged and transferred between networks and users. The Internet Protocol, now at the sixth version, assigns unique identifiers to networks and systems on the internet in order to record their location. A Transmission Control Protocol ensures the orderly, error-checked and reliable transmission of data bytes from one system to another. Dropbox also uses HTTP, another part of the Internet Protocol suite. HTTP allows users to fetch data from Dropbox and have it delivered to their device through systems of hypertext. Part of the suite of standards involved in allowing data to be transferred between Dropbox and its users is the Border [100]Gateway Protocol language it uses. Dropbox runs primarily using a system of iBGP's which only communicate to other routers within Dropbox's system. These routers communicate with the backbone through eBGP's or exterior border gateway protocol languages, which communicate how data should be routed through different autonomous systems or stacks. BGPs make their route decisions based on existing paths, the nature of particular data packages and rules set by network administrators.

All of this data transfer must be standardized at the router architecture level. The Internet Protocol switched from v4 to v6 in anticipation of the greater volume of IP addresses, which meant that Dropbox had to adopt to a routing system that could accommodate both types of data. Dropbox decided to choose an IS-IS routing protocol, which is a protocol-agnostic architecture that focuses on transferring data between systems in the same network, which use the same point links. According to Dropbox, these characteristics make it easier to carry different types of information and support newer protocols, mitigating future TCP/IP updates.

From a user perspective, the most important function of Dropbox is its ability to render data and a variety of different files. Dropbox has developed its own image compression system, which it calls "Lepton". Lepton compresses JPEG files, which are already bit compressing systems, at a rate of 5 megabytes per second and decodes them back to the original bits at 15 megabytes per

[100] Oblumpally, Naveen. "Infrastructure update: evolution of the Dropbox backbone network". *Dropbox.* September 15, 2017. https://blogs.dropbox.com/tech/2017/09/infrastructure-update-evolution-of-the-dropbox-backbone-network/. Accessed February 8, 2019.

second.[101] Lepton also preserves the original file bit-for-bit. Dropbox supports most image files with a limit on the size of image files that can be "previewed" or seen without downloading them onto a computer. For all other files, Dropbox uses an application called zlib, which is also a lossless compression format. These tools are grouped together under the DiVANS suite, which converts raw bit streams into Intermediate Representation that it standardized to allow diversity of use and the safe conversion of the IR into a format that can be efficiently written out as bits.[102]

Thus, Dropbox operates using a complicated mix of internal and external standards that allow diverse networks to communicate with each other, route data and render files and bits. It is interesting to consider how much Dropbox's operations are affected by the data transfer standards set out by the TCP/IP suite. Developed by the Internet Engineering Task Force in 1980, these two protocols are meant to standardize the way that data is transferred between interconnected systems of packet-switched computer communication networks. The need for Transmission Control Protocol was primarily to address unreliability and availability in congestion in military, government, and civilian environments.[103] The Internet Protocol is concerned with how data is transferred by breaking it into smaller packages and the Transmission Control Protocol ensures that this information is transferred reliably- meaning that all packets arrive complete and in the correct order. Internet Protocol is also involved in assigning unique identifies to systems and networks on the internet to ensure routed data can find the correct location.

As indicated by the motivation behind the development of the TCP, the Internet Engineering Task Force was originally created under supervision of the United States government. It branched out in 1993 to become an independent, international activity associated with the internet society, a non-profit organization. Participation in the IETF is open for any individual with technical expertise to contribute. Participants join a working group, organized by topic into several Areas. Areas are then managed by Area Directors, who are members of the Internet Engineering Steering Group. This group is responsible for upholding rules and procedures ratified by the Internet Society trustees and oversees "entry into and movement along the Internet 'standards track'".[104] A separate group, the Internet Assigned Numbers Authority, oversees the assignment of unique parameter values for Internet Protocols.[105]

The Internet Engineering Task Force has taken pains to be an international and independent organization, which is important considering the amount of control it has. In switching from IPv4 to IPv6 Dropbox, as we have seen, was forced to completely redesign its internal routing architecture. Thus, it would be important that a body controlling how data is transferred internationally through

---

[101] Reiter Horn, Daniel. "Lepton image compression: saving 22% losslessly from images at 15MB/s". *Dropbox*. July 14, 2016. https://blogs.dropbox.com/tech/2016/07/lepton-image-compression-saving-22-losslessly-from-images-at-15mbs/. Accessed February 8, 2019.
[102] Reiter Horn, Daniel and Jongmin Baek. "Building better compression together with DivANS". *Dropbox.* June 19, 2018. https://blogs.dropbox.com/tech/2018/06/building-better-compression-together-with-divans/. Accessed February 9, 2019.
[103] "Transmission Control Protocol" Information Sciences Institute University of Southern California. Internet Engineering Task Force. September, 1981. https://tools.ietf.org/html/rfc793. Accessed February 8, 2019.
[104] "Internet Engineering Steering Group". *IETF.* https://ietf.org/about/groups/iesg/. Accessed February 8, 2019.
[105] "About". *IETF.* https://ietf.org/about/. Accessed February 8, 2019.

GO TO INDEX

the internet should have many systems of oversight, checks and balances to ensure that one developments in protocols do not favor one interest group, organization or even one nation.

PRIVACY AND EQUITY

There are several privacy issues that have been raised related to cloud storage and computing. These problems mostly arise from trusting a third party to securely store records and personal information, often at remote locations that are obscure to the user. Other privacy issues related to using any third-party internet service relate to the fact that they collect and store information on each user, often without notifying the user first. The ITU-T report on Privacy in Cloud Computing defines privacy as "the right to self-determination, that is, the right of individuals to 'know what is known about them', be aware of stored information about them, control how that information is communicated and prevent its abuse".[106] ITU-T goes on to indicate that in the Software as a Service model, which Dropbox falls into, the user has "little or no influence how input data is processed". It lists several relevant questions to consider when using a cloud service, as laid out in the Madrid Resolution:

• Who are the stakeholders involved in the operation?

• What are their roles and responsibilities?

• Where is the data kept?

• How is the data replicated?

• What are the relevant legal rules for data processing?

• How will the service provider meet the expected level of security and privacy?

In a report that investigated Cloud Service Contracts, Jessica Bushey et al. raised several issues to look for in keeping records in the cloud. Written for the Canadian Journal of Information and Library Science, the report was more concerned with how organizations could be sure that their records remain authentic and reliable while stored in the cloud. They identify these risks: unauthorized access to information and records, privacy breaches, loss of access to and management of information and records, alteration of information in the cloud, lack of transparency regarding account management, server locations, data destruction and recovery.[107] They further bring up the point that while records may clearly belong to users of the cloud, metadata produced by the user may technically belong to the cloud service provider.

Dropbox does attempt to be transparent, especially in relation to its technology, by maintaining a blog, which details how data is stored and retrieved at every level of operation. They further attempt to address these concerns in their Privacy Policy. Dropbox admits that it collects information regarding your account (name, email address, phone number, payment info, etc.); "Your Stuff"- the

[106] Guilloteau, Stephane and Venkatesen Mauree. "Privacy in Cloud Computing". *ITU-T Technology Watch Report*. ITU: March, 2012. https://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000160001PDFE.pdf, Accessed February 15, 2020.

[107] Bushey, Jessica et al. "Cloud Service Contracts: An Issue of Trust". *Canadian Journal of Information and Library Science*. University of Toronto Press: June, 2015. Vol. 39:2, pp. 128-153.

GO TO INDEX

records and files you store in Dropbox along with "Related information" regarding your profile and "the size of the file, the time it was uploaded, collaborators, and usage activity"; Contacts; Usage Information- actions you take in your account; and Device Information (IP addresses, type of browser and device you use, webpage you visited before Dropbox and your location).[108] Dropbox also mentions that it uses Cookies and pixel tags and targets marketing towards its users and towards others associated with its users. The reasons it claims to collect this information mostly relate to "provide, improve and promote our Services". Dropbox also claims to use personal data for "legitimate business needs". These needs mostly relate to tracking how you use services to target new ones for you, to understand how their services are functioning and to improve them and to track any unusual activity for security reasons. Dropbox claims that it will ask for user consent before processing personal data for other purposes.[109]

In its Privacy Policy, Dropbox also mentions that it shares your information with "Trusted third parties": Dropbox, Inc., Amazon Web Services, Inc. (Infrastructure), Teleperformance A.E. (an international customer experience management company), Salesforce.com (a partnered business management company), Serenova, LLC (cloud contact center), Google LLC (Customer Support), Zendesk, Inc. (customer support tool), Oracle America, Inc (Billing and Customer Support). Dropbox assures that these companies will only access your data "only..to perform tasks on Drobox's behalf" (FAQ). In terms of identifying who the stakeholders in all of this are, the picture becomes more complicated with the wide range of different companies associated with Dropbox who can each have access to your personal information. Added to this is the concern that various governments can technically access both your personal information and any files you've stored through a cloud service. This is especially a concern in the United States where the Patriot Act allows the government to access any data stored in a provider's jurisdiction.

According to Dropbox, you can request that they stop, limit use of or delete personal data- but only if "we have no lawful basis to keep using your data".[110] However, Dropbox defines its legal basis for using your data as: "to provide to Dropbox Services to you pursuant to our contract with you; in furtherance of its legitimate interests in operating our services and businesses" it also mentions that this should all happen "with your consent".[111] Its recommendations for correcting or deleting your personal data are either changing your account information or deleting your Dropbox. It also provides the email address privacy@dropbox.com, where you can complain about use of personal data or request the deletion of it. This same email address is provided for any requests for Data Access. In "your account", a user can opt out of notifications from Dropbox, but otherwise there is no simple way to block access to personal information.

If you do decide to delete your account on Dropbox, it claims that it will delete all of your data within 30 days. This is referring to the files that you store on Dropbox, but does not address all of the personal data and information Dropbox and its affiliates have collected about you. Further, Dropbox does not make it clear in its Terms of Service whether metadata created on its site belongs to its

---

[108] "Dropbox Privacy Policy". *Dropbox.com.* Dropbox: December 17, 2019. https://www.dropbox.com/terms#privacy, Accessed February 15, 2020.
[109] ibid
[110] "The Dropbox Privacy Policy: Frequently asked questions." *Dropbox.com.* Dropbox. https://help.dropbox.com/accounts-billing/security/privacy-policy-faq, Accessed February 15, 2020.
[111] ibid

GO TO INDEX

users or to Dropbox, something that Bushey et al. define as essential in determining the privacy standards of cloud storage. Dropbox does release a report for the amount of times the government has accessed user data including information. It includes what types of requests they've been, what states the requests have been made in, the amount of requests. Dropbox claims that it lets its users know when requests have been made for their information, but their own report shows that they've only been legally allowed to do so in about 40% of the cases.[112]

There have been several data breaches at Dropbox, including in 2016 when 68 million Dropbox Account email addresses and passwords were published online.[113] Dropbox offers encryption services incorporated into its blockchain technology to protect files during transfer. It also tests its infrastructure by using third-party services, such as "Hacker One", which invites internet users to find vulnerabilities and offers a bounty for any found and reported to Dropbox.[114]

There are many concerns related to privacy and security with Dropbox. While Dropbox is fairly transparent about the data that it collects, how it uses that data and who it shares it with, Dropbox gives no indication of how or where that data is stored. Users are fairly powerless to prevent Dropbox and its affiliates from collecting that data unless they simply decide to delete their Dropbox accounts altogether. Once those accounts are deleted, Dropbox does not clarify whether the data it has collected is also deleted or can continue to be used by Dropbox et al. Finally, the level of data security and protection vary slightly depending on how much users are willing to pay for their plans. For example, certain Dropbox Business owners are able to store data in Europe, making it eligible for the EU-U.S. and Swiss-U.S. Privacy Shields Framework laws. This means data is protected by European law, even if Dropbox is a country associated with the United States. These are all very real and relevant concerns for anyone considering using Dropbox as a cloud storage service.

[112] "Transparency at Dropbox". *Dropbox.* June, 2019. https://www.dropbox.com/transparency/reports, Accessed February 15, 2020.
[113] Mendelsohn, Tom. "Dropbox hackers stole e-mail addresses, hashed passwords from 68M accounts". *Ars Technica*. Condé Nast: August 31, 2016.
[114] "Security". *Dropbox.* Accessed February 15, 2020.

GO TO INDEX

COURSES TAKEN

FALL 2019

IS 211: Artifacts and Cultures/ Professor Johanna Drucker
IS 260: Description and Access/ Professor Gregory Leazer
IS 431: Archives, Records and Memories/ Professor Anne Gilliland

WINTER 2020

IS 241: Digital Preservation/ Professor Anne Gilliland
IS 270: Systems and Infrastructures/ Professor Miriam Posner
IS M238: Environmental Protection of Collections for Museums, Libraries, and Archives/ Professor Ellen Pearlstein

SPRING 2020

IS 212: Values and Communities/ Professor Ramesh Srinivasan
IS 262B: Data Curation and Policy/ Professor Jillian Wallis
IS 464: Metadata/ Professor Melissa Gil

FALL 2020

IS 289: Museums in the Digital Age/ Professor Miriam Posner
IS 298A: Doctoral Seminar: Research Methods and Design/ Professor Gregory Leazer
IS 596: Directed Idividual Study: Ethnomusicology Research Data/ Professor Anne Gilliland

WINTER 2021

IS 596: Directed Individual Study: African Objects in Museums/ Professors Ellen Pearlstein, Andrew Apter, Allen Roberts, Glenn Wharton
IS 422: College, University, and Research Libraries/ Professor Robert Montoya
IS 438B: Archival Description and Access/ Professor Jonathan Furner

SPRING 2021

IS 289: Digital Asset Management/ Professor Linda Tadic
IS 438A: Archival Appraisal/ Professor Anne Gilliland
IS 433: Community-Based Archives/ Professor Michelle Caswell

GO TO INDEX

ADVISING HISTORY

ANNE GILLILAND

Dr. Gilliland was my primary faculty advisor during the MLIS program. Apart from providing guidance about a course schedule that would best fit my professional goals, Dr. Gilliland provided invaluable guidance about various opportunities outside of the program. She helped me develop a focus for my summer internship at Sequoia and Kings Canyon, served as the faculty advisor for my directed independent research project in fall 2020, has helped me develop the paper for publication, and will advise me as I participate in an archival appraisal project spring, 2021.

Below is a schedule of advising sessions, in person and via email, with Anne Gilliland.

FALL 2018: Information about the UCLA MLIS program and guidance on writing a Statement of Purpose

FALL 2019: November in-person meeting about winter quarter class schedule and email correspondence about the SEKI summer internship application

WINTER 2020: Email correspondence about funding options for the MLIS program and balancing work and a course schedule

SPRING 2020: March in-person meeting about focusing my summer internship, spring courses, and continuing humanities data management research

SUMMER 2020: Email correspondence about fall courses and organizing a directed independent study about ethnomusicology research data management

FALL 2020: Email correspondence as faculty advisor for ethnomusicology research data management directed individual study, assistance with IRB process

WINTER 2021: March video meeting about spring courses, internship opportunities, and an independent archival appraisal opportunity

PROFESSIONAL MENTORSHIP

I have had several meetings with professionals from the GLAM field who have offered guidance about new directions in the field, MLIS career paths, and professional development opportunities

ASHLEY E SANDS Senior Program Officer, Institute of Museum and Library Service

Dr. Sands was matched with me as my LISAA mentor. We met via video conference on February 16, 2021 to discuss trends in Research Data Management, scholarly publishing, and professional development resources. I have continued to seek advice from Dr. Sands via email.

DEIDRE WHITMORE Digital Archeology Lab and Data Publication Director, UCLA Cotsen Institute of Archeology

GO TO INDEX

Dr. Sands introduced me to Ms. Whitmore as a resource for more information about researcher data management in the social sciences and humanities and data publishing. We met via video conference on March 15, 2021 to discuss these topics and future trends in the field.

CARLEE FORBES Andrew W. Mellon Curatorial Fellow, Fowler Museum

Dr. Forbes assisted with the African Objects in Museums course during winter quarter, 2021. She met with me individually via video conference a few times over the course of the quarter to help me develop my digital exhibit and to discuss issues around museum descriptions, digital databases, and metadata.

SUSAN ANDERSON History Curator, California African American Museum

Anderson will be advising me as I appraise a private family collection that relates to the history of African American leadership in Los Angeles and the Civil Rights Movement.

I would like to acknowledge the help and guidance I received from all of my professors in the UCLA MLIS program. Their thoughtful comments helped me make the most out of course assignments and extracurricular opportunities. Their support and understanding eased the transition to remote learning

GO TO INDEX

CV

EDUCATION

MLIS, UCLA Graduate School of Education and Information Studies
Expected June, 2021
Advisor: Anne Gilliland

BA, Gordon College
June, 2014
Art History major, Studio Art minor
Summa Cum Laude, Kenneth L. Pike Scholar, A. J. Gordon Scholar

WORK EXPERIENCE

Independent Archivist, Shaw Family Collection
Spring quarter, 2021
Inventorying and appraising the contents of the Shaw Family Collection

Reader, UCLA Institute for Society and Genetics
Spring quarter, 2021; fall quarter, 2020
Determined grades and provided feedback for assignments completed in Soc Gen 180, "Data In and Of a Pandemic"

Teaching Assistant, Reader, UCLA Scandinavian Section
Reader: Fall quarter, 2020; Teaching Assistant: Academic Year 2019-20
Taught discussion courses, determined paper and participation grades, and conducted office hours for students enrolled in literature analysis and writing courses

Remote Archival Intern, Sequoia and Kings Canyon National Parks
Summer, 2020
Cleaned data and enhanced metadata to prepare digitized collections for harvesting by Calisphere and the California Consortium of Herbaria

Independent Archivist, Franklin K. Lane Estate
February, 2018 - October, 2020
Inventoried, re-housed, and created a Finding Aid for records belonging to Franklin K. Lane. Advised on steps for preparing the collection for possible donation

Office Manager, Gwen Restaurant
Office Manager: June, 2018 - August, 2019; Office Assistant: September, 2017 - June, 2018
Managed supplies and records for the office and restaurant; assisted with human resources including onboarding, terminations, and health insurance policies

GO TO INDEX

Lead Intern, Royale Projects Contemporary Art Gallery
Lead Intern: November, 2016 - May, 2017; Intern: May, 2016 - February, 2016
Kept an accurate inventory of artworks in the collection and prepared them for storage and installation. Oversaw a team of interns carrying out daily gallery tasks

Principle Cataloger, Smithsonian Environmental Archeology Lab
June, 2014 - February, 2016
Cleaned, researched, and cataloged various artifacts uncovered at archeological sites on the SERC campus

Visitor Services Representative, National Building Museum
June, 2014 - February, 2016
Greeted and assisted museum visitors, facilitated engagement with exhibits, and documented visitor feedback

Gallery Intern, Gallery at Barrington Center for the Arts
January - May, 2014
Assisted in curating, promoting, and interpreting upcoming exhibits at the art gallery

Archival Assistant, Patton Archives Project
September - December, 2011
Researched and archived various artifacts from the General George S. Patton Estate

ADDITIONAL PROJECTS

Archival Advisor, Los Angeles Contemporary Archives
February - June, 2021
Assisted in developing archival and accession policies, developed accession and finding aid forms

African Objects in Museums
Winter quarter, 2021
Interdisciplinary Course researching objects in the Fowler Museum's Wellcome Collection and discussing methods of decolonizing museum practice; My project uses GIS Story Mapping to present provenance data about a Mayombe Nkisi figure

Ethnomusicology Research Data Management
Fall quarter, 2020
Conducted an independent research project about how researchers in ethnomusicology manage, reuse, and share their data; gathered data through semi-structured interviews with faculty, postdocs, and archival staff

Veterans Administration Archives
Fall quarter, 2019
Responsible for drafting a Mission Statement; Defined policies for acquisition, appraisal, organization, and use of archival assets

GO TO INDEX

Analysis of a Delaware Shell Button Factory
2015 - 2016
Gathered data about The Parizek Brothers pearl shell button cutting station through archeological fieldwork and oral history interviews. Presented a corresponding paper at the Mid-Atlantic Archeology Conference, March 2015 and the Society for Historical Archeology Conference, January 2016

Momentum Gallery
Summer, 2014
Developed a pop-up gallery in the North Shore of Massachusetts, which exhibited the work of graduating BFA students from three different local colleges

The Animal Eye Camera
Spring, 2014
Partnered with Dr. Brian Glenney in developing a photography exhibit based around his Animal Eye Camera, which engaged visitors in the philosophy of vision and perspective

The Accessible Icon Project
Spring, 2014
Assisted in developing and installing a pop-up exhibit featuring artists living with disabilities in conversation with Dr. Brian Glenney's Accessible Icon project being exhibited ad MoMA

Les Jeunes Ont La Parole, Musee du Louvre
Fall, 2013
Carefully researched Fra Angelico's *Coronation of the Virgin* to interpret the piece for museum visitors

PROFESSIONAL ORGANIZATIONS AND WORKSHOPS

LEADERSHIP

- Academic Year, 2020-21- Co-Records Manager, Society of American Archivists Student Chapter

MEMBERSHIP

- Society of American Archivists
- Society of California Archivists
- Research Data Access and Preservation Association

LIBRARY CARPENTRY

- Intro to Tidy Data in a Spreadsheet
- OpenRefine
- Unix Shell
- Intro to Git/GitHub

GO TO INDEX

- SQL
- Intro to Python
- Working with APIs in Python

PUBLICATIONS

"Parizek Button Station: Analysis of a Button Cutting Factory in Milford, Delaware". *Maryland Archeology* 50, no. 1 (May 2014):4-8.

ADDITIONAL SKILLS

- Familiarity with XML and Oxygen
- Proficient with managing data in Microsoft Excel and OpenRefine
- Experience with Archives Space and SQL
- Experience with GIS and AutoCad mapping software

GO TO INDEX

ACCESSIBILITY STATEMENT

The online portfolio website was created using the https://www.w3.org/WAI/standards-guidelines/wcag/ Web Content Accessibility Guidelines. It was created to enable easy navigation and with an eye to visual simplicity. All links are clearly labeled and website sections are marked with ARIA Landmarks. Where possible, the downloadable PDF version is tagged to make it readable by accessible devises. Easier navigation is enabled using an index and bookmarks throughout the document. The website has been checked using the https://wave.webaim.org/ Web Accessibility Evaluation Tool.

The HTML portfolio website was made using GitHub, an open source tool that is free to use. These factors allow for easier preservation and long-term access.

GO TO INDEX