Archaeologies of Reading: Modeling and Recreating the Annotation Practices of Gabriel Harvey, John Dee, Jacques Derrida, and the Winthrop Family

Jean Bauer
jabauer@princeton.edu
Princeton University, United States of America

Earle Havens
ehavens2@jhu.edu
Johns Hopkins University, United States of America

Matthew Symonds
m.symonds@ucl.ac.uk
University College London, United Kingdom

Anthony Grafton
grafton@exchange.princeton.edu
Princeton University, United States of America

Jennifer Rampling
rampling@exchange.princeton.edu
Princeton University, United States of America

Katie Chenoweth
kac3@exchange.princeton.edu
Princeton University, United States of America

Christian Flow
cflow@exchange.princeton.edu
Princeton University, United States of America

This panel brings together three projects that explore and analyze annotated texts in digital environments: The *Archeology of Reading in Early Modern Europe*, The Winthrop Family on the Page, and Derrida's Margins: . All three projects seek to give users the experience of how another person read their books (see Notes). While each is fascinating in its own right, collectively these projects span centuries and provide a powerful and instructive lens on data modeling, inter-institutional collaboration, and interoperability.

The connections between these projects, both on a scholarly and an institutional level provide a powerful case study in creating distinct but interoperable digital humanities projects and resources. Both Winthrop and Derrida's Margins are being built in the Center for Digital Humanities at Princeton University, while AOR is a collaboration between Johns Hopkins, University College London, and Princeton. AOR and the Winthrop project share a co-PI, Anthony Grafton, and two books owned and heavily annotated by John Dee are in the Winthrop library and thus came to New England. However, while the AOR project looks at a single renaissance annotator (Gabriel Harvey in Phase I– the project is heavily indebted to the methods and questions outlined in Jardine and Grafton, 1990–, and John Dee in Phase II), the Winthrop project looks at generators of annotators (men and women) making it as much a prosopographical project as a study of annotation. Like AOR, Derrida's Margins focuses on a single scholar's interaction with his library, but Jacques Derrida's annotation practices differ dramatically from Harvey or Dee, and French copyright law poses serious challenges to representing the book pages online.

All three projects make use of the International Image Interoperability Framework (http://iiif.io/), but differ on the rest of their underlying data structures. AOR uses a custom XML schema to encode a critical edition of the annotations. Winthrop will use a custom prosopographical-bibliographical relational database, implemented in Django. While most of these relationships will be attested through annotation, the system will also record other connections (including purchasing records, books referenced in letters, etc). The annotation data model will most likely be created using a graph database, given the multiple types of annotations and the ways they can be expressed. Derrida's Margins will be another custom relational database, built in Django, but will be fully bi-lingual, allowing users to search and browse in English and French. The Winthrop and Derrida database will also be exposed as Linked Open Data.

All three projects are committed to interoperability, with a key goal being the ability to search across the three projects once they are complete. The projects are in close contact the IIIF Editors and the W3C Web Annotation Working Group to ensure adherence to and input on best practices and emerging standards.

The Archaeology of Reading in Early Modern Europe

Earle Havens, Matthew Symonds, Anthony Grafton

The Archaeology of Reading in Early Modern Europe (AOR), is an international collaboration among the Sheridan Libraries at Johns Hopkins University, the Centre for Editing Lives and Letters (CELL) at UCL, and the Center for Digital Humanities at Princeton University Library, with funding from the Andrew W. Mellon Foundation. It is also an interdisciplinary collaboration between historians, librarians, software engineers and data scientists.

While the body of scholarship on the history of early modern reading practices has burgeoned during the past several decades—guided in large part by the initial scholarly work of our project partners, the late Professor Lisa Jardine and Professor Anthony Grafton—as a collective body of knowledge the history of reading has nonetheless remained limited to isolated, partial, and impressionistic studies of single texts read by single annotators.

As researchers, we conduct this work in the conspicuous absence of comparative evidence of the larger range of early modern historical reading practices, strategies, and agendas. Scholars also find it physically impossible to effectively penetrate the dynamic array of information preserved in annotated books for the purpose of systematic analysis owing to their sheer density of content, in many instances, relative to the original texts on which the annotations comment.

During the original planning workshop on annotated books that led to the formation of the project, leading scholars, librarians, curators, and technologists agreed collectively that these constraints imposed upon the study of early modern annotated texts in their original analog form could only be overcome satisfactorily when annotations are treated as data sets that can be mined and analyzed effectively in more versatile, enriched, and readily searchable digital forms.

The AOR team has elected to focus on a distinct and roughly contemporary dyad of clearly identified early modern readers: Gabriel Harvey in Phase 1 (2014–2016) and John Dee in Phase 2 (2016–2018). While the identities of a large majority of early modern annotators remains unknown in extant collections, this focus on known readers will enable the project team to analyze and more precisely situate the processes of reading and annotation within their respective historical contexts.

In order to make these sources more accessible to analysis, the AOR technology team – based at Johns Hopkins University's Digital Research and Curation Center – is working closely with the International Image Interoperability Framework (IIIF) protocol and community to develop features and use cases that will enhance AOR within this larger framework.

IIIF features a set of protocols, application programming interfaces (APIs), and shared technologies for the presentation of web-based images. In the case of AOR, these images are digital surrogates of rare book materials containing manuscript annotations by Gabriel Harvey and John Dee, which users interface through AOR's adapted version of the Mirador (version 2.0) viewer. The technical infrastructure for AOR includes a data archive, an image server, a IIIF image service, a corresponding IIIF presentation service, and the IIIF-compliant Mirador viewer.

The data archive for this project provides the framework for long-term access to, and preservation of, all project content: an important contribution to Digital Humanities more generally, insofar as this issue has not yet been unaddressed within IIIF. The technology team has defined an archival data model that can be mapped to other data models for data access and presentation over time. Another layer of the infrastructure that has been developed in Phase 1 of AOR consists of an image server that accesses content from within the data archive. Currently, the team is utilizing a commercial FSI image server, though comparable image server resources (e.g. djatoka) may also be used.

While AOR has been deployed using the versatile Mirador 2.0 API, the AOR technology team has developed IIIF endpoints for any available image service and presentation service, so long as they can be accessed by an IIIF-compliant image viewer. All AOR data can be accessed through our Mirador2 IIIF-complaint viewer, which has been specifically enhanced to meet the use requirements identified by the team for both current and future users. Over the course of AOR Phase 1, the technology and scholarly teams have worked together closely to define and implement a set of use cases related to image viewing and manipulation, transcription viewing, and dynamic, query-building search capabilities.

The current AOR viewer is the culmination of this iterative development process, which presents users with a wide-ranging set of functionalities aimed at enabling new forms of research on the history of reading practices. One new method is informed by the use of insights gleaned from the data generated throughout Phase 1 to formulate new research questions for the humanities team to investigate.

By creating a corpus of important and representative annotated texts with searchable transcriptions and translations, we can begin to compare and fully analyze early modern reading, and place that mass of research material within a broader historical context. In so doing, we could also approach—not in isolation but as a dynamic, internally and institutionally complimentary, research team—the traditionally subjective study of reading in a demonstrably empirical, comparative, and systematic way.

All visible interventions by readers in the books – marginal notes, underlining, marks and symbols, et cetera – have been marked up according to a non-TEI XML schema developed by humanities and technology teams working in close cooperation. The Phase 1 corpus of thirteen books owned and annotated by Gabriel Harvey has generated 3,355 XML files, each representing a single page of an annotated book. The data contained within these XML files has allowed us to map the language Harvey used in his own annotations and highlighted in the original printed texts, to find interesting correlations, and to use those correlations to form new investigations into the history of reading, Harvey's own biography, and the wider history of ideas in the early modern period.

Derrida's Margins: Annotations from the Personal Library of Jacques Derrida

Katie Chenoweth

This paper will give an overview of Phase One of the Derrida's Margins project currently underway at the Center for Digital Humanities at Princeton University (CDH).

 Derrida's Margins is a longterm project that aims to create a website and online research tool for annotations from the Library of Jacques Derrida, an archival collection housed at Princeton University Library's Rare Books and Special Collections that was first opened to the public in March 2016. Phase One of Derrida's Margins focuses on annotations related to Derrida's landmark 1967 work De la grammatologie, first translated into English in 1976 as Of Grammatology (hereafter OG, see Spivak's 2016 translation. This corpus will serve as a pilot data set for future work, allowing us to establish protocols, workflow, and a relational database model.

Jacques Derrida is one of the major figures of twentieth-century thought, and his personal library represents a major intellectual archive. The Derrida Library, consisting of about 19,000 published books and other materials, represents a lifetime of reading. But

for Derrida, the act of reading was not a passive process: he engaged — even grappled — with what he read, covering pages with notes and cross-references, inserting other handwritten materials, quoting and adapting what he read into what he wrote. As Derrida himself said in an interview later in his life, his books bear "traces of the violence of pencil strokes, exclamation points, arrows and underlining" (for details related to Princeton's acquisition of the Derrida Library, see the article posted by Princeton University Libraries, 2015).

It was in OG that Derrida first articulated a new style of critical reading, which would become the foundation of the philosophy of "deconstruction." Our online research tool will enable scholars to study the development of this philosophy in an unprecedented way by providing comprehensive digital access to the material annotations, marginalia, bookmarks, and other notes from Derrida's library that correspond to each quotation and citation in OG. Beyond making Derrida's annotations available digitally for the first time, this project seeks to enable researchers to understand the relationship between Derrida's published writing and his reading practices. As we move beyond OG in future phases, the website will also allow researchers to gain new insights into Derrida's library and his published writing as networks of texts, citations, and annotations.

OG was chosen as the pilot text not only because it is a foundational text for deconstructive reading, but for two additional reasons: 1) it is among the most widely read of Derrida's works and, according to Google Scholar, by far the most frequently cited, constituting more than 10% of the total citations for Derrida as of November 1, 2016; 2) 2017 will mark the fiftieth anniversary of the initial publication OG, and a number of scholarly conferences and events are scheduled to discuss the text and its legacy.

In Phase One, we began by identifying all instances of citation (quotations, references, footnotes, etc.) in OG that lead us back to books and other reading materials from Derrida's library. Each cited work from OG is entered into a Zotero library, which will eventually form the basis of a public bibliography for OG. We then located these references in Derrida's copy (or, often, multiple copies—different editions, translations, etc.) of each work. Each instance of citation is given a tag in the Zotero library that indicates the source page from OG, the type of citation, the page number(s) in the cited text, whether or not the volume is present in Derrida's Library at Princeton, and the presence or absence of

annotation (here construed as a mark of any kind, verbal or non-verbal). Next, we will transcribe all marginal annotations and other markings, limiting ourselves to tagged pages, i.e., those that are explicitly referenced by Derrida in OG. Digital images and transcriptions of these annotated pages will form the basis of the website, allowing users to go "behind the scenes" of Derrida's reading practices.

Given that Derrida's work is widely read in the United States, France, and around the world by scholars from numerous disciplines in the humanities and social sciences, we anticipate that the audience for this project will be broad and international. The more targeted audience would be specialists of Derrida and deconstruction, as well as researchers in philosophy, literary studies, and intellectual history.

One major risk this project faces is the question of French copyright restrictions. Due to the fact that the majority of books owned by Jacques Derrida are copyrighted works whose copyright is still in force, there is a question regarding the legality of posting digital images of sections of those books on our website. Despite the fact that the copies of these works in JD's library have been annotated by the philosopher's hand, the works themselves are the intellectual property of their respective publishing houses in France. Pending clarification by the Office of the University Counsel regarding the amount of risk we are at liberty to take in this respect, we will make a decision as to the amount of text from those works that it is acceptable for us to display on our public website.

This project is intended as the first phase in a long-term project. Team members in Phase One will create a manual detailing their work process and workflow to facilitate and standardize the project going forward. At the end of our pilot phase with the CDH, we will have the following

The main outcomes and deliverables of this project are the following:

- Bibliography of *De la grammatologie,* made available as public Zotero library
- Digitization of all relevant pages from the Derrida library
- Customized grayed-out images, if necessary
- Annotations transcribed, translated, and tagged
- Custom designed relational database to record and analyze Derrida's annotation practices and the anchor text they reference

- Bi-lingual (French and English) Web portal created in Django to allow users to search and browse the annotations
- Transcriber's manual

The Winthrop Family on the Page

Anthony Grafton, Jennifer Rampling, Christian Flow

"The Winthrop Family on the Page" will employ an extensive database of bibliographic information, biographical data, and marginal annotations to allow digital exploration of early modern lives and learned practices. Centering on the surviving library holdings of the storied Winthrop family, whose representatives included such prominent figures as John Winthrop, a founder and governor of the Massachusetts Bay colony, the final product will be a web platform affording a dynamic sense of how colonial readers interacted with texts and fellow readers. Users will be encouraged to follow the story along multiple axes, both diachronically—as family members communicated across decades, even centuries, in the margins of their shared books—and synchronically, as single readers followed references from one text to the other, leaving a bread-crumb trail of notation along the way. The experience will be further enriched by inclusion of information on how the Winthrops' books surface in other historical sources they left us, including journals and correspondence, ensuring once again vivid access not just to the texts, but to the people who read them.

The bulk of the project's database will consist in entries for some 300 books formerly owned by the Winthrops and currently held at the New York Society Library. Aside from bibliographical information for all of these books, the database will include a second module treating the marginalia they reveal. Here, the project will log the location of all manuscript notes in the collection; in addition, for a select, smaller subset of 50-60 books, high-quality digital images will allow users to view the marginal material directly. Several representative notes from the same sub-set will be translated and transcribed along with relevant anchor text, and all of the collected data will be encoded for search. Users will be able to come to the collection with queries about everything from annotations of a particular era, to those of a particular person, to those presenting a specific sign (e.g., the manicule) or handling a specific theme.

The defining characteristic of the project is its deft exhibition of a very particular source-base. The rich

surviving Winthrop holdings, coupled with the opportunities for collection and display afforded by a digital platform, offer a rare opportunity to excavate and re-assemble an Early Modern book collection, giving a concrete frame to the "intellectual space" within which the family lived and thought. Still more exciting is the chance to fill that space with dialogue: the project's careful curation, transcription, and (when necessary) translation of the marginalia that the Winthrops recorded as they read allows users to follow their intellectual journey between books and between generations. Those generations spanned not just time but space: because the Winthrops were a colonial family, their New England library had its roots and its first readers in Europe. Users of the collection are therefore positioned to pose and answer questions about how Early Modern knowledge made the transition from the Old to the New World. How, we wonder, might our understanding of the Salem Witch Trials be affected when we take into account the fact that Wait Still Winthrop, the Chief Magistrate of Massachusetts, could well have perused the marginal annotations of his European ancestor, Adam Winthrop, in the witchcraft section of the family library? How did reading and annotation practices originally cultivated in England filter into the learned arsenal of later colonial readers?

Queries like these are made actionable by the labors of a highly-skilled team: specialties in book history and annotation, in ecclesiastical history, in alchemical practice, and in the classical tradition will be amply represented. The project itself, which will offer direct access to an array of understudied material, is meant to have wide interdisciplinary appeal. Situated at the nexus of digital humanities and history of the book, straddling Europe and the colonies, the domains of intellectual history and learned practice, it will resonate with historians of Europe, America, and the Atlantic World; book historians; librarians; historians of science, medicine, and religion; and users outside the Academy. And in future years the appeal will only grow more broad-based, for project members hope to build outward from the corpus on which the 2016-17 work is founded. The more material at hand, the richer the web of connections users will be able to make between books and readers. There are several possible modes of expansion: (1) full imaging, transcription and translation of all books in the NYSL corpus (that is, including those beyond the subset currently targeted for the pilot-phase of the project); (2) inclusion of further Winthrop books housed at other major repositories, including items from the eighteenth and nineteenth centuries (when the Winthrops added hundreds of new books to the library, including a collection of almanacs now housed at the Houghton Library, and the bulk of the collection of Winthrop books at Allegheny College); (3) development into a broader study of Colonial reading practices and knowledge networks, involving other family collections of annotated books, such as the Mather Family Library, now held at the American Antiquarian Society; the library of Thomas Prince, now at the Boston Public Library; and the library of James Logan, now at the Library Company of Philadelphia.

Notes

1. There are a number of projects dedicated to recording annotations or marginalia, notably the crowd sourced Annotated Books Online and Book Traces. The three projects discussed in this panel proposal differ in that they are specifically grouped by annotator and want to answer specific questions about the worldview and read patterns of known annotators. They are also designed to facilitate more complex queries. These projects are also addressing a different question from other annotation projects such as Annotation Studio which are designed for modern readers to annotate digital texts.

Bibliography

Derrida, J. (1967), *De la grammatologie* (Paris: Editions du Seuil).

Derrida, J., (2016) *Of Grammatology*, trans. Gayatari Chakravorty Spivak (Baltimore: Johns Hopkins University Press).

Jardine, L., and Grafton, A. (1990)"'Studied for Action': How Gabriel Harvey Read His Livy," *Past & Present*, no. 129: 30–78.

Princeton University Libraries (2015). Princeton University Library Acquires Jacques Derrida's Personal Library. 31 March, http://library.princeton.edu/news/2015-03-31/princeton-university-library-acquires-jacques-derridas-personal-library