Group 3 Data Management Plan

Group 3: Britta Anderson, Andre Cachoua Cervello, Billy Chao, Ceilidh McCallum, Hailey Rosquist

The Bicycle Revolution: Fashions and Freedoms

Project Description

Our team utilized Google Sheets to develop and organize our metadata, GitHub to host and manage the final website, and CollectionBuilder to structure and present our thematic digital project. Together, these tools supported a collaborative and adaptable workflow from start to finish.

Our project focuses on the personal freedoms people discovered during the Bicycle Revolution of the 20th century. By combining visual materials with structured metadata, we created a collection that allows scholars to explore questions about everyday fashion, bicycle customization, and diverse modes of transportation. This approach reveals how individuals used personal style and mobility to express independence.

Through CollectionBuilder and GitHub, we developed an accessible and visually engaging website that invites all audiences to discover how people used bicycles as a form of identity, self-expression, and participation in a changing cultural landscape.

Roles and Responsibilities

Andre - Repository Manager: The repository manager is responsible to make sure the collection follows the MAP and file format requirements for storing so it can be uploaded to CollectionBuilder, adding the collection description, data management plan (DMP), and data

documentation to the GitHub repository. The team also validates and troubleshoots any issues that come up during the upload process. In addition, I assist the project manager with finalizing the DMP and data documentation, take notes during group meetings, and manage the final setup of the GitHub repository and CollectionBuilder implementation, also leading the technical troubleshooting and handling any platform related problems.

Billy - Object preservation manager: The preservation manager is responsible for how different objects are added into the project's storage space and also works together with the development manager to finalize and choose what is the best file format for the project. I also must keep the standard of file naming and preservation formats. Most importantly, the object preservation must ensure backup copies of all project materials.

Britta - Collection Development Manager: The Collection Development Manager is responsible for deciding which materials should be included in the collection and works with the Object Preservation Manager to choose the best file formats for the project. They make sure that all copyright and licensing information is properly recorded and organized, and they write descriptions that explain what the collection contains. The manager also helps decide on research topics and defines the overall scope of the collection to ensure it stays meaningful and useful for the collection.

Ceilidh - Project Manager: Coordinated communication with instructor and team, and ensured infrastructure setup for the team’s internal virtual database. Prepared and coordinated the team for the Progress Report, Data Management Plan, documentation, and Final Project. The

project manager also ensured all meeting notes are recorded and stored properly in the Google Drive; also managed by them. Additionally, the project manager was responsible for submitting all of the team's final work to the instructor. To support the rest of the team, Ceilidh also helped refine and clean up all group work before submission.

Hailey - Metadata Manager: The metadata manager ensures that all objects follow the project's metadata application profile. Essentially, they make sure that all objects are organized, correct, and done in the same format each time. The metadata manager standardizes all data entry based on the requirements described by CollectionBuilder. Additionally, it is the metadata manager's job to create a master spreadsheet where all of the information is stored and maintained. There are also official standards that the metadata manager needs to follow such as Dublin Core and CollectionBuilder that help make the collection easy to search, read, and understand.

Andre:: Anticipated Data

For our data we standardized all image files to be under JPEG format, the configuration for the site is done on yml files, built on html, item information were parsed into csv's, with finally markdown files for the repo structure and organization. The license used on the repository was MIT formatted, about half of our items fell into the no rights category, while the rest had a copyright statement declaration that unless expressly stated otherwise, the organization that has

made each Item available made no warranties about the Item and cannot guarantee the accuracy of this Rights Statement, mostly followed from no creator items. Additionally 4 items had a creative commons rights license with a display of no rights reserved. Total data volume including repo and items comes to about 6.3 after compression.

JPEG images (.jpg): ~16 MB

(Standardized image files created from downloading archive items)

Metadata (.csv): <2 MB

(Metadata spreadsheet exported from Google Sheets)

YAML configuration (.yml): <1 MB

(Site configuration files for CollectionBuilder)

Markdown documentation (.md): <1 MB

(MAP documentation, project notes, DMP, and supporting documentation)

HTML/CSS/JS (website): ~3–5 MB

(Static site output produced automatically by CollectionBuilder)

Documentation and Metadata

Our project uses a structured metadata system that is based on our Metadata Application Profile (MAP). This incorporates Dublin Core and CollectionBuilder standards. All of our metadata is recorded in a master CSV spreadsheet that ensures each item is described consistently and is formatted correctly. This spreadsheet is published in our GitHub repository

where users are able to download it or view individual metadata fields through the item pages on our CollectionBuilder site.

To ensure that we are absolutely clear on our standards, we also provide a complete Data Dictionary. This table defines every field we used in our metadata, including its purpose, formatting rules, and required inputs. Additionally, it serves as the main documentation for understanding the extent of our data. The full dictionary is included below.

Field	DisplayName	Schema Map	ExampleValue	Definition	Data Type	InputRules	Obligation	Search
objectid			ba_laplpc_1886_photo_bikingacrosstown	A unique identifier used to ID each object, cataloger initials, repository acronym, date, format type, and object name.	Text	Use a short, consistent ID with no spaces. Follow the project's naming convention (for example: all lowercase, use underscores if needed). IDs should be stable; once assigned, they never change. Each ID must be unique across the entire collection.	Required	No
filename			bc_smithhm_1895-1899_samuelmurrayandbenjamineakinsonbicycles.jpg	The name of the digital file uploaded to the system.	Text	Follow the File Naming Standard Use the exact file name, including the extension (.jpg, .tif, .png, .mp4). No spaces are allowed. Use underscores if needed.	Required	No
title	Title	headline	The new woman and her bicycle	The name of the item as it should appear in the collection.	Text	Write a clear, concise title that summarizes the item. Capitalize major words and avoid overly long captions or descriptions. The title	Required	Yes

						should be short. Do not include dates, locations, or rights information in the title field because those belong in their own metadata elements. If untitled, give a brief descriptive title.		
creator	Creator	creator	“Doe, John” “Doe, John; Doe, Jane” “Unknown”	The name of the photographer or creator of the object. Also includes additional Creators such as composers, illustrators, and/or lyricists.	Text	Write the creator’s name as “last name, first name” to be formatted for DublinCore. The name must also be written completely in lowercase to be formatted for CollectionBuilder. Objects with multiple creators can be separated by using a “;” [semicolon] If an object has a blank creator field, or no known creator; use “Unknown” in place of.	Required	Yes
date	Date Created	dateCreated	Year: YYYY (1997) Year and month: YYYY-MM (1997-07) Complete date: YYYY-MM-DD (1997-07-16)"	Calendar date of creation of the object.	Integer	A resource may have several dates associated with it. The date covered by this table refers to the creation of the original resource, that is, when the resource was first created, before undergoing any conversion. If the date is unknown, specify an estimated date or date range	Required	Yes
description	Description	description	Black and white portrait photograph of Betzy Kjelsberg sitting on bicycle with both feet on pedals. She is	Brief written representation based on the image’s observation.	Text	Write a comprehensive, objective summary. Using complete sentences to focus on content, scope, key features, and primary	Required	Yes

			wearing a dark, long skirt and a bodice with large puffed sleeves, a high collar, and a dark cap.			subjects. While avoiding repeating the title.		
subject	Subjects	keywords	“Man; Portrait; Penny-farthing”	A set of keywords or topical terms that describe what the item is about.	Text	The topic of the resource. Typically, the subject will be expressed as keywords or phrases that describe the subject or content of the resource. The use of controlled vocabularies and formal classification schemes is encouraged. For multiple subjects, separate values with a semicolon.	Required	Yes
location	Location	contentLocation	New York, New York	Geographical location of the image.	Text	Use a clear place name, written as City, State/Region, Country when applicable. Use consistent formatting across all items. If the exact location is unknown, provide the most specific, reliable level (e.g., “Oregon,” “United States,” or “Pacific Northwest”). Do not include full street addresses.	Required	Yes
source	Source		Oregon Historical Society. Library. (Digital Collections)	Place of origin of creation or publishing for the image/object.	Text	The official name of the repository. Include sub-collection or location code if needed for unique identification within the repository.	Required	No
identifier	Source Identifier		b10826013	A unique accession number for individual objects.	Text	Use a consistent format across all items. Enter either a full handle path for each object; the identifier	Optional	No

						should reliably point to the digital object based on the individual who logged it. Should not change once published.		
type	Type				Text		Required	No
format	Format	encoding Format	image/jpeg	How the data and objects are stored and saved in the computer.	Text	Defines how data and content are presented to the user, and action as a blueprint for its appearance and behavior.	Required	No
language	Language		English	A language of the resource.	Text	Typically in English	Required	No
rights	Rights		The University of Minnesota believes that this item is in the Public Domain under the laws of the United States, but did not make a determination as to its copyright status under the copyright laws of other countries. The item may not be in the Public Domain under the laws of other countries.	Legal or moral permissions for tertiary use of the image or item. Human-readable copyright text.	Text	Look for any additional rights information provided by the source repository either on the individual object or within their Copyright information.	Required	No
rightsstatement	Rights	license	http://rightsstatements.org/vocab/InC/1.0/	This field is a standardized rights statement, designated in the form of a URI.	URL	Use a rights statement that matches the copyright status for an item. Required to use the RightsStatements.org URI addressed as the authority control. Do not hyperlink. CollectionBuilder will automatically transform the URI into hyperlinked text.	Required	No

display_template			image	Reusable format used to format how the data is presented.	Text	Use a CSV file for metadata to ensure all field names are lowercase without spaces or special characters. Plus specific naming conventions like object ID.	Optional	No
image_alt_text			A woman stands smiling beside her bicycle on a paved street. She wears a short sleeve blouse, a knee length skirt, and flat shoes. The bike has a basket on the front and a bag on the back. Behind her are flags, trees, and a large rounded building with tall windows.	A descriptive text that conveys the meaning of an image or graphic for people who cannot see it.	Text	Write 1–2 clear sentences describing what is visually happening. Keep it objective and non-interpretive unless context requires meaning. Use plain language and past/present tense naturally.	Required	No
citation	Reference		Oregon Historical Society. "Group of Riders at Cycle Trades Field Day in Portland, Oregon." OHS Digital Collections, 1926, digitalcollections.ohs.org/orglot1003-f11-001.	The citation or reference for each object/source where it can be found digitally.	Text	Using MLA format: Author, “Title of Web Page”, Title of Site, Publisher, Date, DOI	Required	No
cataloger			John	Name of the person who found and cataloged the object.	Text	Using only the first name of the student.	Optional	No

Storage and Backup

Google Drive: We used a shared Google Drive as our daily main workspace. It acted as a central hub where all 5 members of our team could collaborate. In the Drive, There are certain folders that holds the original, raw images we downloaded from the archives, our planning documents, and the master spreadsheet where we write out metadata. This Google Drive allows everyone to work on files at the same time and ensures we don't lose anything while we are gathering research.

GitHub: GitHub is where the finished project is held. It acts as our final storage and publishes our website. GitHub held the final objects, MAP, and metadata CSV file. GitHub is especially useful because it keeps a history of the changes we make. It can also protect our data from accidental deletion.

Personal computers: The Object Preservation Manager downloaded raw files from the drive to their computer to rename and reformat them. The computer acted as a third backup. While they processed the files, copies exist on my hard drive, keeping them safe while moving them from Google Drive to GitHub

Data Sharing

How we share our data: we share our entire collection publicly through our project website. Because the site is built on CollectionBuilder and hosted on GitHub Pages, access is open to anyone with an internet connection. There are no paywalls or logins required.

What we make available: we can make three things available: images, digital copies of photographs, and prints we processed. The metadata is the CSV spreadsheet that describes things like dates, locations, and descriptions of every object. The source code, since our project is hosted on Github, the code for the website itself is open source. Other students or developers can see what we used (CollectionBuilder).

Reuse Potential: We can expect others to reuse the data that we collected. Historians and researchers studying the bicycle revolution can use our data to analyze trends in transportation and social norms. Fashion students, in our collection, used the term “Fashion & Freedoms,” making it a resource to study the evolution of sportswear or women’s clothing reform. Teachers or educators can also use the collection. As a collection builder example or as a demonstration of the link between bicycles and social change.

Archiving strategy: We can use GitHub as our long-term public archive, unlike Google Drive, which is just for our private teamwork. The GitHub repository acts as a permanent public record. Even if the website were to go down, the repository on GitHub would preserve all our files, metadata, and code, allowing anyone to download or copy the entire project for their own use.

Period of Data Retention

The Project Manager has agreed to do monthly checkups on the object files and metadata to ensure consistency and no changes have been made from the source until the end of the current school year, ending in June of 2026.

Licensing and Ethical Issues

All of our objects are at minimum available for non-commercial use; including No Copyright, Copyright Undetermined, Copyright Not Evaluated, CC0 usage, and an In Copyright item that allows public use. Our purposes for these objects are entirely educational, non-commercial, and are to the benefit of open-access sources and information. For the

Licensing and Ethical standpoint, our group understands our project to fall under “CC0”; or “No Rights Reserved” as we do not have the final say in the publicality of these objects. Any further reference or use of these objects requires further research and communication with the original repository’s Licensing and Copyright agreements.

Appendix: Data Dictionary

Metadata Application Profile Available in our GitHub Repository and Linked as group3-map.

File Naming Conventions Available in our GitHub Repository and Linked as file-naming-standard.

Data Backup Strategy Available in our GitHub Repository and Linked as data-backup-strategy.