

MATERIAL AWARENESS

*Exploring the entanglement of
library digitization and
digital textual scholarship*

Merisa Martinez

18 November 2025

Lund University

Brief intro to my academic background

- BA (Hons) in Modern European History with a secondary focus on French Literature
- MSc in Modern British and Irish History, focusing on 20th century warfare, philosophy and mathematics
- PhD in Library and Information Science, focusing on library digitization and digital textual scholarship
- Currently: Temporary Senior Lecturer in Digital Humanities at Uppsala University

The Digital Scholarly Editions Innovative Training Network (DiXiT ITN)

- A Marie Skłodowska-Curie Action including intensive training courses and conferences
- 12 Early Stage Research Fellowships and 5 Experienced Research Fellowships
- 10 Partner Institutions:

Högskolan i Borås (SE); University of Antwerp (BE); University of Cologne (DE); École des Haute Études en Sciences Sociales (FR); University of Graz (AU); King's College London (UK); Oxford University (UK); Università di Roma 'La Sapienza' (IT); National University of Ireland, Maynooth (IE); Huygens Institute (NL)



You can find information about MSCA PhD positions through the EURAXESS portal

- MSCA PhD Fellows are required to be mobile: you cannot perform your PhD in the country where you currently live – so you would have to move outside of Sweden
- They have generous salaries and travel stipends with good placements for research secondments
- An excellent way to make international connections and life-long academic partnerships

EURAXESS Portal: <https://euraxess.ec.europa.eu>



My PhD Project in the DiXiT ITN

My PhD Project: Critical transmission activities in memory institutions

- Placed at Högskolan i Borås, working with two PhD supervisors and a Postdoc in the same project
- Aimed to track and map how critical decision making was occurring in the digitization process in libraries and relate this to digital textual scholarship
- Performed a preparatory study at Cambridge University Library and three published studies branching off from that preparatory study

Research Problem

Library digitization has advanced significantly in the past 30 years, bringing new partners, new methods, new workflows. The same has happened in digital textual scholarship, yet there has yet been no major study exploring how the people involved in these activities collaborate, share knowledge, and how their concerns overlap.

Research Questions

1. How and where does advanced library digitization work performed by library staff involve **critical decision making** that contributes to the scholarly landscape?
2. How and where does such work **affect, overlap, and interplay** with the work of other actors such as textual scholars and their editing/publication projects?

Research Methods

- Participant Observation

Working closely with experts for a prolonged period, and learning their skills along the way

- Semi-Structured Interviews

Asking pre-determined questions while still allowing room for unexpected insights to develop

- Surveys

Using insights gained from the two methods above to spot and analyze trends in more detail

What is digital textual scholarship?

What is digital textual scholarship?

- Digital Textual Scholars use computational tools to 1) study how what we think of as 'finished' or 'established' texts came to be, and 2) to establish texts based on scholarly editing
- Look at writers' documents, including draft materials, correspondence, marginalia in their books, as well as comparing multiple published versions of their work
- Can include skills such as paleography, textual criticism, codicology, bibliography, typography, collating textual differences across manuscripts/printed texts
- Requires the ability to think abstractly about the creative process and historical factors while building a concrete, disciplined understanding of how writers organize their thoughts and ideas

Example: The Anne Frank Manuscripts

AnneFrank MANUSCRIPTS

Manuscripts Guide Textual relationships About this edition 🔍 📖 NL
EN

**Scholarly edition of Anne Frank's diary, of her
adaptation of it to 'a novel about the Back House' and of other texts
with which she hoped to become a famous writer.**



Example: The Anne Frank Manuscripts

Manuscripts

By scrutinizing the original manuscripts, Anne's development as a writer and as a person can be followed. Plus, the way in which she rewrote her diary is also made clear. Explanations and illustrations clarify the historical context of the events and circumstances she describes.

Diary

The red and white checkered diary and two other notebooks with the original version of Anne's diary.

The Back House

Anne's rewritten version of her diary, which she hoped to publish as 'a novel about the Back House'.

Tales

Short stories and sketches, written down in her notebook Tales, and Occurrences from the Back House.

Cady's Life

Unfinished, fictional story, of which five chapters and some fragments for the continuation of the story survive.

Book of Beautiful Sentences

Anne's anthology of quotes and fragments that inspired her.

Example: The Anne Frank Manuscripts

AnneFrank MANUSCRIPTS

ManuscriptsGuideTextual relationshipsAbout this edition

NL
EN

Textual relationships

Anne rewrote parts of her diary, which she hoped to publish as 'a novel about the Back House'. She also incorporated certain occurrences and passages into other texts, for example into her tales. This overview provides a schematic representation of all the relationships that exist between the texts in this edition.

	 Diary	 The Back House	 Tales	 Cady's Life	 Book of Beautiful Sentences
June 1942	fri 12 June sun 14 June mon 15 June tue 16 June fri 19 June	sat 20 June sat 20 June sat 20 June sun 21 June			

Understanding the development of Anne Frank's creative writing process and how her manuscripts are connected to and influencing each other through a process of **collation** and **genetic criticism**

Example: Mary Shelley - Frankenstein

Bodleian MS Abinger c.56

Author(s) : Mary Shelley

Date Written : [August or September]-[?December] 1816

Title/Literary Work : Frankenstein

State : draft

Institution : Bodleian Library,
University of Oxford

Hand(s) : Mary Shelley, Percy Shelley

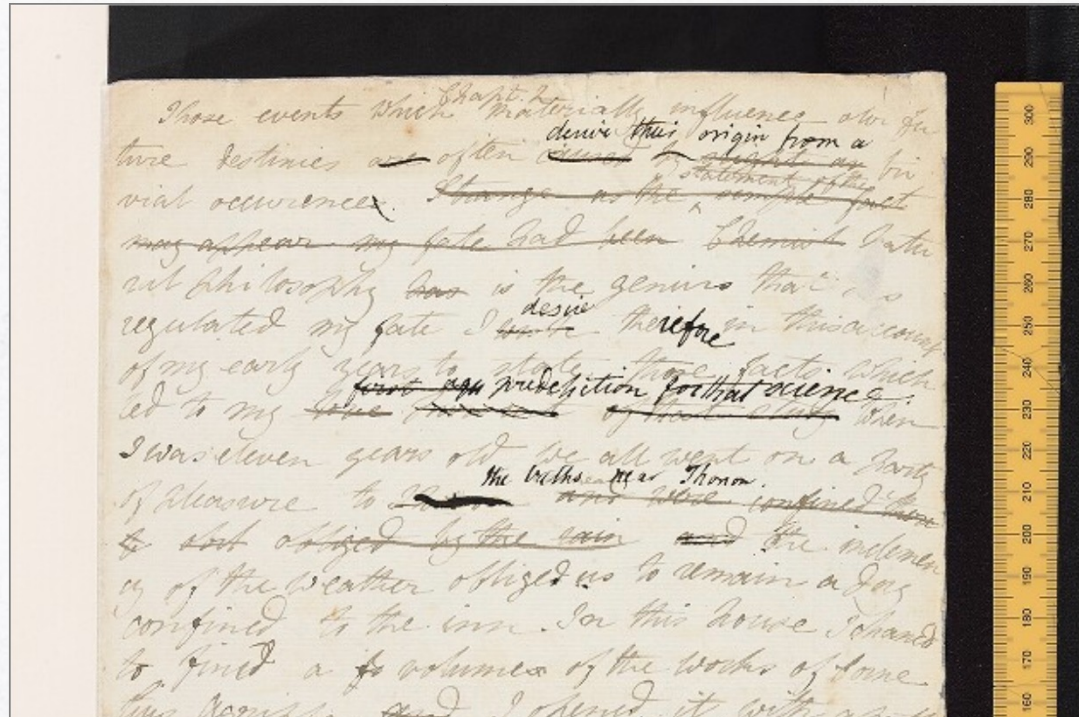
Shelfmark : MS. Abinger c. 56

Folio : 1r

Transcription Status: ● ● ●
Metadata Status: ● ● ●

Q Q Q ↶ ↷ LIMIT VIEW: ☒ All ☐ Mary Shelley ☐ Percy Shelley

⏮ ⏪ ⏩ ⏭



1

Chapt. 2

Those events which materially influence our future destinies are often caused by slight or trivial occurrences. Strange as the simple fact may appear my fate had been Chemist Natural philosophy has is the genius that has regulated my fate I wish therefore in this account of my early years to state those facts which led to my love pursuit of that study. When I was eleven years old we all went on a party of pleasure to Thonon and were confined there by the rain and the inclemency of the weather obliged us to remain a day confined to the inn. In this house I found a few volumes of the works of some time ago and I devoted it with some

An example of a diplomatic transcription of Mary Shelley's draft manuscript for Frankenstein

Tensions in Digital Textual Scholarship

Critical

Interventionist

Interpretive

Facts as interpretation

Idiographic (individual)

Contingent tools

Material document

The one text: being
selective

Non-Critical

Non-interventionist

Factual

Facts as separate from interpretation

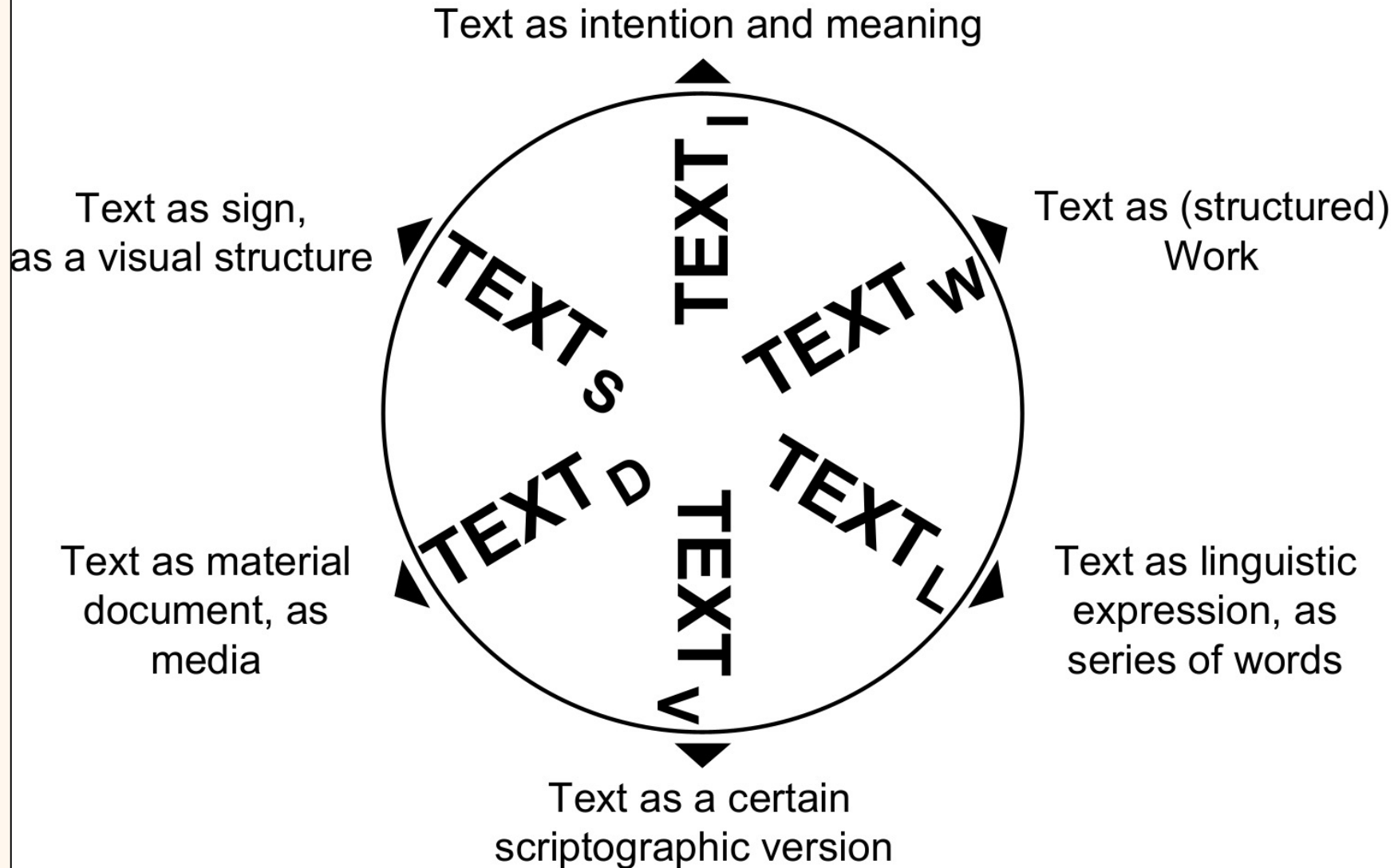
Nomothetic (general)

Universal tools

Abstract text

The many texts: being
comprehensive

Notions of text



Patrick Sahle,
2006, – The
Wheel of Text

In textual
scholarship,
‘text’ can
have many
meanings

What is library digitization?

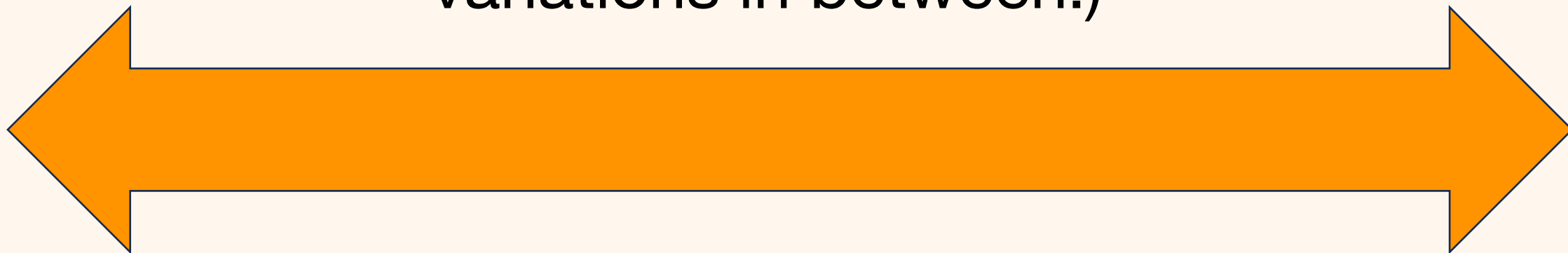
A (very) brief timeline of library digitization

- 1920s-1930s: Library of Congress begins to microfilm collections from the British Library for consultation by researchers in the U.S.
- 1940s: Library staff across the U.S. set about creating networked catalogues of information held across libraries
- 1950s: first electronic catalogues with bibliographic information appear
- 1960s-1970s: reproduction photography and electronic cataloguing are continually refined and standards are implemented in libraries

- 1980s: forays into creation of digital files from printed sources
- 1990s: electronic library catalogues on CD-ROMs can be purchased for home use
- 1990s: 'the decade of digitization' for libraries, focus on unique or canonical text-based sources
- 2000s: Google Books and mass digitization
- 2010s-2020s: Theorizing critical/slow digitization and development of interactive digital models of physical materials viewed through graphical user-interfaces

What types of library digitization are there?

Fast, commercial mass digitization on the one end of the spectrum, and boutique, slow single object digitization on the other (with many variations in between!)

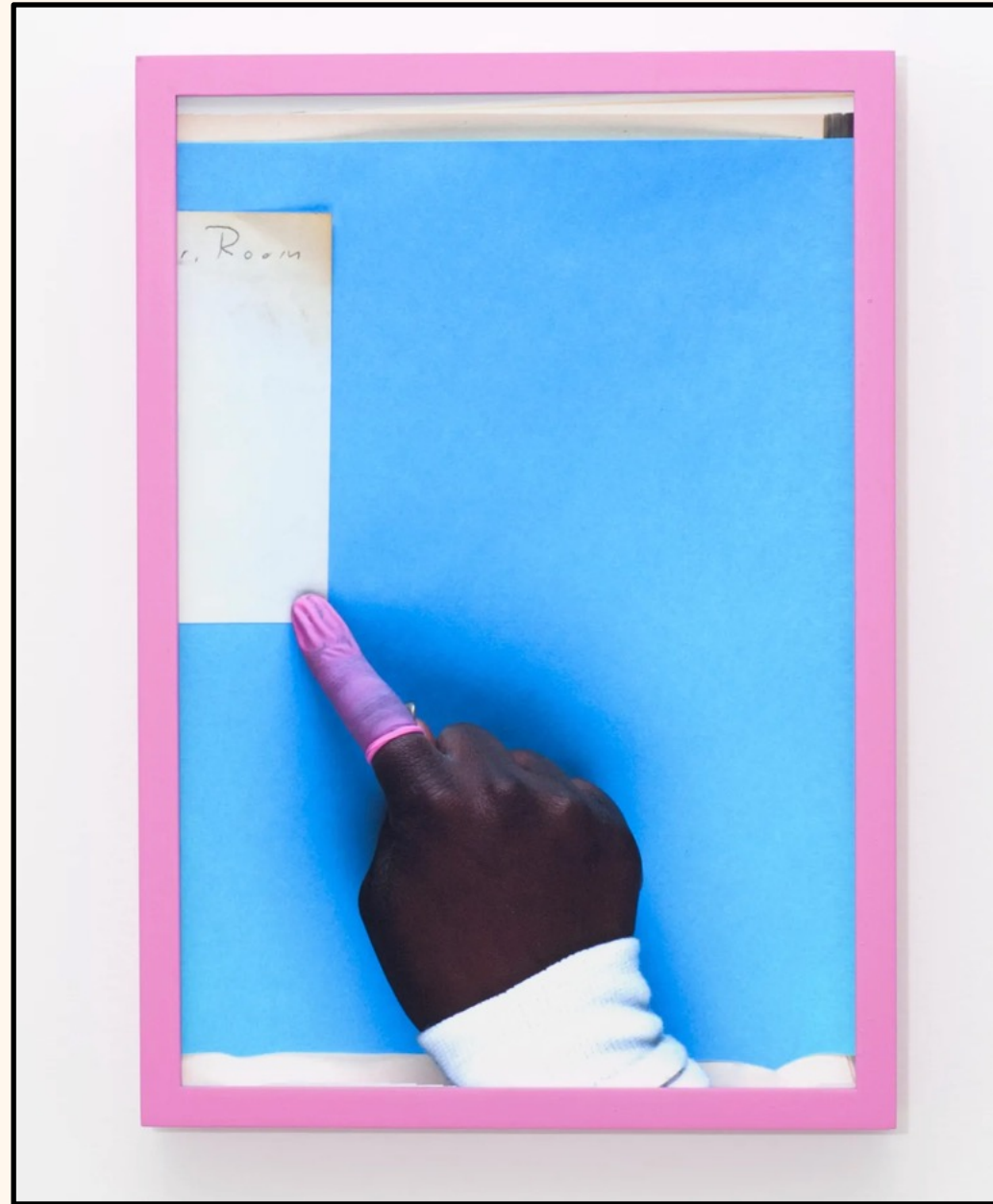


However!

I found during my PhD research that these types of vocabulary (mass vs. critical) need to be reassessed given the variations of activity across different digitization labs as well as the rapid development of equipment and standards in the field of cultural heritage digitization



What
types of
library
digitization
are there?



Mass
digitization
(Google
Books)

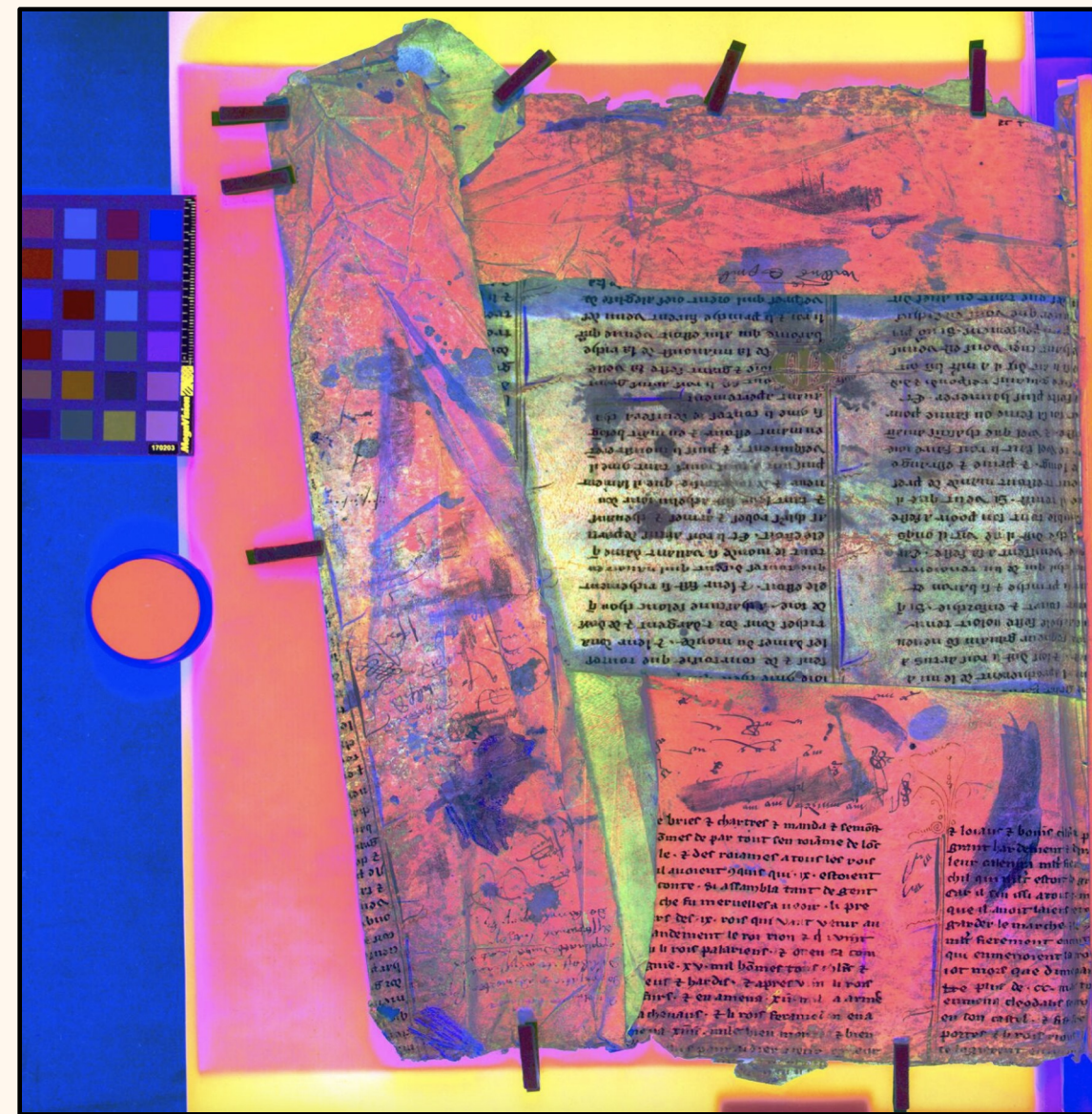
Invisible
labor, often
female PoC

Critical digitization (of cultural heritage materials)

Performed by highly
trained photographers,
and can also often be
characterized as invisible
labor (Martinez, 2024;
Smith & Whearty, 2023)



3D scanning of the *Merlin* manuscript in Cambridge UL, Cultural Heritage Imaging Laboratory.



Multispectral imaging of the Merlin Manuscript



7 of 8





Zoom: 18 %

Home | Browse | CUL.1,52,155

About | Contents | Thumbnails | Transcription | Translation | More

Chinese Oracle Bones (CUL.1,52,155)

Over three thousand years old, the Chinese Inscribed Oracle Bones in the [Hopkins Collection](#) are by far the oldest written materials in Cambridge University Library.

The oracle bone texts are the oldest extant documents written in the Chinese language. Inscribed on ox shoulder-blades and the flat under-part of turtle shells, they record questions to which answers were sought by divination at the court of the royal house of Shang 商, which ruled central China between the 16th and 11th centuries B.C.E.

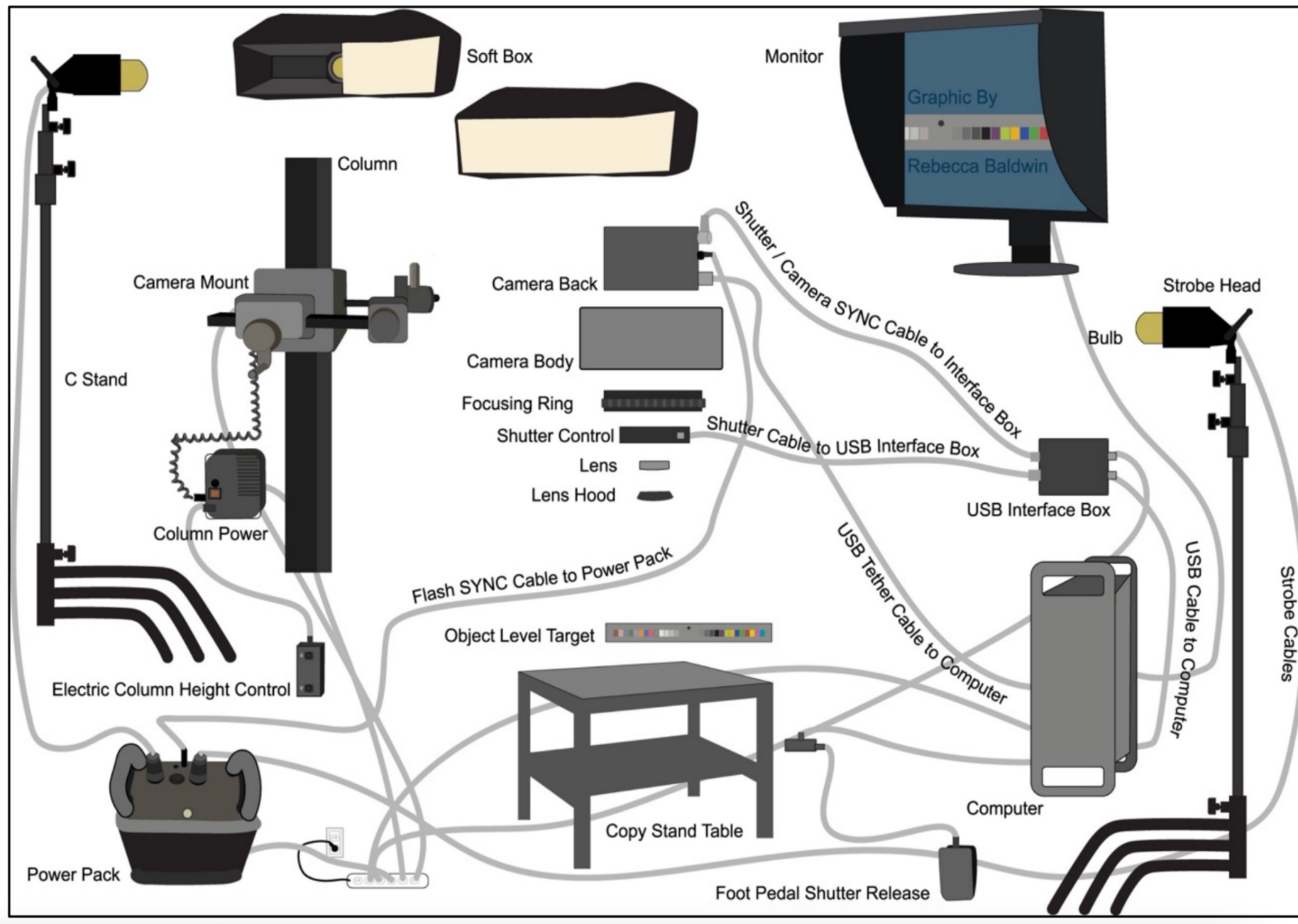


[Cambridge UL Oracle Bone CUL.52 Hi Res on Sketchfab](#)

[Download Image](#) [Request Rights](#) [Download Metadata](#) [IIIF Manifest](#) [Open in Mirador](#)

Page: 7

The Chinese Oracle Bones [\(link\)](#)



Rebecca Baldwin et al., 2021:

(Some) equipment needed to perform cultural heritage digitization

Tensions in Library Digitization

Critical (or Slow) Digitization

Primarily manual
Prioritizes depth
Maximizes editing/ metadata
Interpretation
Critical selection
Focus on the document
Qualitative
Idiographic (specific)

Mass Digitization

Primarily automated
Prioritizes breadth
Minimizes editing/metadata
Clone
Random selection
Focus on the text
Quantitative
Nomothetic (general)

When is mass digitization not recommended?

- For older, unique, and very fragile materials that do not respond well to light, heat, humidity
- For manuscripts with special material aspects (such as palimpsests)
- For books that do not conform to standard sizes
- For bound books that cannot lie flat/open widely enough to be placed on a scanner
- For materials with unique aspects such as fold-out maps or other material oddities
- For those whose scripts will not be captured by current OCR scanning

What happens in a library digitization project?

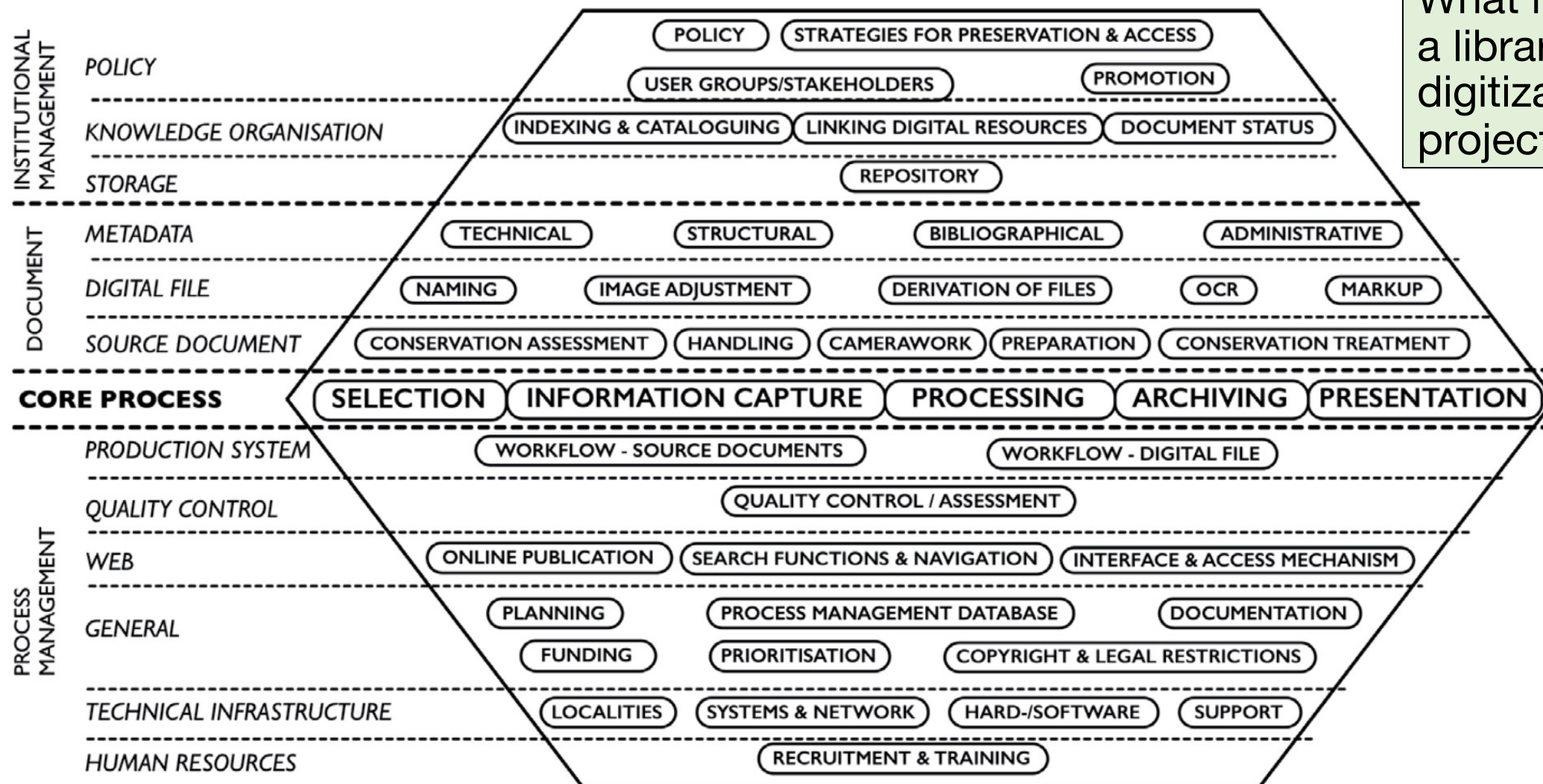


Figure 29. A conceptual model of the digitisation process outlining the stages in the core process and three sets of strata: *Process management*, *Institutional management*, and *Document*.

Björk, 2015, p. 145

Studies and Results from my PhD

My PhD: Visiting Libraries

National Library of Sweden

- What the digitization workflow looks like, how it is incorporated into every aspect of the organization

National Library of Denmark

- Digitization as a teaching opportunity for students interested in history and textual scholarship; Digitization as a form of preservation of unstable materials (nitrate film)

Uppsala University Library

- How multi-institutional digitization collaboration works, how research infrastructure for digitization is built and implemented

Cambridge University Library

- Participating in the digitization process first hand, observing as photographers and conservationists perform critical and interpretive work

Study 1 (Martinez & Terras, 2019)

Orphan Works

Studied how restrictive copyright frameworks for orphan works – those works for which a copyright holder is not known or cannot be located – is hampering the digitization, dissemination and reuse of cultural heritage materials

Found that current frameworks are either so confusing or restrictive that cultural heritage institutions are avoiding the digitization of these (sometimes extremely fragile, high academic value) materials – this is affecting both digitizers and textual scholars

Study 2 (Martinez et al., 2019) Access and Accessibility Survey

Found that textual scholars and library staff have different meanings for the words access and accessibility.

E.g., **access** can mean:

access to the transcriptions of a digital project, access to the images, access to metadata, access to project notes, access to the code that makes the project run, access to the tools that the project partners used. And what level of access? Is reuse of these aspects allowed? Opinions differed greatly!

Study 2 (Martinez et al., 2019) Access and Accessibility Survey

Accessibility can mean:

making the edition/digital collection more easily understandable to non-experts in the subject, AND that it can refer to web accessibility (providing ways to interact with projects for people with (in)visible disabilities). Textual scholars were more skeptical than library staff that web accessibility was necessary.

These differences contribute to **misunderstandings and missed opportunities** for deeper collaboration in projects, as well as interacting with a wider user base!

Study 3 (Martinez, 2019) Digitization vs. Digital Textual Scholarship in Swedish Libraries

Met with and interviewed photographers and researchers working at Uppsala University Library, Gothenburg University Library, and Litteraturbanken.

Realized that many scholars with PhDs and extensive experience working with textual scholarship and textual criticism are in the library, and are directly influencing textual projects – and that photographers are also influencing these projects by making suggestions for how to organize the materials digitally

Results – The Digitization Process

“Why can’t we just digitize everything?”

Because library digitization is heavily influenced by local, institutional contingencies including access to staffing (for e.g., photography, licensing, metadata, web development...) equipment; skills’ expertise; support from administrators; grant funding that can often shift and change over time, etc., etc., etc....

Results – Finding Common Ground

- **Access and accessibility** – translating across communities
- **Copyright and licensing** – textual scholars and library staff are both subject to this, though textual scholars more likely to find ways around it
- **Infrastructure and tools** – textual scholars and library staff rely on the same set of objects to conduct their work
- **Service Orientation** and its effect on scholarly communication – library staff are less likely to publish regularly and to promote individual methods and work, and therefore **their labor often becomes invisible** unless they have long-term partnerships with academics or strong support from administrators

Discussion – Material Awareness

While digital textual scholars have a strong focus on **textual awareness** (through the creation of digital text collections), and library digitizers are preoccupied with **image awareness** (through the creation of digital facsimiles), I argued that both have a shared, though unbalanced, focus on ***Material Awareness***.

This material awareness is characterized by **the knowledge of *what was*** (documents that no longer exist and therefore cannot be consulted), ***what is*** (the current makeup of both documents and the materials needed to digitize them), and ***what will be*** (the necessity to constantly update equipment and plan for the digitization of new and materially complex documents). The service orientation of libraries means that the responsibility lies largely with the library staff to plan for, accommodate, and preserve both physical and digital documents for the long term.

So what is needed now?

More focus on **paradata**: capturing aspects of the research process in digitization studios in photos, video, audio and documenting how decisions are made, what vocabularies are agreed on!

Better understanding of each other's vocabularies (e.g., 'access' and 'accessibility' for what, exactly? The underlying code of a project? The images? The transcriptions? The metadata schemas? Documentation about the tools? Documentation about research practices and decision making?)

Historico-Mathematical Case Study: Alexandre Grothendieck



Alexandre Grothendieck, pictured on the left, with Alain Schwartz on the right.

At the end of my thesis, I wrote about my previous study of Alexandre Grothendieck, who was a proponent of building bridges across different areas of mathematics with a common, shared vocabulary. To break out of silos and perform better, more enriched work. I used this as an example of someone who, by taking a bird's eye view of lots of different areas, could find valuable connections.

Historico-Mathematical Case Study: Alexandre Grothendieck



Alexandre Grothendieck, pictured on the left, with Alain Schwartz on the right.

Alexandre Grothendieck is widely considered the most consequential mathematician of the 20th century

When he died, he left behind 70,000 pages of manuscripts to be digitized
But should they be?

Historico-Mathematical Case Study: Alexandre Grothendieck



This is disputed!

Because he disavowed his
work

To work with these
materials one needs skills
in textual criticism,
paleography, archival
science, and mathematics

Bringing my thesis into the classroom

1/2

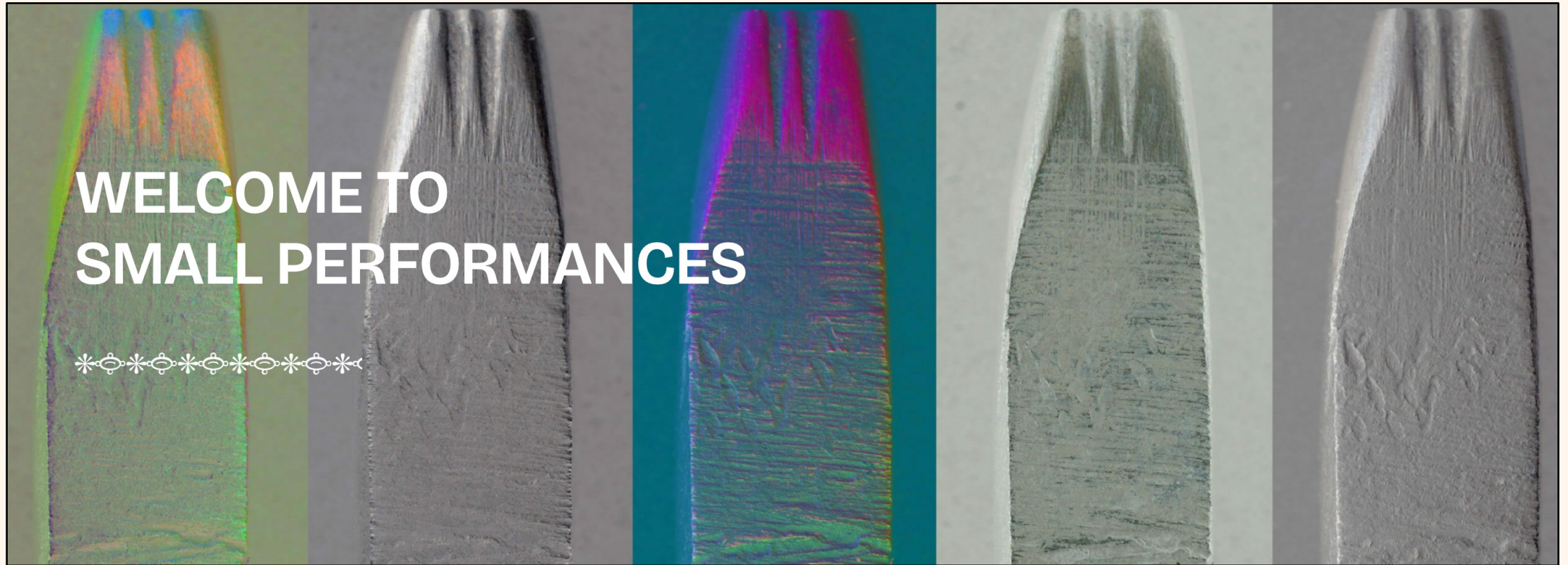
- Digital Cultural Heritage – Digital Humanities Master's Thesis Course
 - First Year Students
 - Focus on: Invisible Labor in the digitization process
 - Photography as highly trained, intuitive, collaborative act of scholarship framed by local contingencies and access to support
 - The ever-present question “Why don't we just digitize everything?” demonstrates the need for continued research and engagement in this area
 - Includes field trips to cultural heritage digitization labs in order to engage students in the intricacies of the digitization process

Bringing my thesis into the classroom

2/2

- Digital Text and Image Analysis– Digital Humanities Master's Thesis Course – First Year Students
- New course; previously Distant Reading and Visual Analysis; now a combination of those two with a broader focus on textual criticism, textual scholarship, and digital literary studies beyond DR
 - Discuss how texts are 'established,' explore how understanding this can affect our analysis of large-scale corpora of texts

Small Performances: The John Baskerville Punches Project (AHRC)



Performing RTI imaging on John Baskerville's typographic punches to understand how they were made (as no documentation still exists explaining their creation)

Small Performances: The John Baskerville Punches Project (AHRC)

Taking my insights from my thesis into my current research project Small Performances: the Baskerville Punches Project (funded by the Arts and Humanities Research Council in the UK until 2027)

Working with cultural heritage photographers, bibliographers, archaeologists, metallurgists, blacksmiths, jewelry makers to understand the material aspects of the punches and how they were made

We have a group vocabulary and hold regular meetings to train each other in our respective disciplinary backgrounds!

Thank You!

Any questions? Please email me: Merisa.martinez@abm.uu.se 😊

Textual Sources

- Björk, L. (2015). *How reproductive is a reproduction? Digital transmission of text-based documents*. Doctoral Dissertation: University of Borås, SE. Valfrid no. 59.
- Braid, A. (1994). The Problem of Digitising Non-electronic Media. *LIBRI: International Journal of Library and Information Studies*, 44(4), pp. 311-316.
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnapp, J. (2012). *Digital Humanities*. MIT Press.
- Dahlström, M. (2010). Critical editing and critical digitization. In van Peursen, W., Thoutenhoofd, E., and van der Weel, A. (eds.) *Text Comparison and Digital Creativity: The Production of Presence and Meaning in Digital Text Scholarship*. Amsterdam: Brill, pp. 77-97.
- Dahlström, M. (2011). Editing Libraries. *Bibliothek und Wissenschaft* 44, pp. 91-106.
- Martinez, M. (2024). *Material Awareness: Exploring the entanglement of library digitization and digital textual scholarship*. Doctoral Dissertation: University of Borås, SE. Valfrid no. 80.
- Prescott, A. and Hughes, L. (2018, September). Why do we digitize? The case for slow digitization. *Archive Journal*, 2018.
- Smith, A. and Whearty, B. (2023). All the Work You Do Not See: Labor, Digitizers, and the Foundations of Digital Humanities. In Gold, M.K. and Klein, L.F. (eds.) *Debates in the Digital Humanities 2023*. Minneapolis: University of Minnesota Press.
<https://dhdebates.gc.cuny.edu/read/debates-in-the-digital-humanities-2023/section/f1946d2b-93ca-4ea0-ab8c-54fb835750ad>
- Terras, M. (2015). Cultural heritage information: Artefacts and digitization. In G.G. Chowdhury & I. Ruthven (eds.) *Cultural heritage information: Access and management*. London, UK: Facet, pp. 63-88.
- Thylstrup, N.B. (2018). *The Politics of Mass Digitization*. The MIT Press.
- Van Horik, R. (2005). *Permanent Pixels: Building blocks for the longevity of digital surrogates of historical photographs*. Doctoral Dissertation: Technische Universiteit Delft.
- Vickery, B.C. (2000). *Scientific Communication in History*. Scarecrow Press.

Interactive Sources

The Shelley Godwin Archive: <https://shelleygodwinarchive.org>

The Anne Frank Manuscripts (to access the manuscripts, you must first indicate that you have the right to access them based on your current location in Europe – this is due to copyright restrictions. If you have any trouble using the copyright declaration page in English, try using the page in Dutch, and that should work).

<https://www.annefrankmanuscripten.org/en/tekstverbanden>

Cambridge University Digital Library: <https://cudl.lib.cam.ac.uk>

Citing these slides

This presentation was given by Dr. Merisa Martinez as a guest lecture to students in the Digital History Master's course (administered by Pelle Snickars and Mathias Johansson) at Lund University on 18 November 2025.

Suggested citation:

Martinez, M. (2025, November 18). *Library digitization and digital textual scholarship*. [Lecture Slides] Lund University, Digital Historia.