

Bodies in Transit:

A digitization project of New York Department of Health records currently held by the
New York City Municipal Archives

Overview

The New York City Municipal Archives' "Bodies in Transit" collection documents the movement of corpses through New York City between 1859 and 1894. Little is known about the ledger itself, which was obtained from the Department of Health and whose creation was the result of laws mandating records of transported bodies (in part in order to track and prevent the spread of disease). The digitization, transcription, and application of these records — housed in both print and microfilm format — will occur in several, sometimes concurrent, phases. In order to provide useful deliverables, the team must identify historical context for the ledgers; produce digital files for a reasonable part of the content of three (3) microfilm reels; transcribe the ledgers from the digital files; and suggest analysis, infographics, and/or user interfaces based on the information gathered.

Client Description and Objectives

The Municipal Archives, directed by Sylvia Kollar, house and provide access to NYC municipal documents. According to the institution's website, its collection includes a mix of "office records, manuscript material, still and moving images, ledger volumes, vital records, maps, blueprints, and sound recordings," in formats such as printed material, micrographics, and photograph negatives. Because the vital records collection holds information about reported births, deaths, and marriages in all five boroughs, it is often consulted by individuals conducting genealogical research. The fundamental objective of the "Bodies in Transit" digitization process is the reproduction of the ledger in a format that can be searched, catalogued, and integrated. Although we cannot expect to completely transcribe the entire microfilm collection, our work will begin the move toward a versatile database. Because we will not be working within the Archives' database or information management systems, the primary deliverables will consist of digital images based on the ledger microfilm and a spreadsheet document housing transcriptions and metadata about the documents. This information can be used in conjunction with other data held by the Municipal Archives to produce on-site exhibitions, drive projects, and ultimately support a user-facing web-based interface. Digitized images will be provided for approximately one (1) microfilm reel, and we project that transcription will be completed for the first one hundred (100) frames. Finally, we expect to produce ideas and/or prototypes for potential data applications, outlined below.

Project Team and Responsibilities

The team consists of Rubi Mora, Shannon McDonald, Victoria Harty, and Grace Afsari-Mamagani, with additional support from Professor Wolf and possible Rob Koehler. Rubi is serving as a project manager by facilitating communication between the group and the New York City Municipal Archives. Shannon is both facilitating project development by providing frameworks for milestones/deliverables and contributing to preliminary contextual research. Victoria will play the role of primary research manager, collecting historical and theoretical information and compiling the research contributed by other team members in order to provide a rich context for the content being digitized. Grace will serve as the data manager, monitoring details like naming conventions,

conducting initial analyses, and testing potential technology solutions that might prove valuable to the client. All group members are responsible for conducting research on post-mortem transportation in/through New York City, converting microfilm into digital format in the Bobst Library Microfilm Center, and transcribing data into a usable format.

Component I: Microfilm Conversion and Information Architecture

Objective and Methods: Prepare data from Municipal Archive microfilms for transcription by producing accessible digital images. Resources used include microfilm readers and scanning systems enabling the conversion of microfilm frames to PDF documents to be shared and converted into plain text. The resulting files will be stored locally on removable storage and in the cloud via Google Drive. For each page digitized, file metadata is entered into a shared Google Sheet using an established naming convention.

Metadata components necessary for proper retrieval of data:

- Digital File Name/Identifier (“collection_volume_frame.pdf”)
- Microfilm Reel Number (Location in Collection)
- Approximate Frame Location (on reel)

Bounds and Timeline: Based on initial attempts, approximately 100 microfilm frames can be digitized and expert over the course of two hours. Each team member (Grace, Rubi, Shannon, Victoria, and Nick/Rob) should spend a total of approximately four (4) hours digitizing over the course of two weeks.

- Goal: ~1000 digitized microfilm frames by the first week of November 2014; ideally, the entirety of one reel

Component II: Data Transformation/Transcription

Objective and Methods: Replicate the information structure of the Department of Health ledgers in a digital, searchable format using a collaborative Google Sheet. Transcription will take place by hand rather than using optical character recognition due to the nature of the documents. The resulting spreadsheet constitutes a primary deliverable, as it can be exported to a .csv or other file format and imported into a cataloging system. Furthermore, the data can be easily adapted for research or visualization and integrated with other data sets for presentation in the exhibitions that the Municipal Archives are currently planning. Until exported, stored data is only accessible to team members; the final system architecture will be determined by the client post-project.

Bounds and Timeline: Transcription can occur concurrently with digitization. Each team member will transcribe ~20 ledger frames over the course of one (1) month, mid-Nov. to mid-Dec. 2014.

- Goal: Approximately 100 fully transcribed and proofread ledger pages; nearly 5,000 individual records

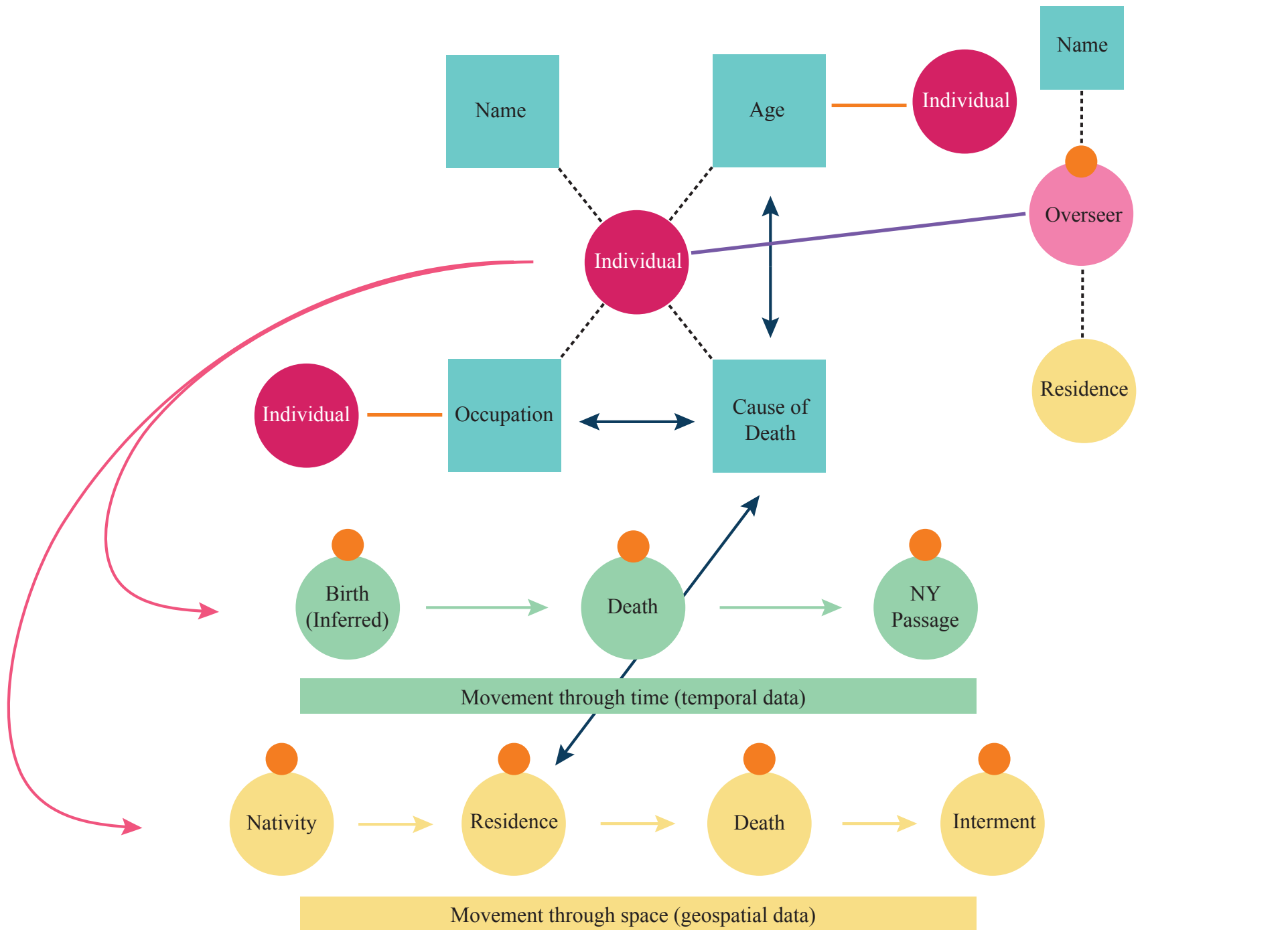
Data Model

Date of Passage through New York	Name	Age Year	Age Month	Nativity	Residence	Place of Death	Disease
----------------------------------	------	----------	-----------	----------	-----------	----------------	---------

Date of Death	Place of Interment	Name and Residence of Person having charge of the Body	Occupation	Remarks
---------------	--------------------	--	------------	---------

Data Structure

Data Types: Dates, Numerical data (age), Locational data, Names, Category (cause of death, occupation)



..... Characteristic/Attribute

———— Caretaker relationship

————> Geographic movement

↔ Potential causality

————> Chronological movement

● Points of potential connection via coincidence/shared traits (individual-individual)

Component III: Research and Prospective Applications

Objective and Methods: Establish a contextual framework through extensive research in which to situate the “Bodies in Transit” data, and consider ways in which the data might be fruitful for the client, external researchers, and the general public. If possible, retrieve provenance and creator information. Collect relevant documents and paths to other sources primarily using Google Drive; work with archive employees and subject librarians to identify useful sources and potential uses for transcribed data. Provide examples of technological interfaces, tools, and visualizations, such as maps of body routes or infographics linking occupation to cause of death.

Resources: library materials, online databases, Sylvia Kollar at the Municipal Archives, similar DH projects produced elsewhere (see model projects list), Margaret Smith (Physical Sciences and History of Science Librarian) and connections at the Department of Health vital records division

Bounds and Timeline:

- Ongoing research from all team members from mid-Oct. to late Nov., ~6 hours/person
 - topics of interest include public health, history of disease and causes of death over time, history of post-mortem transit records, immigration, occupational hazards, the Civil War, and others
- Test potential visualizations and provide UI recommendations in late Nov., ~10 hours
 - technologies of interest include KML/Google Maps, Neatline, network visualizations (Gephi, D3.js), timelines and sliders representing data against historical events/trends, static visualizations like graphics and charts, and robust database searchability
- Metadata potentially of interest: subject, relation, date, accrual, marginalia, gaps/page wear

Data Formats, Dissemination, and Preservation of Access

While in use by the team, all data is shared in a private cloud-based environment. Final data access decisions lie with the client, but information can be easily exported in sustainable formats (like the .csv) for storage or integration into cataloging systems and other data sets.

This project is less concerned with user experience than with plain data entry for — initially — internal use/knowledge. Once the data is added to the Archives’ collection, the user interface will likely feature search functionality. As digitization continues over time, the data might allow for a user-facing digital portal that provides analysis, allows for user manipulation, creates links between data types, and hosts other visual representations to supplement the bare data (which itself can be made downloadable).

Bibliography

- Cliff, Andrew and Peter Haggett. "Time, travel, and infection." *British Medical Bulletin* 69.1 (June 2004): 87-99.
- Condran, Gretchen. "Changing patterns of epidemic disease in New York City." In *Hives of sickness: public health and epidemics in New York City*. Eds. Rosner, D. New Brunswick, NJ: Rutgers University Press, 1995.
- Condran, Gretchen and Eileen Crimmins-Gardner. "Public Health Measures and Mortality in U.S. Cities in the late 19th century." *Human Ecology* 6, no 1 (Mar 1978): 27-54.
- Duffy, John. "Nineteenth Century Public Health in New York and New Orleans: A Comparison." *Louisiana History: The Journal of the Louisiana Historical Association* 15, no 4 (Autumn 1974): 325-337.
- Finseth, Ian. "The Civil War Dead: Realism and the Problem of Anonymity." *American Literary History* 25, no 3 (Fall 2013): 535-562.
- Harris, Elisha. *Report on plans for securing complete and authentic records of deaths and the causes of death in the United States*. Cambridge: Riverside Press, 1878.
- Humphreys, Margaret. "No Safe Place: Disease and Panic in American History." *American Literary History* 14.4 (Winter 2002): 845-857.
- Laderman, Gary. "Locating the Dead: A Cultural History of Death in Antebellum, Anglo-Protestant Communities of the Northeast." *Journal of the American Academy of Religion* 63.1 (Spring 1995): 27-52.
- Lederberg, Joshua. "Infectious History." *Science* 288.5464 (April 2000): 287-293.
- Lee, Chulhee. "Prior Exposure to Disease and Later Health and Mortality: Evidence from Civil War Medical Records." *Health and Labor Force Participation Over the Life Cycle: Evidence from the Past*. National Bureau of Economic Research. Chicago: University of Chicago Press, 2003.
- Pike, Martha. "In Memory Of: Artifacts Relating to Mourning in Nineteenth Century America." *Journal of American Culture* 3, no 4 (December 1980): 642-659.
- Scott, Sean A. "'Earth Has No Sorrow that Heaven Cannot Cure': Northern Civilian Perspectives on Death and Eternity during the Civil War." *Journal of Social History* 41.4 (July 2008): 843-866.
- Susser, Mervyn and Ezra Susser. "Choosing a future for epidemiology: I. Eras and paradigms." *American Journal of Public Health* 86.5 (1996): 668-673.
- Second Annual Report of the Board of Health of the Health Department, City of New York*. New York: David H. Gildersleeve, 1872.

Woods, Margaret E. and Peter S. K. Chi. "Sanitary Reform in New York City in 1866 and the Professionalization of Public Health: A Case Study of Social Reform." *Sociological Focus* 19.4 (October 1986): 333-347.

Projects of Interest

<http://www.tlarchive.org/htm/home.htm>

The Living City, a Columbia University project meant to “capture the experience of life, health, and urban transformation during the decades between the end of the Civil War and the end of World War I.”

<http://connecticuthistory.org/death-and-mourning-in-the-civil-war-era/>

Connecticut History, a Connecticut Humanities project in collaboration with the UConn Digital Media Center, that includes information about “Death and Mourning in the Civil War Era.”

<http://caribbeancholera.org>

19th Century Caribbean Cholera TimeMap, a medical history map and timeline from Duke University’s Haiti Project.

<http://dsl.richmond.edu/civilwar/>

Hidden Patterns of the Civil War, a collection of related projects at the University of Richmond Digital Scholarship Lab, including emancipation data that shows migration patterns/vectors.

<http://orbis.stanford.edu>

ORBIS, The Stanford Geospatial Model of the Roman World, which integrates both travel through geographic space and network analysis