

Bibliographic and textual studies and the personal library

Alexandra Wingate, Indiana University, Twitter @alextheknitter, ORCID 0000-0003-3863-8141

John A. Walsh, Indiana University, Twitter @jawalsh, ORCID 0000-0002-1824-210X

Caroline Nurkkala, Indiana University

Daniel Evans, University of Illinois at Urbana-Champaign, Twitter @djohnnevans

Alyssa Mertka, Indiana University

Jennifer Christie, Indiana University

Category: long paper

Attendance: In-person

Introduction

This paper will discuss the application of digital humanities methods to understanding private libraries in three contexts: a corpus of post-mortem library inventories from early modern Navarre, Spain, the library of Isaac Newton, and the library of Victorian poet Algernon Charles Swinburne. The particularities of each case facilitate different analysis techniques from the visualization and quantitative analysis of private libraries to comparative textual analysis of full-text corpora of an author's library and the author's own works. The number of viable techniques is affected by factors such as how much can be known about the library's contents, how much is known about the owner, and the availability of quality OCRed texts of the books in the library from sources such as the HathiTrust Digital Library.

Navarrese private libraries

In the case of Navarrese private libraries from the 16th and 17th centuries, information about the contents of the libraries comes from post-mortem inventories: court documents generated about a deceased person's goods soon after death. Since these inventories are created for legal and economic purposes and not for historical reconstruction of the libraries' contents, they often do not list ephemera or provide complete descriptions of individual books. Furthermore, these inventories only provide a snapshot of the library's contents at the death of the owner, meaning that any additions or deletions from the library over time are unknown without other sources of information about the library. The owners are also most often relatively ordinary people—they have not been the subject of years of meticulous study as with figures like Newton and Swinburne. This hinders the ability to use biographical information to contextualize the quantitative analysis of their book collections.

Despite these drawbacks, it is possible to assign metadata to the books listed in inventories—the amount of metadata will simply vary based on how much information is given. For example, these two entries from two different inventories represent two extremes in the amount of information given about books:

Iten veynte y un libros pequeños y grandes y medianos de Romance, todos de diferentes tomos (Archivo Diocesano de Pamplona, C/618 N. 10, f. 82v) *Also, twenty-one books, small and large and medium in Spanish, all different works*

“Ytten vocabulario juris pequeño ynpression de Leon del año de sesenta y una en pergamino” (Archivo Real y General de Navarra, no. 89984, f. 46v) *Also, a small vocabulario juris impression of Lyon in the year '61 [bound] in parchment*

On one hand, the first example only gives useful metadata about the number and language of the books (21 and Spanish). While the description does describe size, it does not give the exact format (folio, quarto, octavo, etc.) nor the number of books of each format, making it less easily analyzable from a quantitative standpoint. On the other hand, the second example, gives size (small), an approximate title, place and year of publication (Lyon, 1561), and binding information (parchment), which makes it possible to identify the book as Antonio de Nebrija’s *Vocabularium utriusque juris* (Lyon: Jacques Faure apud Jacques Giunta, 1561) in octavo (USTC 153208). Even with varying amounts of metadata derived from each inventory entry, we can compile statistics and generate visualizations about individual libraries, libraries of a particular subset of the corpus (e.g., by profession), and the complete corpus.

Isaac Newton

The second case, the library of Isaac Newton, provides more opportunities for analysis since both Newton and the contents of his library are much more well-known and more fully documented. Newton’s personal library has been documented in John Harrison’s *The Library of Isaac Newton*, and we have transformed the data from this print catalogue into a spreadsheet. By cleaning this data and adding additional metadata to the catalog entries, as with the Navarrese libraries, we can analyze and create visualizations of Newton’s library.

In addition, as part of the *Chymistry of Isaac Newton project*, Newton’s alchemical manuscripts and his citations to alchemical secondary sources are being transcribed and encoded in TEI alongside a TEI-encoded bibliography of the works cited in the manuscripts. By comparing Newton’s cited sources to his library and his patterns of citation across the manuscript corpus, we can begin to identify books Newton used but did not own at his death as well as trace the development of his library and his documentary and reading habits in the context of his alchemical study.

Swinburne

Our third case is a digital literary research effort to study influence and intertextuality in the works of Victorian poet, novelist, and critic Algernon Charles Swinburne. This line of research is particularly relevant for Swinburne, a poet whose work and life are steeped in books, documents, and literature. Swinburne’s friend William Morris wrote in an 1882 letter: “I never could really sympathize with Swinburne’s work; it always seemed to me to be founded on literature, not on nature.” Morris draws a sharp distinction between literature and nature, but Swinburne recognized no such distinction. For Swinburne, poetry is omnipresent in the natural world, and books are—like a flower, a child, or his beloved sea—living things. In a poetic manifesto, the “Dedicatory Epistle” that introduces his collected *Poems* (1904), Swinburne writes:

The half-brained creature to whom books are other than living things may see with the eyes of a bat and draw with the fingers of a mole his dullard’s distinction between books and life: those who live the fuller life of a higher animal than he know that

books are to poets as much part of that life as pictures are to painters or as music is to musicians, dead matter though they may be to the spiritually still-born children of dirt and dullness who find it possible and natural to live while dead in heart and brain. Marlowe and Shakespeare, Aeschylus and Sappho, do not for us live only on the dusty shelves of libraries.

Swinburne's own library, or much of it, was sold at auction on 19-21 June 1916. The catalogs from the Sotheby's sale are published in volume 6 of *Sales Catalogs of Libraries of Eminent Persons* (1972), edited by John Woolford. We transcribed these auction catalogs to compile a digitally encoded bibliography. The bibliography, encoded as TEI/XML, is populated with links to full-text sources—from the HathiTrust Digital Library and other repositories—of the works in Swinburne's library. With the TEI-encoded bibliography, we may apply XSLT stylesheets and XQuery queries to explore and visualize attributes of the library, such as languages, publication dates and places, and frequency of works by particular authors.

Using the full-text sources, we have compiled a corpus of the contents of Swinburne's library and are applying text data mining and computational analysis to the corpus. We will then compare these results to a similar analysis of Swinburne's own works. This comparative analysis will, we hope, reveal similarities in vocabulary, syntax, and diction and identify similar passages or paraphrases that might otherwise escape detection. Our paper will provide an overview of the project and our findings to date.

It is worth noting that this comparative analysis is possible not only due to the TEI-encoding of Swinburne's works, but also because so many of the books in Swinburne's library are from the machine-press era, leading to fewer errors in the OCR'd text. Newton's alchemical manuscripts are TEI-encoded as well, but the lower OCR quality of early modern printed books makes a similar analysis more difficult.

Conclusion

The application of digital humanities methods to private library research allows scholars to go beyond listing and describing the books in a library or libraries. Combining and remixing data sources facilitates new analyses that shed light on library owners, the libraries themselves, book usage, and even books consumed by the owner beyond those associated with their library at their death.

Our presentation will provide additional detail on the particularities of each project and the current results of research into each library, illustrated by data visualizations derived from our digital representations of the libraries.

Bibliography

[Cura de almas vs. Domingo de Benedo]. (1656). Procesos (C/618 N.10), Archivo Diocesano de Pamplona, Pamplona, Navarre, Spain.

Harrison, J. (1978). *The Library of Isaac Newton*. Cambridge University Press.

[Pedro de Casanova y Antonia García Jiménez contra Beatriz de Azcarraga y Leache]. (1620-1629). Procesos judiciales. Solano-Pendientes (no. 89984), Archivo Real y General de

Navarra, Pamplona, Navarre, Spain.

Swinburne, A. C. (1905). Dedicatory Epistle. In *The Poems of Algernon Charles Swinburne*. London: Chatto & Windus.

Woolford, J. (Ed.). (1972). *Sales Catalogs of Libraries of Eminent Persons*. London: Mansell.