

Archival Data Assignment

Overview

Earlier in this class, we read Stefan Müller-Wille and Isabelle Charmantier¹ (a.k.a M-W & C) on how Carl Linnaeus stored plant descriptions in various different paper systems, how it connected his scholarly community, and how it shaped his evolving classification. As many of you noted, one of the big differences between their work and Ann Blair's was how firmly M-W & C described the actual paper operations of an individual data collection.

For this assignment, you will practice the same skills by going into the archives to find and investigate the origins of a dataset by looking at the paper it was stored on.

Project Timeline

Task	Date	Is this part graded/checked?
Choose an archival collection and set up a time to visit.	Weds March 2-4	Discuss in class on Monday March 2; register via e-mail Wed March 4.
Visit an archive	Week of March 2 or March 9	No
Present a problem in classes	March 23th or 25th	Obviously!
Write-ups due	Friday April 3rd by 5pm	Obviously!

Part I: Visit a dataset in the archives.

Find a pre-1945 organized, unpublished data source at a local archive or museum, **or (post-COVID) website** and look at it. We can easily each explore a different archive in the city. You may only use NYU's own archives by special permission.

Example materials

There's no universal definition of an "organized, unpublished data source."

Organized means you should strive to find something more structured than simply (say) all of a person's letters. It should be organized by its creators, not by the archivists after the fact. And if it is not obviously data, be prepared to argue why you think it should count.

Data is plural. It doesn't have to be numbers: but I roughly mean, "a record of many things of the same kind." Or think about it functionally—would this thing live in a database today?

Unpublished means you can't use a printed volume.

Some useful keywords to use in your search may be "ledger", "account book", "logbook", "log", "tables," or (sometimes) "file."

¹Stefan Müller-Wille and Isabelle Charmantier "Natural History and Information Overload: The Case of Linnaeus," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43, no. 1 (March 2012): 4–15, <https://doi.org/10.1016/j.shpsc.2011.10.021>.

In previous iterations of this assignment (in Boston), students have looked at:

1. 'Summitting' slips at the Appalachian Mountain Club archives: pieces of paper kept in glass bottles atop mountains on which 19th-century climbers wrote their names and other information to record their feat.
2. The records of a pharmacy at the Harvard Medical School's Countway Library for the History of Medicine. This was an overflowing book into which a pharmacist glued the actual scrip every time they filled a prescription.
3. Muster records from the Civil War, stored at the National Archives, showing how the citizens of particular Massachusetts towns enrolled in regiments into the Grand Army of the Republic, in formats that changed as the war went on and the draft began.

Take care in choosing a source; you probably want to look at a few. Double entry bookkeeping is notoriously difficult to understand, for example: and handwriting from before about 1860 can be extremely difficult to read. Don't be afraid to change your source in the archives for a better one.

Possible Archives

Post-COVID: Likely use online sources. Most of you have found these already.

Since our class is small, each student can choose a document from a *different* archive.

New York City probably has more and more varied archives than any city in the world. Only national capitals can compare. Many of the museums and libraries on the tourist track have associated archives that let you experience.

There are a number of general purpose libraries that will have all sorts of different things.

1. [The New York Public Library](#) has several different research centers, not just the one in the building with the lions. For example, if you want to work at the Schomurg Center for Research in Black Culture in Harlem, go to [their website](#), find the Rare Books division, and click on manuscript collections to a sense of what's there.
2. [Columbia's Rare Books and Manuscript Library](#)
3. The [New York Historical Society](#) has extensive manuscript records about 18th and 19th century New York.
4. [The National Archives and Records Center](#) has a major branch in downtown New York with US Government records. Materials created in Washington remain in D.C.; but this branch holds records created not just in New York and New Jersey, but also Puerto Rico, and the U.S. Virgin Islands. It is located in one of the more beautiful downtown buildings, the Alexander Hamilton U.S. Custom House, which also holds the (free admission) New York branch of the Smithsonian's National Museum of the American Indian.
5. The [Municipal Archives](#) collect government records from New York City as far back as the Dutch Colonial Era.

Many organizations have much more specific-purpose collections; either to preserve their own materials, or to collect in just one area. Pretty much every museum along Museum Mile on Fifth Avenue, from [The Metropolitan Museum of Art](#) to the [Museo del Barrio] has some special collections and archives. So do organizations like the New York Philharmonic, the there is an institution or subject area you're interested in, we'll see what we can find.

1. If you're not sick of plants yet, the [New York Botanical Gardens](#) have extensive personal papers and institutional records related to botany and plants.
2. The [Brooklyn Public Library](#) is distinct from the NYPL, and has hundreds of thousands of items related to Brooklyn specifically.
3. [The Wildlife Conservation Society](#), which runs the Bronx Zoo, the Central Park Zoo, etc., has archives which are closed in spring 2020.
4. [The New York Transit Museum](#) has information about the building and operating of New York's subways, buses, and elevated trains. They request two week's advance notice.
5. The [Museum of the Chinese in America](#), tragically, recently suffered a fire in its special collections.

Archival Visit Etiquette

Archives are different than libraries because most items in an archive are the *only copy* in the world. Archives have two jobs.

1. To help the public access and understand the irreplaceable materials in their collections.
2. To keep the items in their collections safe.

As a visitor, try to help them with both. Here in the United States, most archivists will be excited to help you find collections as long as you approach them professionally. The most important advice is that, especially if you are visiting anything other than the

- Many archives require advance appointments, especially smaller ones.
- Some archives contact the archive before you arrive to make sure they have the material on site, and that it's open to the public.
- Online descriptions of archival holdings can be scant: usually the fullest description of an archives contents is in a **finding aid** that you may have to visit the institution to see.

Archival policies on computers, photography, and so forth vary widely. At some archives, you can take all the pictures you want; at others, you can only take notes using *their* pencils on paper *they* provide. Follow the rules. If you have a digital camera, it's a good idea to bring it; ask whether you are allowed to include pictures in a class report, **and** if you could post it on the Internet.

Part II. Class Presentation

We will set aside time on three days for class discussion.

Everyone is scheduled for March 23 or March 25. Fallback may happen to March 30, but be aware that this leaves you little time to write up your paper.

If the first two letters of your NYU netid begin with the following letters, you're on Monday the 23rd: 'ri' 'as' 'mw' 'gs' 'ns' 'nb'

If it begins with the two following letters, you're on Weds. March 25. 'jr' 'dm' 'rm' 'ag' 'ry' 'bl' 'il' 'sb'

Please upload to NYU classes: there is a folder called "Archival Images and questions" where you should be able to create a thread with an image.

Part III. Writeup.

After presenting, you should write up your project as a 2000-2500 word paper. This will include your description of the artifact and any open questions you raised (hopefully slightly better developed than the presentation.)

This will be roughly 6-8 pages. It often makes sense to include images; if you do, they should be directly discussed in the text and as close to the paragraph where you discuss them as possible.

Your paper will include substantially the same material as your presentation, but should be structured more centrally around the single artifact you are discussing. Any additional items you viewed in the archive should be discussed *in the context* of that artifact.

The point of this is not historical argumentation, but *close and detailed description* that points to the limits of what *you* know about the artifact and what *can be known* about it. You will structure your paper in ways appropriate for the artifact.

You will also, as described below, offer a *plan for digitization*. This can simply be a separate section of the document. It should rarely

You may use the first person in describing the materials and your opinions, but you should not describe your personal experience. (Hopefully you will feel a shudder of delight when you crack open an old leather book; but keep that language out of the paper.) Do **not** include a description of the archive/institution in general, except in cases where it actually affects the data you're looking at.

Connection to secondary sources

Also, branch out beyond the source to provide some minimal context based on published sources. (These should probably be book or journal articles, and probably not web pages. Try to use the library catalog, not Google, for this.) You should be predominantly engaging with your source; one of the purposes of this is to help you see if there are connections to explore further in later work in this class.

Be sure to cite these sources properly. (See below).

Digitization Plan

Outline—but do not implement!—a plan for digitizing the data here into a form that could be used for further research. If you were going to store it in a digital spreadsheet or database, what sort of fields would you collect information on. Could you store the information in a spreadsheet or database? What sorts of questions could you answer by having the entire dataset digitized. (Assume that you have all the technical analysis capabilities needed to do so). How much time and effort would it take to create a digital version? What aspects of the document might be lost in the transition?

Submission

Turn in your paper via e-mail to bs145@nyu.edu by 5pm on Friday April 3rd.

Citations and acknowledgements

You should cite the works that you quote and refer to in the text in a *consistent* format. I recommend the [Chicago documentary note format](#): with it, you give a full citation the first time you use a text, and smaller ones later.

For short papers like this, you can omit the final bibliography. If you prefer to use a social-science author-date format with final bibliography, that is also acceptable.

If you are worried about formatting your citations correctly or keeping track of the sources you use, I strongly recommend the open-source citation software [Zotero](#). This will automatically pull citations from the web, and you can drag and drop into a paper to get a formatted citation. Just be aware that online library sources may give you extraneous information, such as the language or a URL. Edit the fields in the library until drag-and-drop gets you good results.

Your primary source for this assignment will be the archival document(s) that you describe: be sure to cite it according to the standards of the archive. This will mean you've described it comprehensively enough so that a future researcher could easily find it. In general it includes a description of the collection, box, and folder.

You should also acknowledge any archivists or peers who help you to better understand the materials. Such acknowledgements would typically come either as a footnote to the first paragraph (for general assistance) or as a footnote to the specific place you received help. Do not thank me, though.

Müller-Wille, Staffan, and Isabelle Charmantier. "Natural History and Information Overload: The Case of Linnaeus." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43, no. 1 (March 2012): 4–15. <https://doi.org/10.1016/j.shpsc.2011.10.021>.