San Francisco Earthquake and Fire Collection

Group Project Proposal

INFO 284

Heather Johnson

Kari Lease

Joshua Simpson

I. Conceptualize

Goals of Collection

The Great San Francisco Earthquake and Fire occurred in the early morning hours of April 18, 1906. It is considered one of the most significant earthquakes of all time and provided a wealth of knowledge to the field of seismology. Analysis of the crust and fault line led to the formulation of the "elastic-rebound theory of the earthquake source, which remains today the principal model of the earthquake cycle" ("The Great 1906," n.d.). More than 3,000 people died, over half the population of 400,000 people became homeless, 28,000 buildings were destroyed, and there is estimated to be upward of $400 million worth of property damage, which would equal $11.4 billion in 2019. ("Casualties and damage," n.d.).

While these statistics put the scope of the structural destruction into perspective, they don’t capture the staggering human toll. The devastating impact of this event on the people of San Francisco are revealed in harrowing detail in our collection of photographs and correspondence. The destruction of a bustling and vibrant city can be seen and felt through these powerful images and words. Visitors to our digital repository will attest to how this event became a turning point in the development of "The City" and how this event marks a rebirth to what it is today.

Our Institution

This institution is a new digital archive that is focused solely on destructive seismic activity in the San Francisco Bay Area. The archive is a digital humanities project that will be jointly funded by the San Francisco Public Library (SFPL) and San Francisco Parks and Recreation. The office space and any physical versions of the objects will be housed in the main branch of the SFPL and the digital archive will be held on the SFPL servers; the Parks and Rec department will have a page on their website that informs the public of the project with a brief description and links them to the actual archive website and app. While the dedicated employees of the archive will be solely focused on the digital archive, departments from both funding institutions will be free to use the archive materials in displays, exhibitions, etc. throughout the parks and library branches.

Our designated community are patrons of the San Francisco Public Library, researchers of San Francisco history, California school districts, and the general public interested in the study of natural disasters and their repercussions on society.

Our collection

The collection consists of primary sources from the days and months following the 1906 San Francisco earthquake and resulting fire. Objects collected will focus on both the immediate aftermath of the disaster and the subsequent rebuilding efforts of the city. Secondary sources will be excluded as the archive wants to focus on the experiences and reflections of those who witnessed the disaster first hand or arrived shortly thereafter. The objects consist mainly of letters, postcards, and photographs. Newspaper clippings may be included on a case by case basis if it is determined that they fit the standards of the project.

How the Collection Fits Within the Scope of the Institution

The digital archive will focus on the human impact of catastrophic seismic events in the San Francisco Bay Area, along with their after effects like the 1906 fire. The archive will begin with the earthquake of 1906, with plans to include at a later date, collections from other local earthquakes from the 20th century.

Part of the missions of both the SFPL and SF Parks and Rec is preserving and disseminating SF culture and history to Bay Area residents and visitors, and they believe this archive directly relates to that mission. Earthquakes and earthquake lore are a large part of Californian life. Californians often talk about “The Big One”, which references the next large earthquake that would leave devastation akin to the 1906 earthquake. Children perform earthquake drills in school and learn about past devastating earthquakes in history classes, and most everyone either has an earthquake kit or frequently talk about how they need one. The funding institutions believe that a space dedicated to this part of California life would be beneficial to current and future generations.

II. Create and Receive

All objects have been acquired with the consent of the owner of the records and in consultation with the archivist in charge and governing body of that repository. It has been verified that the transferor has proper authority or title to transfer the objects.

Metadata

The objects received included attached existing descriptive and administrative metadata in XML format. The amount of metadata to compile and write will necessitate a dedicated metadata librarian. Due to the size of the collection, this is a six month temporary position and will utilize staff already employed with the San Francisco Public Library.

METS

METS (Metadata Encoding and Transmission Standard) is a standard for encoding XML that is maintained by the Library of Congress. It encourages interoperability by providing a language for exchanging digital materials.

PREMIS

PREMIS Preservation Metadata: Implementation Strategies. The information collected is representative of what is needed for long-term preservation. There is wide applicability and technical neutrality. It is often used in conjunction with METS and is supported by the Library of Congress.

How the Metadata Will Be Created and/or Enhanced

We will be using METS and PREMIS for our metadata. The decision to use METS was made to provide metadata for the display of the collection and the exchange of information between repositories. PREMIS will be used to create the preservation metadata while METS will be used for the descriptive and administrative (technical rights, provenance, and source)

metadata (Guenther and Wolf, 2009). There is metadata attached to some of our objects from other repositories, all XML-based. METS will be used as the container for that metadata since it allows for other metadata packages to be entered as a complete package, in one place. This will save time by not needing to crosswalk the metadata from different schemas. Every attempt will be made to avoid redundancy. In the case of duplicate administrative information in both PREMIS and METS, we will choose the proper and equivalent METS sections. Since there are possible redundancies that can occur between PREMIS and METS, and since the display and the preservation of the information are of equal importance, we will err on the side of caution and leave those redundancies in place if necessary to ensure the proper description of the objects (PREMIS, 2017).

The METS documents will include: Descriptive <dmdSec>, Administrative <amdSec> (covers provenance and rights), File Groups <fileGrp> (lists any and all files that comprise the digital object), Structural Map <structMap>, Structural Links <structLink>, Behavior <behaviorSec>, and Header <metsHdr> (Amaral, 2010). The Metadata Extraction Tool will be used to automatically extract data from the various file formats we have to work with. Our controlled vocabulary will come from the Thesaurus for Graphic Materials found on the Library of Congress website.

A data dictionary will be compiled identifying the metadata fields that will be required information from each object. This data dictionary will ensure that our metadata is consistent. Information from the objects will include: ID#, title, author/photographer, date, location, subject, historical notes, digital collection, repository, repository collection, repository collection guide, object type, and physical description.

Metadata Storage

Metadata created by the metadata specialist will be stored in a PostgreSQL relational database and will also be stored as an XML file stored with the content data files. The database allows for quick access and updating as well as easy querying and reporting (PREMIS, 2015, p. 25).

The collection currently includes the following firsthand accounts (letters, postcards, newspaper clippings, etc) and photos from directly after the San Francisco 1906 earthquake and fire.

Data	Data Type	Format/ Size	Planned Access
Story of an Eyewitness	Manuscript	docx, 3pgs, 20 kb.	Open access via institution repository
Thousands Flee From Blazing City	Manuscript	.html, 5pgs, 200 kb.	
images_from_1906_earthquake	Digital images	.jpg, 15 mb	
Ruins of "Angel of Grief" statue, Stanford University, Palo Alto.	Digital image	.tif	
New Homestead Claim	Digital Image	.jpg, 1674 kb.	
collection_of_letters_1906_earthquake_aftermath	Manuscript	.pdf, 40 pgs, 3 mb.	
Collection of Newspaper Clippings	Digital Image	.pdf, 7 mb	
City Hall	Digital Image	.tif, 9 mb.	

III. Appraise and Select

Our designated community (DC) has a lot of bearing on the decisions we make about appraisal and selection. Our DC has been defined as patrons of the San Francisco Public Library, researchers of San Francisco history, California school districts, and the general public interested in the study of natural disasters and their influences on society. Our collection is primarily being used for academic and research purposes, therefore it is imperative that it remains accessible to the general public. In addressing our appraisal and selection policy, we look to five themes:

Future Users

Our goal with this collection is to show the human impact of the San Francisco Earthquake and Fire of 1906. It will combine items from multiple repositories about the event and allow teachers, students, and researchers to access a centralized location for these objects. It is our hope to put together a narrative with these objects and to create important discussions about how to use this knowledge to prepare ourselves for future earthquakes.

Feasibility of Preservation

At the moment our collection is limited in scope to only the 1906 event. It does not require a large amount of storage for preservation. It is well within our ability to preserve the digital objects and the descriptive data that is attached to it, within the repository we have constructed.

Legal and Intellectual Property Rights

We have full and complete legal and intellectual property rights as laid out in our contract with the donating parties. Part of this contract is to preserve the documents and keep them accessible to the general public. There are no restrictions on any of the current objects.

Mission

It is critical to our mission that these objects remain accessible to the public. There are important websites that maintain scientific information on seismology and specifically on the California earthquakes. There are also various repositories with objects documenting the event as it happened and in its aftermath. We are bridging the gap between these two worlds and educating the public on the science of the events with our main focus being on the human impact they have on our community.

Associated Data

The Descriptive Information associated with our objects is imperative to preserving their authenticity. They include the provenance of the objects which allow the tracking of ownership. The data also allows for the continued discovery of the objects within our database.

Data to be Archived

All objects pertinent to the 1906 earthquake and fire will be archived within our repository. This includes the postcards, photographs, and correspondence from other repositories and from donations.

Data not Archived

Any data that is outside of the city limits of San Francisco or secondhand accounts will not be included in the archive. Amongst the items listed above, there are some irrelevant objects within the newspaper clippings, photos, and correspondence. The newspaper clippings that are not within our remit include stories of people that live or work beyond the borders of San Francisco and include information irrelevant to our collection. Some of the photos received in the donation are unclear or damaged beyond recognition. Of the correspondence, two letters are not concerning the earthquake and one is damaged and illegible.

Rationale and criteria for the decisions

Our collection is very limited in scope so our requirements are defined by location, time period, and event. In order to be archived it must be from San Francisco, from April to May of 1906, and about the SF earthquake and fire and its human impact.

IV. Ingest

When the SIP (Submission Information Package) has been moved into the ingest action a health check will be performed. First it will first be placed into quarantine for fourteen days to

ensure that the objects are not corrupted and that no viruses or other malicious software are present within any of the objects. Once they have been quarantined the process of preparing the objects for long term access and storage will begin.

The objects will be assessed to ensure that all the required metadata is present and has been transferred correctly. The data will also be re-checked to ensure that no duplicates are present. Copies of the original bit-stream for each of the objects will be made at this point. After the quality of the digital objects has been re-assessed, fixity information will be attached to each object. The fixity information will be in the form of a check-sum and is placed to ensure that the data does not change during long term storage.

The digital objects will then be transformed, as necessary, to conform with the archive format policies. All images and postcards will be converted into TIFF format. All letters and newspaper clippings will be converted into the PDF format. Since we plan to only receive bitmap image files into our archive (and undigitized physical objects that will be scanned), ImageMagick (a free, open source image tool) will be used to convert any digital image we receive to a high quality open standard format. Of course, it's not possible to increase the resolution of an original image or its digital surrogate, but they can be enhanced for clarity to some extent. We've chosen to convert files to TIFF and PDF formats because they are mature, interoperable, stable, and widely used. These file types are among the most likely to be widely used for the foreseeable future.

V. Preservation Action

After the digitized images and their associated metadata which make up our collection - referred to in this context as Content Data Objects - have been ingested into the repository, a

process begins whereby Descriptive Information needed for search and retrieval is developed, and Representation Information is added to each content data object to create Content Information. Representation Information is fundamental "to [understanding] and [rendering] both the digital material and the associated metadata" (Introduction to Digital Preservation, n.d.), a definition which Oliver and Harvey distill even further to say that Representation Information "imparts meaning to an object's bit stream" (2016). The bit streams or "raw data" which make up digital objects are unrecognizable without the Representation Information which allows us to interpret them.

Representation Information can take a few forms: structural information, semantic information, and other representational information. By interpreting an object's bits in order to represent its file(s) type - with the assistance of a tool like DROID or JHOVE, which identifies and reports the specific file format versions of digital file - structural information pairs this detail with a specification of the particular data format. By using TIFF, the open standard format for high quality images, and PDF, the leading document format, we avoid complications with proprietary file formats . Semantic information supplies further details about the Content Data Object, including definitions of relationships between parts in more complex digital objects. These semantic details may not have any bearing on the renderability of the data, but can be very useful for understanding it. The semantic information within the Representation Information that will be enjoined to the Content Data Object can rescue digital objects from a lack of context. Other kinds of Representation Information include details about software, hardware and storage media; encryption or compression algorithms, or otherwise unrecorded documentation. In spite of what might seem like resource-demanding complexities inherent to Representation

Information, it's actually much easier to automate its creation than Descriptive Information which typically must be carefully created and added manually.

This new Content Information needs to be paired with Preservation Description Information (PDI) to ready the data object for long term preservation. Descriptive Information which supports the discovery and retrieval of Content Information by the repository's designated community will also be associated with the Content Information. PDI encompasses five components: 1) Reference Information; 2) Context Information; 3) Provenance Information; 4) Fixity Information; and 5) Access Rights Information. Reference Information provides mechanisms to create persistent identifiers which enable local and global access to specific content information. Context information describes the relationships between the Content Information and its environment, i.e. - other Content Information, or even documentation about why it was created in the first place. Descriptive Information, a separate object from PDI which is required for archival search, is derived in part from PDI and Content Information. It encodes information about the intellectual content of a digital object, most significantly the resource identifier that uniquely identifies the object and allows it to be discovered locally and globally via finding aids and search tools. Description Information also contains metadata to track essential provenance details (e.g. origin, enhancement and annotation) as well as key elements like title, author, date of publication, subject, publisher, description, and other fields from the data dictionary which were added to the object's metadata earlier (Gourley & Zhang, 2008). Packaging Information binds Content Information and PDI into a single logical package, and when combined with the Descriptive Information an Archival Information Package (AIP) is created for long term preservation.

Migration

Policies are needed to define what kinds of data and digital objects are candidates for migration, what file formats should be used and how often migration should take place. While migration is likely the most widely used preservation strategy, and one that is particularly well-suited to our collection insofar as it handles large collections of digital objects of similar / simple types very well, it nevertheless usually requires "artisanal, hand-crafted" work on each individual file, which is very labor intensive (Oliver & Harvey, 2016). This is primarily due to the substantial documentation that must be added to the various layers of metadata, a challenge that can't be side-stepped as it is imperative for ensuring the integrity and authenticity of migrated data. That said, the simplicity and uniformity of the types of digital objects in our collection offsets this requirement to some extent as they need less complex documentation and can recycle frequently used metadata as "classes" applied, for example, to Representation Information created during the Preservation Action stage.

The collection will be regularly monitored to ensure that all objects are in the latest non-proprietary file formats. Because migration fundamentally changes digital objects, introducing risks of compromising its integrity, we will save original bit streams of those objects as well. Riley describes how some repositories "[signal] when preservation actions such as a format migration or an integrity check should be undertaken" (2017), thus taking some of the burden off archivists to perform regular inspections of file formats and fixity information, though they can initiate manual checks at any time and trigger a migration if needed. If an archivist becomes aware of any impending hardware, software or file format obsolescence, a preemptive migration may make sense. We don't expect that a large percentage of our digital objects will require migration because, given their prior custody in many cases by sophisticated contributing

institutions, they are likely arriving at our repository already in the open and current standard formats that we require every digital object to use. Those that don't will be converted ("normalized") during ingest to either TIFF or PDF format. After any migration administrative metadata must be updated to register that event. Strict quality checking and enforcement will help minimize unwanted alterations to the object and guarantee its authenticity and integrity.

After a repository or archivist triggers a migration / reformatting of the Content Data Object, they are then moved from the Migration to the Transform stage. After this last transition is complete, the objects begin the curation lifecycle again by entering the Create and Receive stage for the second time. The reformatted objects will need validation and fine tuning of their existing metadata in the form of minor updates to Descriptive Information, Representation Information and PDI metadata to prepare for ingest and preservation.

Reappraisal

If errors or discrepancies are found in the process of reviewing an object's metadata and descriptive information, particularly in the administrative and integrity data, the Content Data Object should be reappraised, unless the archivist can amend these problems. Triggering reappraisal sends the files back to the Appraise and Select stage. Because we are investing strongly in editing / writing / encoding all the necessary metadata at each step in the receiving and creating, appraising and selecting and ingest stages of curation, we anticipate that very few ingested Content Data Objects will lack appropriate metadata or Descriptive Information.

Preservation metrics

Digital preservation in general can benefit from the formulation and use of metrics that provide insight into both public-facing performance indicators (number of visits, website hits, requests for reuse of archival materials, etc.) and "the efficacy of preservation processes as they unfold" (Dempsey and Lavoie, 2004) which would offer funder / stakeholders insight into the costs of ongoing preservation and new acquisitions. Looking at these metrics together might answer the question(s) of the organization's ROI. Further insight could be gained by benchmarking these statistics against similarly sized cultural heritage institutions.

VI. Store

The sustainability of our digital collection depends on long-term, secure and redundant storage. Refreshing data in and of itself - that is, copying digital objects to new storage mediums without changing the object - while certainly useful in many cases, particularly early on in the curation lifecycle, falls well short of a robust storage strategy. The emphasis in data refreshing on checking the accuracy of the digital objects on their new medium and documenting the refresh event in the objects' metadata, however, are best practices that apply to all efforts to copy archival data to more reliable storage. Refreshing should be considered a minimal intervention and all hard drives / USB sticks used should be removed from the office to some secure location after copies have been made.

A rich ecosystem of digital repository platforms, tools for managing data in archives, and cloud storage vendors present many options for creating a sustainable preservation system to

accommodate the needs of our digital collection. There are also numerous commercial and open source solution providers that offer preservation systems, some integrated with commercial cloud storage and various features and options for customization to solve specific problems faced by archivists in different organizational niches. There are several possible scenarios including self-hosting a repository on an office computer and backing up the Content Data Objects and exporting metadata for storage on local media devices, but this approach comes with excessive liabilities and risks. After evaluating a number of open source platforms like DSpace, Archivematica and Fedora as well as cloud storage providers, DSpaceDirect stands out as an excellent match for us for a variety of reasons. DSpaceDirect creates and hosts DSpace repositories in DuraCloud, a company created by DuraSpace, which also manages DSpace (DSpaceDirect, n.d.). Crucially, DSpace supports the use of the PostgreSQL relational database that contains our metadata and supports all the file and mime types found in our collection. Other advantages include the ability to build an open repository to display our image archive online, though there are other website solutions like Access to Memory, which works in conjunction with the archive and database that may offer a better fit for us (and we may want to keep our archive and website completely separate). Among its many features and benefits, DSpaceDirect creates additional backup copies of archived data objects and replicates the copy to DuraCloud, where automated health checks are performed regularly. DuraCloud itself is a "cloud of clouds" which offers cloud storage across multiple commercial and non-commercial providers (Amazon, Rackspace, Microsoft Azure, etc.) as well as preservation tools that enable security, access controls, transformation and sharing. Our security requirements are not as stringent as they would need to be for a repository storing private, sensitive or copyrighted data. It's difficult to gauge exactly how strong DSpace's security features are but according to DSpace Software

Support Policy documents they routinely monitor for and patch any security flaws. Based on the projected size of our collection, the DSpaceDirect Small plan, which offers 75GB of storage and costs $3490 per year, is more than adequate. Also attractive are the 10 free support requests per year and the ability to tap the large community of DSpace users and developers who are active on the DSpace forums. DSpace is a mature platform which has been around for nearly two decades (its initial release was in 2002), which is particularly reassuring for an organization like ours with a small, non-technical staff facing a somewhat steep learning curve in creating an sustainable, industry standards-based digital repository.

Physical Archives

Once a digital surrogate of a physical object has been made and added to DSpaceDirect, the original photo or correspondence will be stored in the SFPL's archives. A permanent office space will be needed for the archivist, which the SFPL has agreed to provide at one of their branches. Another possibility under discussion is hosting the archive within the Main SFPL branch's "History Center." In any case, the SFPL and SF Parks and Recreation have agreed to provide 75% of the initial costs of the collection. The remaining 25% will be funded through grants. The parent organizations will cover the salary of the permanent archivist collection. The ongoing costs of the collection upkeep and expansion will be largely covered by grants. However, the parent organizations may dedicate more money to the archive if it proves to be a popular resource with the San Francisco Bay Area community.

VII. Access, Use, and Reuse

The resources in the archive will be free for the general public to access; there are no intellectual property restrictions or copyright issues that would hamper open access to the collection. The SFPL and SF Parks and Recreation are both institutions with a long history of trustworthiness; partnering with these organizations will show users that the information provided is accurate and has been thoroughly vetted for provenance and other concerns.

General Access

Patrons of the SFPL will be allowed to sign into their online account and bookmark items within the archive that they would like to come back to. However, logging in is not a requirement. Visitors from outside of San Francisco or those who do not have a library account will still be able to access the full collection. The website will ask permission to use the visitors' IP addresses to determine their location, but this will not be required nor will it affect the patron's experience if they decline. Location information that is gathered will be used to determine the geographic location of users. This data will be used in the future when applying for additional funding, grants, etc. If the archive proves to be popular with Bay Area residents, it is likely that more funding will be allocated by the parent organizations to expand the archive.

During the ingest process, findability aids and fixity information (hash values), will be added to all the digital objects to allow continuous, authenticated access to the digital objects on the server side (see the ingest section above for more details). This section will focus on how the collection will ensure access on the client-facing platform. Since the collection will be hosted on a website most clients will not be consulting with the collection's archivist. It is important that

the objects on our website are findable and accessible to the general public. In this case, that means making sure that there are multiple avenues available for patrons to find what they are looking for on the website and making sure that the website is accessible for persons with disabilities.

Design

The archivist will work with SFPL's webmaster and programmers to design the information architecture of the archive's online platform. The archive will be accessible via a website, but depending on demand, may expand to include an app in the future. There will be multiple avenues for patrons to search for and find the information they need via the website. An overall directory will be created. The directory will organize the objects by medium (postcard, newspaper clipping, letter, etc.). Here all of the items will be listed alphabetically; this will be ideal for users who want to browse the collection without a clear idea of what they are looking for. Each item will also be "tagged" with descriptive information that includes the location, subject matter, and other features. Users will be able to use the search bar function to help them find specific items. For example, if a person was searching for objects that are about or feature the famous cable cars in San Francisco, they could type "cable cars" into the search bar and all items that show or talk about cable cars will appear. This function will be useful for patrons who know what they are looking for and have specific objects in mind. Each object will also link to related objects within the collection.

The archivist will also group related objects into mini-collections that will be featured on the homepage. This will include topics such as post-disaster encampments. Within each of the

categories all relevant postcards, photos, letters, and newspapers will be grouped together along with a brief description of the mini-collection and its relation to the overall disaster.

The webpages will have a uniform look and be kept as simple as possible while still allowing for an aesthetically pleasing look. The entire goal of the archive is to allow patrons to access the information they need with as easily as possible (Krug, 2013). If funds allow, user experience research will be conducted to further evaluate the accessibility of the digital objects by users.

Accessibility

The website will be designed with accessibility in mind. Information on how to make the website accessible will be taken from WebAIM's articles and tutorials and Dowden & Dowden's 2019 book on web accessibility. The most time consuming task will most likely be making sure the objects are accessible to screen reading devices used by those with vision impairments. This means that all photos and postcards will need to have <alt> text programmed in that gives a thorough description of the images. All newspaper clippings and letters will also need to be transcribed since screen readers will not be able to read the text within them. We will also need to make sure that all colors used pass accessibility compliance; this is especially important for users who experience color blindness. If there is any audio or video included, closed captions and transcripts will be provided to allow users with hearing impairments to interact with the content.

VIII. Transform

The collection will be regularly monitored to ensure that all objects are in formats that conform with the most recent standards. When a format standard is changed, the objects in the collection will be migrated to the new file format. The collection will always aim to use open source file formats to ensure that the majority of patrons will have access to the files without having to download software.

The archive is dedicated to open access and data sharing and reuse. It is the hope of the archive that historians, museum curators, and others will be able to use the archive's collections in the future to expand the reach and scholarship of the archive. An AIP that includes all metadata related to the digital objects will be made available to patrons upon request to use for exhibits, papers, or other projects.

Required Resources

We estimate that the initial acquisition, appraisal, and ingest process for the collection will take six months time and will require two full time employees (an archivist and a metadata specialist). The archivist will remain a full time employee and be responsible for working with IT and the parent organizations to ensure that the collection remains up to date with relevant standards. They will also be responsible for expanding the collection in the future, researching possible new objects to be added, and finding and applying to grant opportunities.

Sources

Amaral, M. (2010, June 30). METS for transferable metadata [blog post]. Retrieved from: https://easydigitalpreservation.wordpress.com/category/metadata/

Casualties and damage after the 1906 Earthquake. (n.d.) in Earthquake Hazards Program. Retrieved November 27, 2020 from https://earthquake.usgs.gov/earthquakes/events/1906calif/18april/casualties.php

Day, M. (2005). *Instalment on "metadata"*. DCC Digital Curation Manual. Retrieved from: http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata

Dempsey, L. & Lavoie, B. (2004). Thirteen ways of looking at...digital preservation. *D-Lib Magazine*, *10*(7-8). https://doi.org/10.1045/july2004-lavoie

Dowden, M., & Dowden, M. (2019). *Approachable accessibility: Planning for success*. Apress. https://doi.org/10.1007/978-1-4842-4881-2

DSpaceDirect - Hosted Repository Solution (n.d.) Retrieved December 5, 2020 from https://duraspace.org/dspacedirect/

Gourley., D. & Zhang, A.B. (2008). 4 - Metadata strategy. In *Creating Digital Collections* (pp. 31–53). Elsevier Ltd. https://doi.org/10.1016/B978-1-84334-396-7.50004-3

Guenther, R., & Wolfe, R. (2009). *Integrating metadata standards to support long-term preservation of digital assets: Developing best practices for expressing preservation metadata in a container format*: iPRES 2009. Retrieved from https://escholarship.org/uc/item/0s38n5w4

Introduction to Digital Preservation: Digital Curation Lifecycle (n.d.) in *Oxford LibGuides*. Retrieved December 6, 2020 from https://libguides.bodleian.ox.ac.uk/digitalpreservation/digitalcurationlifecycle

Krug, S. (2013). *Don't make me think, revisited: A common sense approach to web usability* (3rd ed.). New Riders.

National Archives. (2004). *The National archives' standard for record repositories*. British Crown. Retrieved from: https://www.nationalarchives.gov.uk/documents/archives/standard2005.pdf

Oliver, G., & Harvey, R. (2016). *Digital curation* (2nd ed.). American Library Association.

Prater, S. (2017). How to talk to IT about digital preservation. *Journal of Archival Organization, 14*(1-2), 90-101. DOI: 10.1080/15332748.2018.1528827

PREMIS Editorial Committee. (2015). *PREMIS data dictionary for preservation metadata* (Ver. 3). Retrieved from Library of Congress website: https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf

PREMIS Editorial Committee. (2017). *Guidelines for using PREMIS with METS for exchange*. Retrieved from Library of Congress Website: http://www.loc.gov/standards/premis/premis-mets.html

Riley, J. (2017). Understanding metadata: What is metadata, and what is it for? In *National Information Standards Organization Primer Series.* Retrieved December 1 2020 from https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf (Links to an external site.)

The Great 1906 San Francisco Earthquake. (n.d.) in Earthquake Hazards Program. Retrieved November 27, 2020 from https://earthquake.usgs.gov/earthquakes/events/1906calif/18april

OCLC/RLG Working Group on Preservation Metadata., OCLC., & Research Libraries Group. (2002). *Preservation metadata and the OAIS Information Model: A metadata framework to support the preservation of digital objects*. Dublin, Ohio: OCLC/RLG Working Group on Preservation Metadata.

WebAIM. (n.d.). *Constructing a pour website: Putting people at the center of the process*. Retrieved December 2, 2020, from https://webaim.org/articles/pour/