

# LaTeX in the Digital Humanities

---

Sarah Lang  
2024



# Overview

1.  $\text{\LaTeX}$  and the Digital Humanities
2. Use Case 1: Conference Proceedings in  $\text{\LaTeX}$
3. Use Case 2: Printed Books & Archaeological Catalogues
4. Use Case 3: Print Versions for Digital Scholarly Editions
5. Challenges & Popularizing  $\text{\LaTeX}$  in DH

This talk explores the intersection of LaTeX typesetting and the digital humanities.

More specifically, it asks why LaTeX typesetting is used so infrequently in the digital humanities, despite its clear advantages and applications.

This talk, on the one hand, aims to understand the reasons why LaTeX isn't more widely used in the digital humanities and, on the other hand, presents three examples of relevant use cases to illustrate its value for the (Digital) Humanities.

# $\text{\LaTeX}$ and the Digital Humanities

---

# What are the DH? (& why don't they use $\LaTeX$ ?) i

A difficult question. No one knows and there are rumors that, in fact, the field doesn't want to define itself (Piotrowski 2019).

## Pragmatic Definition of Digital Humanities

Digital Humanities means using digital methods for humanities data and to answer humanities research questions in the field.

# What are the DH? (& why don't they use $\LaTeX$ ?) ii

Following Piotrowski 2019 & Roth 2019, according to Lang 2021a:

## 1. Humanities of the Digital

Media studies &c. (especially in the Anglophone area); not the primary focus of my flavor of DH.

## 2. Digitized Humanities

As the name suggests, this involves doing what humanities scholars would have done traditionally, but in the digital realm (modelling, annotation, maybe with bigger corpora than before).

## 3. Computational Humanities

Applying computational methods such as Computer Vision, large-scale text analysis (*distant reading*) or Machine Learning more broadly to Humanities data. This often proves surprisingly difficult despite the increasing availability of usually highly performative algorithms.

# Three Use Cases for $\text{\LaTeX}$ in Digital Humanities

## 1. Conference Submissions

Using LaTeX in conference submissions is very pragmatic and doesn't necessarily pertain to the research topics themselves, but rather in how research is communicated and published.

## 2. Archaeological Catalogues

LaTeX is useful for typesetting big scholarly monographs in the humanities. Even though some humanities disciplines that publish big monographs might not seem very digital, LaTeX proves valuable because:

- Monographs with lots of images don't work well in MS Word.
- Humanities scholars sometimes already have their catalogues formatted in CSV, which can be leveraged by LaTeX packages to easily create a catalogue (Lang 2021b)

## 3. Print Versions for Digital Scholarly Editions

Complicated. More on that later... (Lang 2022d; Lang 2022a)

## Use Case 1: Conference Proceedings in $\text{\LaTeX}$

---



Conference proceedings that require full paper submissions, exemplified by the *Computational Humanities Research* Conference present a good reason to get acquainted with the necessary  $\LaTeX$ .

## But there are some issues:

- Resorting to  $\LaTeX$  (specifically) has been criticized as exclusionary by some DH discourses (Dombrowski 2022).
- I agree that not all DH work involves coding but don't see how  $\LaTeX$  in particular has to take the blame for some things wrong with the field, especially given it's not even that hard to learn compared to other things that are becoming increasingly common in DH.
- I think this sentiment of feeling alienated by  $\LaTeX$  has less to do with the technology or the  $\LaTeX$  community (that most DHers probably don't know about, in my opinion) but rather the association with Computer Science as a problematic and exclusionary field.

## Pros according to me:

- Allows outsourcing typesetting work to users through LaTeX templates. (Understandable as the conference began as a small community effort.) Some more burden is placed on the editors / conference organizers.
- Humanities scholars, even those with minimal technical background, can learn LaTeX at the necessary level to format their submissions effectively, by using it essentially as a markup language.
- Humanities scholars often understand markup languages intuitively. Learning LaTeX is much easier compared to programming languages like Python, which are increasingly required in the DH.
- I even made tutorials (Lang 2019; Lang 2022c).
- Fast open-access publication of the proceedings (which is good for early career scholars and doesn't rely on exploitative practices in commercial academic publishing).

## Use Case 2: Printed Books & Archaeological Catalogues

---

# Physical Books in the Humanities

Despite early predictions that digital media would make physical books obsolete, we observe in 2024 that this transformation has not yet materialized and seems unlikely to occur in the near future.

- The value of the book as a material object and a symbol of cultural capital has surged, especially among younger audiences on platforms like BookTok.
- Enduring preference for printed materials in many Humanities fields and for many tasks such as close reading.

In fields like archaeology, it is commonplace to compile extensive catalogs of objects or findings as integral components of larger research projects, such as PhD dissertations.

- Such catalogs often contain numerous images and can become unwieldy when managed with standard word processing software like Microsoft Word due to their size.
- Although this application is not exclusively within the DH, it underscores the relevance of LaTeX across broader humanities disciplines, as has been discussed previously on the LaTeX Ninja blog (Lang 2021b).
- LaTeX is great at:
  - Effectively typesetting large monographs.
  - Managing extensive catalogs of objects or findings (e.g. `csvsimple`, Lang 2021b), including lots of images.

## Use Case 3: Print Versions for Digital Scholarly Editions

---

# Digital Scholarly Editing

The third and most important use case for  $\text{\LaTeX}$  in the DH is in digital scholarly editing (DSE). There is a sizable DSE community within the field. One could even say that DSE is one of the core fields of work for DH (at least from a Graz perspective where it's quite dominant) despite the subfield of Computational Humanities being on the rise.

## My prognosis

Digital scholarly editing will continue to be a core task and technology within Digital Humanities, thus,  $\text{\LaTeX}$  will continue to remain relevant to the Digital Humanities.

Maybe even more so at a time where we have to acknowledge that print has not died out and doesn't seem to any time soon...

# Transforming TEI-XML to $\LaTeX$

Transforming digital scholarly editions in TEI-XML format to  $\LaTeX$  allows us to produce print versions from digital editions, even when the primary intent of the edition is not to publish in print.

- **XSLT Transformation:** Through the transformation of TEI data using XSLT (potentially even employing large language models to generate  $\LaTeX$  code) scholars can create high-quality printed materials with little effort.
  - **Utility:** This functionality is crucial for the digital humanities community, as it responds to the frequent preference of more traditional Humanists for accessible, printable formats, which are particularly useful in teaching settings.
  - **Tools:** The `reLedmac` package is a particularly useful tool for Digital Humanists.
- Effective for ensuring accessibility and usability of scholarly resources in both digital and print formats.



## Wait...What's TEI?

When it comes to LaTeX and digital scholarly editing, I maybe first need to introduce the *Text Encoding Initiative* (TEI):

- **TEI:** The standard XML format used to archive data for digital scholarly editions.
- **Functionalities:** TEI supports different types of texts and critical apparatus for documenting variants in a stemma of different textual witnesses.

Some in the LaTeX community may already be familiar with TEI or at least the issue of typesetting critical apparatus, as there are multiple packages, such as `reledmac`, that deal with this.

## Text Encoding Initiative

### .XML

XML-Standard, i.e. convention on how to use XML so that resulting data will be interoperable between different projects. (founded in 1987, consortium since 2000)

“ The Text Encoding Initiative (TEI) is a text-centric community of practice in the academic field of digital humanities, operating continuously since the 1980s. The community currently runs a mailing list, meetings and conference series, and maintains the TEI technical standard, a journal, a wiki, a GitHub repository and a toolchain. (Wikipedia) ”

## TEI minimal example

```
<TEI> <!-- root element -->
  <teiHeader>
    <!-- author, title, dating,
           sources, edition rules, etc.
    </teiHeader>
    <text> ... </text>
</TEI>
```

## TEI Core

- **div** (division)
- **p** (paragraph)
- **head** (heading)
- **lb** (linebreak)
- **pb** (page break / beginning)
- **hi** (highlight)
- **bibl** (bibliographical information)

# Typesetting critical editions with the `reledmac` package

A critical edition/apparatus documents the variants of a text in different witnesses of textual transmission.

`reledmac` package (Wilson and Rouquette 2024) (reviewed in Dunning 2020) and critical apparatus example (from Overleaf tutorial document)

5 Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce sed dolor libero. Aenean rutrum vestibulum lacus ut pretium. Fusce et auctor lectus. Ut et commodo quam, quis gravida orci. Nullam at risus elementum, suscipit enim a, pellentesque mi. Morbi commodo, ligula vel consectetur accumsan, massa metus egestas velit, eu fringilla leo ante in turpis. Vivamus ut tellus sollicitudin, facilisis ipsum sit amet, tincidunt odio. Maecenas tincidunt dolor sed ante blandit tincidunt. Etiam vulputate ultricies facilisis.

---

1 Lorem ] A critical note  
1 ipsum ] An other critical note  
1-3 Fusce...mi ] A long critical note

---

1 Lorem ] Critical note in series B  
1 ipsum ] Other critical note in series B

1-3 Fusce...mi ] Again B

# Critical Apparatus i

...aims at documenting the variants of a text in the witnesses of textual transmission.

A critical apparatus is encoded with:  
<app>, <rdg> / <rdgGrp> and  
<lem>. <lem> can contain text not  
documented in any textual witness.

```
<app>
  <rdg wit="#Sh1">Then</rdg>
  <lem wit="#Sh2">Than</rdg>
</app> is my deede to my most
<app>
  <lem wit="#Sh1">painted</rdg>
  <rdg wit="#Sh2">pained</rdg>
</app>word:

<app>
  <lem>deed</lem>
  <rdg wit="#Sh1 #Sh2">deede</rdg>
</app>
```

```
<app>
  <rdgGrp type="orthographic">
    <rdg wit="#Sh1">giue</rdg>
    <rdg wit="#Sh2">give</rdg>
  </rdgGrp>
  <rdg wit="#AS1">have</rdg>
</app>

<listWit>
  <witness xml:id="Sh1">
    <bibl>Folger STC 22276</bibl>
  </witness>
  <witness xml:id="Sh2">
    <bibl>Huntington 69304</bibl>
  </witness>
</listWit>
```

There are different options:

1. location referenced + internal
2. double endpoint + internal
3. location referenced + external
4. location referenced + internal
5. parallel segmentation

# From TEI to $\text{\LaTeX}$

TEI encodes text in a semantic way using XML, but this needs to be translated into the presentational type of markup used in LaTeX.

## TEI Apparatus Toolbox: transformation tool

- <http://teicat.huma-num.fr/>
- by Marjorie Burghart
- **Print an edition of a TEI XML edition**, with a TEI-to- $\text{\LaTeX}$  and PDF transformation (`reledmac` → XSL is here).
- Based on running an XSLT script customizable via web form.
- Limited freedom in typesetting customization.



**Critical  
Apparatus  
Toolbox**

## TEI Critical Apparatus Toolbox

[Home](#)[Check your encoding](#)[Display parallel versions](#)[Print an edition](#)[Annotate an image](#)[Get statistics](#)[Help](#)[Download](#)[Credits](#)

The XSLT template for transforming TEI is quite complex, covering all the many possible encoding options.

## Simplified XSLT

I created simplified XSLT documents for teaching purposes, specifically for transforming TEI, although still challenging for complete beginners.

This is the result of the first (and, thus far, last) in-person L<sup>A</sup>T<sub>E</sub>X Ninja workshop (Lang 2022b) that also features other simplified documents help learners get used to XSLT.

# Challenges & Popularizing $\text{\LaTeX}$ in DH

---



# Popularizing $\text{\LaTeX}$ in Digital Humanities

## Challenges

- Digital humanities scholars often seem to have a difficult relationship with print.
- Despite the practicality,  $\text{\LaTeX}$  remains underused in the field.

## Opportunities

- XSLT-to transformations can produce print outputs from TEI data, especially given that the XSL alternative **XSL-FO** (*Formatting Objects*) was discontinued in 2012.
- Large language models may even assist in writing these transformations.
- Print outputs are valuable, especially for textual scholars who prefer working with printed materials.

To provide some insights and issues I've encountered with LaTeX:

- **Digital Scholarly Edition Project:** Created  $\text{\LaTeX}$  print outputs, especially relevant for teaching materials (e.g. *Grazer Repitorium antiker Fabeln (GRaF) 2020*).
- **GAMS Repository:** Our editions are archived in the *Humanities Asset Management System* (GAMS, Stigler and Steiner 2018) that archives XML data from which it generates all representations (HTML, print, RDF) on the fly.

# Practical Implementation Challenges ii

## Issues Encountered

- Difficult to debug errors due to lack of visibility into the underlying LaTeX code  
→ cryptic and time-consuming to fix.
- Errors sometimes occurred even in previously tested documents.
- Often had to run the XSLT stylesheet locally to check results, contrary to the system's intended automatic transformation.
- Sometimes uploaded static PDFs created through the dynamic workflow to ensure successful compilation.

## GAMS Version 4

The original implementation by the previous lead programmer was very complex. The new version did not carry over some  $\LaTeX$  functionalities due to limited use, complicating the debug process further.

If at all, the team would locally compile  $\LaTeX$  and use static uploads in the future.

# Proposed Solutions and Future Directions

## Tutorial Resource

I consider creating a tutorial on the Programming Historian to teach  $\LaTeX$  basics & TEI-to- $\LaTeX$  XSLT transformations and explain their value for digital humanists. This may include:

- Basics of  $\LaTeX$  and why DHers should care.
- Practical use cases, especially in digital scholarly editing.
- Step-by-step guide to transforming TEI to  $\LaTeX$ .

## Future directions for GAMS

- Moving to a static workflow where the XSLT transformation is done locally and the resulting PDF is uploaded to the archival system.
- A tutorial and XSLT starter documents might lower the barrier to entry.
- Figure out when and how print workflows make sense or are needed.

# References

---

- [1] Quinn Dombrowski. **“Does Coding Matter for Doing Digital Humanities?”** In: *The Bloomsbury Handbook to the Digital Humanities*. Ed. by James O’Sullivan. Bloomsbury Handbooks. London: Bloomsbury Academic, 2022, pp. 137–146. DOI: [10.5040/9781350232143.ch-13](https://doi.org/10.5040/9781350232143.ch-13).
- [2] Andrew N. J. Dunning. **“Reledmac. Typesetting technology-independent critical editions with LaTeX”**. In: *RIDE – A review journal for digital editions and resources* 11 (2020). Tools and Environments. URL: <https://ride.i-d-e.de/issues/issue-11/reledmac/>.
- [3] **Grazer Repositorium antiker Fabeln (GRaF)**. Ursula Gärtner (ed.) Graz, 2020. URL: <http://gams.uni-graz.at/graf>.
- [4] Sarah Lang. **A Humanities’ seminar paper with LaTeX – in 10 minutes**. The LaTeX Ninja Blog. Jan. 2019. URL: <https://latex-ninja.com/2019/01/22/a-humanities-seminar-paper-with-latex-in-10-minutes/>.

- [5] Sarah Lang. *Enough reledmac to be dangerous: Scholarly Editing with LaTeX & XSLT*. The LaTeX Ninja Blog. Sept. 2022. URL: <https://latex-ninja.com/2022/09/03/enough-reledmac-to-be-dangerous-scholarly-editing-with-latex-xslt/>.
- [6] Sarah Lang. “Experiments in the digital laboratory. What the Computational Humanities can learn about their definition and terminology from the History of Science”. In: *Fabrikation von Erkenntnis: Experimente in den Digital Humanities*. Ed. by Manuel Burghardt et al. Esch-sur-Alzette: Melusina Press, 2021. DOI: <https://doi.org/10.26298/melusina.8f8w-y749-eitd>.
- [7] Sarah Lang. *First ever LaTeX Ninja workshop at Harvard: “Beyond TEI: Digital Editions with XPath and XSLT for the Web and in LaTeX”*. The LaTeX Ninja Blog. Apr. 2022. URL: <https://latex-ninja.com/2022/04/17/first-ever-latex-ninja-workshop-at-harvard-beyond-tei-digital-editions-with-xpath-and-xslt-for-the-web-and-in-latex/>.

- [8] Sarah Lang. *LaTeX for Archaeologists: An archaeological catalogue from a spreadsheet*. The LaTeX Ninja Blog. July 2021. URL: <https://latex-ninja.com/2021/07/16/latex-for-archaeologists-an-archaeological-catalogue-from-a-spreadsheet/>.
- [9] Sarah Lang. *The LaTeX Newbie's Guide to Using Overleaf for Conference Paper Submissions*. The LaTeX Ninja Blog. Jan. 2022. URL: <https://latex-ninja.com/2022/01/09/the-latex-newbies-guide-to-using-overleaf-for-conference-paper-submissions/>.
- [10] Sarah Lang. “What you really need to know about Digital Scholarly Editing”. In: *The L<sup>A</sup>T<sub>E</sub>X Ninja Blog* (2022). URL: <https://latex-ninja.com/2022/10/30/what-you-really-need-to-know-about-digital-scholarly-editing/>.
- [11] Michael Piotrowski. “Ain't No Way Around It: Why We Need to Be Clear About What We Mean by “Digital Humanities””. In: *Wozu Digitale Geisteswissenschaften? Innovationen, Revisionen, Binnenkonflikte* (November 20–22, 2019). Lüneburg: Leuphana University Lüneburg, 2019.

- [12] Camille Roth. “Digital, digitized, and numerical humanities”. In: *Digital Scholarship in the Humanities* 34/3 (2019), pp. 616–632. URL: <https://doi.org/10.1093/llc/fqy057>.
- [13] Johannes Stigler and Elisabeth Steiner. “GAMS – Eine Infrastruktur Zur Langzeitarchivierung Und Publikation Geisteswissenschaftlicher Forschungsdaten”. In: *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare* 71.1 (2018), pp. 207–216. URL: <https://doi.org/10.31263/voebm.v71i1.1992>.
- [14] Peter R. Wilson and Maïeul Rouquette. *reledmac – Typeset scholarly editions*. Version 2.40.1. Comprehensive TeX Archive Network (CTAN). <https://ctan.org/pkg/reledmac>, May 2024.