

Patrick Polglase

Rameen Kamazani

Bryce Danecker

Jess Fayer

William Drnach

Fernando Sierra

INST 442 - 0101 Digital Curation Across Disciplines

Professor Fenlon

Group 1 [Librarians]

Recommendation Report for Future Digital Curation Efforts

Introduction and Motivation

Our group worked with boxes 11 and 6. Each of the boxes contained a wide variety of documents such as different format types (vhs tapes, physical folders, guidebooks, etc.) The main directive of the project was to implement digital curation practices that documents the information gathered from the archived records. Our group accessed documents and attempted to find the most efficient way of cataloging the information stored within each item as well as the items themselves. The purpose of this report is to share our research and recommendations to the 4 H council so future digital curators may continue information preservation efforts with these archived boxes.

The items that we documented may still be of significant use to 4H members across the country. The information found within the items primarily consist of records of the organizational structure and protocol of the 4-H organization in the 1980s to the 1990s. The 4-H organization prides itself on storing and preserving knowledge so that it can then be used in communities that may not have readily available access to such beneficial information. Preserving information regarding the organizational structure of the 4-H organization allows people to view the development of practices and methodology of 4-H agents throughout the years. Some items in the boxes contained statistical data on tourist events and community

programs that information can serve as a helpful reference. By accessing the past, organizational directors within the current 4-H organization may choose to follow or disregard previous practices of 4-H protocol. The items also may be useful to the 4-H members across the country and in the future because some of the items are guidebooks that can be implemented in elementary schools within underrepresented areas of the general populace. The 4H itself seeks to foster data preservation efforts that can further contribute to the equitable dissemination of information for all groups of people.

Description Process

Our group discussed how to create categories. We first sought to identify the most common item types such as folders, vhs tapes, guide booklets, newsletters, and brochures and filed the items under said categories. The standardization of each of their descriptions involved documenting relevant information pertaining to that item type. For instance, a folder would have its title and year documented including a brief description as to the nature of the contents within the folder. Entries into the column titled “team member” would also adhere to a standard format of “first name” followed by “last name.”

To create the spreadsheet we first had to scan in all the documents using the EPSON V300 scanner and the macbook that was provided to us in our lab. After that we created the spreadsheet then as a group we decided what columns and rows to put and the name of them so we could differentiate them based on the documents and tapes. Once we all decided on what to name our columns and rows we then proceeded to finish scanning everything in and putting it under its designated column. The choice for our categories came from common similarities each item had with each other. For instance, some items were in folders, so we labeled any item with a folder holding multiple documents as folder items. Our fields came from the need to identify who curated the item and which day the item was published. We created a sheet that explains what each field and category means so that clients can know the rules of why we labeled them as such.

Standardization of Categories and Rule for Use

Field name	Definition	Rules for use
Inscribed Date	The date of the item as read on any writing physically inscribed on the item.	The date for this column will be read of the physical item and represents the date found on the item prior to any form of standardization. The initial date found on the item.
ISO 8601 Date Format	The date, in ISO 8601 format, of when the item was created or published.	The dates were formed based on what was written on the item. When entering the date for the item, the ISO 8601 standard format was used to input the year, month, and day into the format YYYY-MM-DD. The year is entered into the space YYYY, the month entered into the space MM, and day entered into the space DD. For items that omit a month and day, the year would simply be inputted as such a format is still compliant with the ISO standard. If no date was found then, we would just write "N/A".
Name Description	The names of each item recorded.	We wrote down the title of the item that was usually on the

		front cover or written on the cassette cover. If no name was clearly present on the folder, our group would simply use the title of the first page in the folder detailing the folder's contents as the name.
team member	The names of the team members that assessed and documented the item.	The rule was that each team member had to write down their name of who recorded what item. The entry would be in the order of First Name followed by Last Name.
Item Description	A brief description of the items contents, category, and appearance.	The description was primarily based on factors that would allow each item to be easily discernible from other items within the boxes. The description generally included the title of the item (if applicable), what category the item belonged to (folder, vhs, etc.), and the appearance of the item.
identifier	The order in which each item was accessed by the group.	When an item is accessed, a group member must document the number pertaining to the order of when the item was obtained from the box. For

		example, "1" indicates that the item was accessed first. A letter would be next to the number so we can group the items. The letter would range from a through f.
video length	The length of the video, in minutes, of items that consisted of a video format. If the item was not in a video format, "none" is applied.	This was exclusive for cassette tapes and if there was no run time written down, then we would write "none". We would write them in "hours:minutes" format.

Category name	Definition	Rules for use	mapping to complete list
VHS cassette tapes	Items that are stored in the VHS format that can only be watched using a VHS player.	Document the title of the vhs and document the time interval to which the VHS records.	1a, 1d, 1e, 1f, 2a, 2b, 2c, 2d, 2e
Folders	Items that consisted of a file folder with several other documents inside. This folder might contain the title of the subject along with the date.	Any item that contains several other items inside its file. Document the main purpose of the file and provide a brief description as to the nature of its contents.	1b, 1c, 2a, 2f, 3a, 3b, 3c, 3e, 3f, 4b
Booklets	Items that consisted of	Any item that is	4c, 4d, 4e, 4f, 5a, 5b,

	a guide format with pages and a table of contents. They usually have a plastic cover with staples on the spine.	consistent with a guide or book format. Document the main title of the guide and provide a brief description as to the purpose of the guide booklet.	5c, 5d, 5e, 5f
Newsletter	Items that consisted of a press cover and information catering to members of the 4 H organization.	Document the years the newspaper was published, the purpose of any events it advertises, and general information contained within the newsletter.	3d
Brochure	An informative paper document (often also used for advertising) that can be folded into a template, pamphlet, or leaflet	Any item that was small enough as a guest guide and had crease lines would be considered to be a part of the brochure category.	4a

Digitization process

During the digitization process we decided we wanted to digitize as many different types of files as possible. In our box we had VHS tapes, a brochure, books, and folders containing loose papers. We decided to digitize every different type of file except for the VHS tapes because we did not have the tools to digitize that file. First, William digitized a newsletter called “Epsilon

Sigma Phi Edition 59". Then, Patrick scanned in a folder called "1985 Annual Conference Ft. Worth, Texas". The folder, unlike the newsletter, contained multiple files that were related to the 1985 annual conference. After that Fernando, digitized the brochure. He scanned each page of this brochure. This brochure is called "Explore! Columbus Center" and it details a new center at the time in Baltimore. Next, Bryce scanned in a booklet called, "NAE4-HA On a Fast Track in Kentucky". Lastly, Jess scanned in another booklet called "climbing up" because there was no other type of file to digitize left. The digitization process went smoothly and we were able to digitize many different files of varying format types.

(6) Provide a detailed list of steps you took in digitization. This is the same process you submitted in Phase 4 but cleaned up and improved based on our feedback on your Phase 4 Submission.

Creating Digital Copies of Physical Media

Abstract

The goal of our team was to collaborate on ways to create digital objects of the physical objects within each of our archival boxes. During our lab session, our group used a single EPSON scanner and Mac laptop to process the items and convert them into a pdf format. The pdf files of the digitized items were then stored into our group's folder within the shared google drive for the project. Throughout the course of the lab, our group devised ways to best categorize and document various types of items. Below is our documentation of our methodology and process for creating digital copies of the physical objects.

Equipment/Tools:

- One EPSON V300 Flatbed scanner
- One Mac laptop

Steps for Digitization

Step 1 Prioritization:

When accessing the documents from each of the boxes, the group had to prioritize which of the contents of the item would be the most significant to scan. Our group did not have the extensive resources nor timeframe to scan and process many of the individual documents within items consisting of a collected folder. Our group decided that the best course of action would be to initialize the digitization process by identifying the most significant aspects of the document that would be used to discern the information contained within each item. When processing folders containing various types of flyers, transcripts, and employee protocols, our group would scan the first few pages found within the folder and attribute metadata to the scanned document file detailing where item the folder and its contents were sourced.

Step 2 Scanning Items:

Items were scanned primarily by the initial page or front cover of the item. For instance if the item was a guidebook, the front cover of the guidebook would be scanned using the EPSON V300 Flatbed scanner. The back cover of the guidebook would also be scanned to provide the person accessing the boxes a photographic representation of what the item looks like. In the case of folders, the cover of the folder itself should be scanned to provide an image of what exactly the item looks like. Secondly, the first document of the folder detailing its contents should be scanned to allow the item to be processed with the necessary information of what topic it is covering.

Step 3 Filing Digitized Items:

The final step of the digitization process entails creating the appropriate metadata properties of each now scanned item within the file name of the pdf.

- 1.) The file was processed from the EPSON V300 Flatbed scanner to a Mac laptop using a usb to type c cord.

- 2.) A pdf version of the scanned document would then be named with the title of the item and then what position it is located within the item. An example of the file name would be “Cat Guidebook - Playing with your Kitten cover.pdf.”
- 3.) Items with multiple scanned images would be concatenated into one single pdf file respective of each item. The combined pdfs would then be added to the digital repository in the form of our team’s collective shared google drive folder.

Future Considerations

(7) Describe any tasks your group intended to complete but did NOT complete as part of this project. If certain files were not described in your list, please provide detailed information about how the groups coming after you can pick up where you left off.

The aspects of the project that our group lacked in completion include further digitization of each of the item’s contents using the EPSON V300 Flatbed scanner. Initially, we intended to scan every page of the item from the folders and guidebooks assigned to our group. Our group lacked the proper time frame to accomplish this task so we instead sought to create easily understood curation protocols. Through adhering to the guidelines set in our spreadsheet and this report, a future curation team group may continue our group’s preservation initiatives

Disposition Plan

Category	Disposition option
VHS cassette tapes	We would suggest that any future digital curators would aim to create digital copies of the contents of the VHS tapes and then discard the physical formats due to their obsolete nature.

Folders	<p>We would suggest that any future digital curators digitize the contents of each folder as well as archive the most significant physical items relating to the historical organizational structure of the 4-H organization. For items that relate to negligible information about the organization such as banners, flyers, and event notices, we would suggest that these items are to be discarded upon digitization of their contents.</p>
Booklets	<p>Digitize the guidebooks so that their instructional information can be disseminated more widely via the internet. The physical copies of the guidebooks can either be archived for future use within educational environments or even donated to educational systems for their own use.</p>
Newsletter	<p>For newsletters, our group would recommend that these items be digitized and preserved in a storage database for future citation. The newsletters once served to inform 4-H employees about upcoming events, changes in organizational structure, and fundraising initiatives. The information possesses historical value to the organization as it provides information on how the 4-H organization was once operated and the events they held.</p>
Brochure	<p>We would recommend that brochures are digitized because of their historical content regarding 4-H hosted events. Once the items</p>

	are digitized they are to be discarded in order to alleviate physical crowding of storage facilities from physical items.
--	---

Our Recommendations

Recommendation A: Hazards, challenges, or concerns to be aware of when preserving digital vs. physical information in general.

The hazards and challenges that we encountered during the digital curation process were processing items that consisted of many individual pages. Due to the lack of time, our group was unable to process more than the initial pages of each item. The technology used to scan in the documents may vary for future digital preservation efforts. However, in the case of our group we found that it is vital to adjust the settings on the pdf processing software and the EPSON V300 flatbed scanner so that you are able to scan multiple documents if needed. Another problem we ran into was creating categories that we could group the items by. We would advise adhering to a metadata schema similar to that of Dublin Core when processing any of the items and their descriptions. Finally, we would advise future curation teams to implement standardized formats that will facilitate consistent information to be entered into each category. It is important to find operable machines that can process and access outdated format types such as the case with VHS tapes. We as students were unable to have access to a device that could process the information stored on the VHS tapes.

Recommendation B: Guide to formatting and preserving spreadsheets of data, best practices for spreadsheets, and what may go wrong and how to avoid it.

One recommendation we can think of is creating a simple identity for each column of data. For example, the “Name Description” column describes the title of the item while the “Item

Description” describes the physical look of the item itself. One recommendation for the best practice for spreadsheets is to have separate tables for different purposes. For instance, the table of the first sheet details our curation of our items. The second sheet has two tables, one for the field names and one for the categories. If those two tables were combined, it would be hard to tell which value is related to which row title. One way to avoid confusion of values is to communicate with teammates and discuss how they can navigate the table. If it is easy for them, then you do not need to change anything. If not, then you need to implement their thoughts into making the data simple enough to comprehend for most people.

Recommendation C: Future considerations or next steps with the archived boxes, or with other boxes in the collection. How should 4H move forward to serve their audiences?

One recommendation with the boxes is having more tools to find certain information in items. There were instances where we could find descriptions on one item but not one on a similar item. For example, when curating the VHS cassette tapes, some had the time length and others did not. Having a VHS player or contacting the organization who created the item would help curators find that information. 4H can have more tools to help curators with digitizing items into spreadsheets. A VHS recorder can record old footage where it can be uploaded into google drive for digital preservation.