



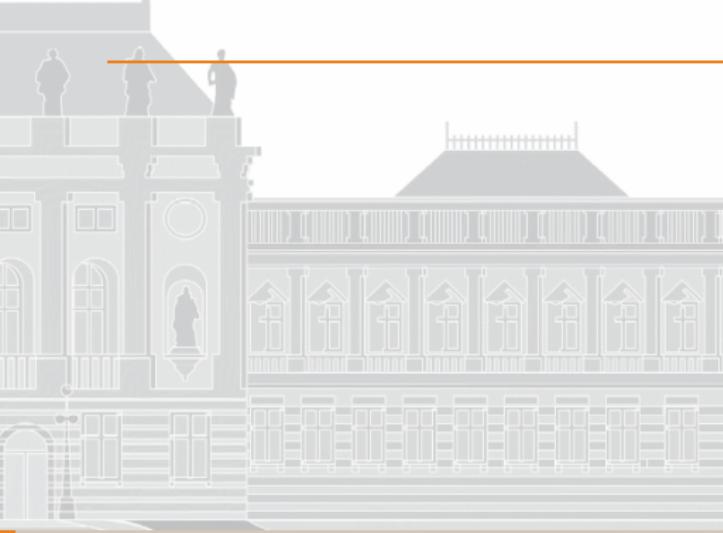
ZENTRUM FÜR
INFORMATIONSMODELLIERUNG
AUSTRIAN CENTRE FOR
DIGITAL HUMANITIES

KARL-FRANZENS-UNIVERSITÄT GRAZ
UNIVERSITY OF GRAZ



Digital Editing & XML/TEI

for premodern books



Sarah Lang

2022

Overview

- 1. The TEI & Digital Editing Session
- 2. Data modelling
- 3. Digital Scholarly Editing
- 4. Editorial Schools and Workflows
- 5. Digital Archives
- 6. Library catalog data
- 7. Annotating with XML markup
- 8. Text Encoding Initiative
- 9. Transcriptions using Transkribus
- 10. TEI for Digital Editing

The TEI & Digital Editing Session

Goals

1. understand basic modelling theory for digitization
2. understand the differences between a digital scholarly edition & other digital resources (such as archives or library catalog data)
3. understand the role of XML/TEI in Digital Editing
4. be able to use the Text Encoding Initiative standard to encode descriptions of pre-modern books

Goals for the next session

1. understanding modelling theory relevant for digitization
 2. understanding the terms 'digital edition', 'digital archive' versus library catalogs
- ✖ *<msDesc>* → next session
- ✖ transcriptions in TEI & Transkribus → later session

Data modelling

3 properties of a model following Stachowiak i

1) Mapping

„Modelle sind stets Modelle von etwas, nämlich Abbildungen, Repräsentationen natürlicher oder künstlicher Originale, die selbst wieder Modelle sein können.“ „Der Abbildungsbegriff fällt mit dem Begriff der Zuordnung von Modell-Attributen zu Original-Attributen zusammen.“ (Stachowiak 1973, 131–132)

Models are always models of something, i.e. mappings from, representations of natural or artificial originals, that can be models themselves. “

2) Reduction

„Modelle erfassen im allgemeinen nicht alle Attribute des durch sie repräsentierten Originals, sondern nur solche, die den jeweiligen Modellerschaffern und/oder Modelbenutzern relevant scheinen.“ (Stachowiak 1973, 132)

Models in general capture not all attributes of the original represented by them, but rather only those seeming relevant to their model creators and / or model users. “

3) Pragmatism

“Eine pragmatisch vollständige Bestimmung des Modellbegriffs hat nicht nur die Frage zu berücksichtigen, *wovon* etwas Modell ist [Abbildungsmerkmal], sondern auch, *für wen, wann und wozu* bezüglich seiner je spezifischen Funktionen es Modell ist.“ (Stachowiak 1973, 132)

Models are not uniquely assigned to their originals per se. They fulfill their replacement function

- a) for particular – cognitive and/ or acting, model using subjects,
- b) within particular time intervals and
- c) restricted to particular mental or actual operations. ”

Stachowiak's General Model Theory

Stachowiak's notion of a model

“ Alle Erkenntnis ist Erkenntnis in Modellen oder durch Modelle und jede menschliche Weltbegegnung überhaupt bedarf des Mediums Modell. ”

→ All knowledge-making is knowledge-making in or through models and all human perception of the world needs models as a medium.

Data modelling

Model = snippet of the real world but it only covers the attributes I chose to be relevant for the task at hand. Thus, the model and the aspect of the real world it models (its subject) diverge.

Digital, standardized & formal models

Standardized models allow us to exchange and analyse data, search/query data. → only *formal models* can be processed digitally, i.e. every digital model is a formal model.

→ Models are simplified representations of parts of the real world.

Why should we care about modelling theory?

- Modelling is a pivotal task in the Digital Humanities.
- When we create digital representations of material objects, those are models.
- Models are by definition subjective, abstracted and not universal.
- Their quality has to be judged in relation to their purpose (Stachowiak's 3rd criterion): →

Different motivations for digital projects

research-driven:

individualized for answering a research question, work-intensive & relatively expensive.

curation-driven

mass-digitization, cookie cutter approach which covers the most important elements for most use cases but can easily miss features relevant to subject-matter experts.

- despite the objects being digitized, users might have to go back to the material objects to fill in the blanks
- but objects become more discoverable & (hopefully) somewhat comparable to larger corpora of similar objects
- superficial digitization can lead to the creation of misleading datasets, e.g. errors or bad tagging → **reparative librarianship** is needed
- potential application for machine learning approaches

What are data?

Definitions of 'data'

Plural of Latin *datum* ('given').

- data aren't exactly a given but rather constructed or created.
- There has been a discussion in the DH whether we should call them *capta* (Drucker 2011)
- Most data (so-called 'givens') are constructed by phenomenotechnical devices (Bachelard 1968), i.e. perceiving devices which translate the (often quite abstract) things they see into data.
- data has to be interpreted.
- Stachowiak's pragmatism criterion means we selectively capture data most important to us, not all possible aspects (!).

Data resulting from cataloging & digitization contains interpretations & is thus, *per definitionem* always subjective & incomplete (!).

data \neq original!

Digital Scholarly Editing

What is a Digital Scholarly Edition? i

Digital editions (as per Sahle 2016):

- discipline-independent & not exclusive to text
- overcome the limitations of print (financial, inter-/ multimediality, etc.)
- follow a digital paradigm like printed editions follow the print paradigm (“shaped by the technical limitations and cultural practices”)
- “A digitised edition is not a digital edition.” → it’s not about the storage medium (just being on the internet doesn’t make it a digital edition!)
- “A digital edition cannot be given in print without significant loss of content and functionality.”
 - different views, interactivity, searchability
 - facsimiles alongside diplomatic transcription and reading versions
 - all generated from the same source data: **single source principle** → How? Data transformation!
 - ≠ Digital Archive, text corpus or facsimile → has critical engagement

What is a Digital Scholarly Edition? ii

We need to start at the beginning:

What is a scholarly edition?

- originally stems from ecdotics (*ecdotica, ecdotique, Editorik / Editionswissenschaft*) → tradition of Philology
- **goal = textual criticism:** reconstruct the original version of a text transmitted to us via textual witnesses (*Lachmannian paradigm*)

Sahle 2016

“ A scholarly edition is the critical representation of historic documents.
Edition ist die erschließende Wiedergabe historischer Dokumente. ”

What is a Digital Scholarly Edition? iii

Criteria from Sahle 2016:

1. **Representation:** recoding a document and its transformation (in the same or other media)
 - Visually: image reproduction, facsimile
 - Textually: transcription.
 - varying degrees of closeness/abstraction with regard to the original document
2. representation \neq presentation
3. **Critical engagement** (based on scholarly agenda). “Critical engagement without representation is not an edition—but an examination, a catalogue or a description.”
 - textual criticism
 - historic criticism
 - bibliographic criticism
 - material criticism
 - visual criticism
 - etc.

What is a Digital Scholarly Edition? iv

4. **Documents:** “every non-abstract object that is the subject of an edition can be called a document.”
5. **Historic:** editions “explain what is not evident to the present-day reader. In short, they bridge a distance in time, a historical difference. Texts that are created today do not need to be critically edited. They can speak for themselves.”

Sahle 2016

“ A representation without [critical] treatment or the addition of information is not an edition—but a facsimile, a reproduction or—nowadays—a digital archive or library.

Critical representation as a compound notion of editing aims at the reconstruction and reproduction of texts and as such addresses their material and visual dimension as well as their abstract and intentional dimension (Sahle 2016). ”

What is a Digital Scholarly Edition? v

Questions to ask if you want to know if it's a Digital Scholarly Edition:

1. **Is there a full representation** of the subject in question?
2. **Is it critical?** → processing rules stated and applied, scholarly knowledge included to make the document easier to understand, regarding material, document genesis/creation, context and reception?
3. **Is the edition of academic quality?** → transparent and rigorous edition process, responsibilities stated, enables future research on a reliable basis.
4. **Does the edition follow a digital paradigm?** → makes use of the possibilities of the digital, not printable without major loss of content or functionality.

What is a Digital Scholarly Edition? vi

Ideally, a DSE should also implement the FAIR criteria

Findable. in library catalogs, discovery systems or repositories, with a persistent identifier

Accessible. free for any user, no access restrictions? (from open access to usability, language selection, etc.)

Interoperable. data in standardized and widely used format (for example TEI or similar standards), allowing for reuse and data exchange outside of the project.

Reusable. data accessible (individual download, aggregate download, repository, API)? Licenses allowing reuse. Data creation, modelling and processing documented adequately so that others can make sense of it?

- RIDE (review journal for digital editions and resources)
- RIDE Criteria for Reviewing Digital Editions and Resources

Further reading

Patrick Sahle. "What is a Scholarly Digital Edition?" In: *Digital Scholarly Editing: Theories and Practices*. Ed. by Matthew J. Driscoll and Elena Pierazzo. Cambridge: Open Book Publisher, 2016, pp. 19–39.
URL: <https://books.openedition.org/obp/3381>

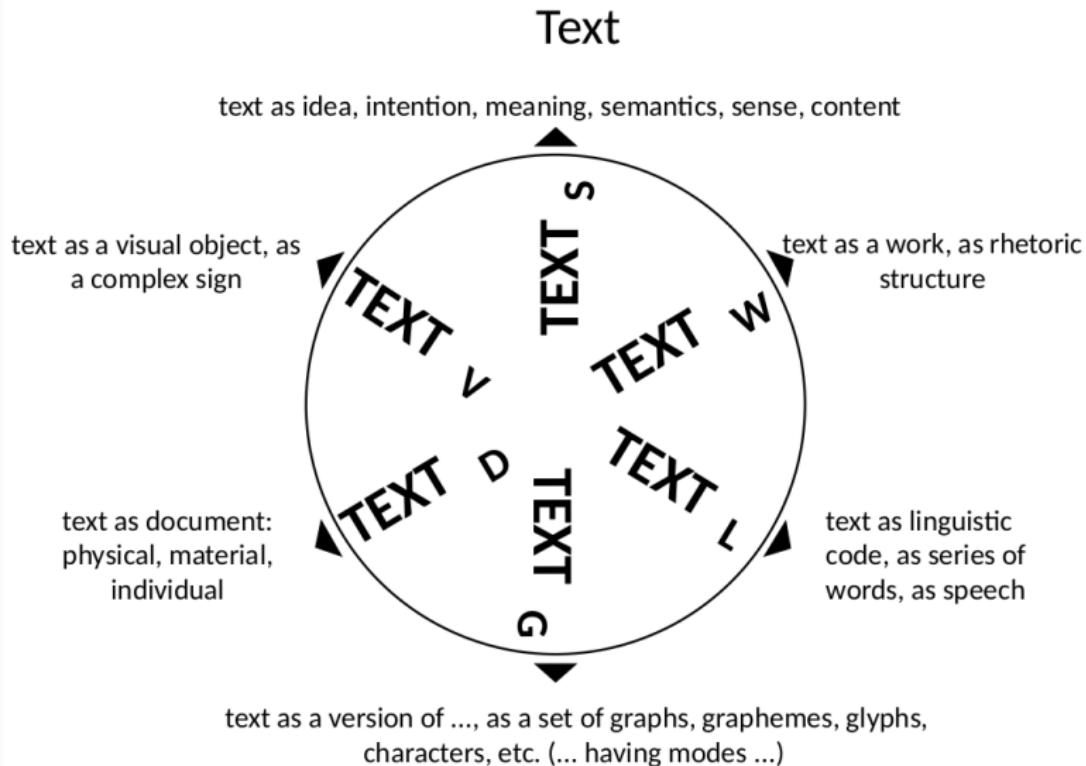
Blogpost (short summary)

What you really need to know about Digital Scholarly Editing, The L^AT_EX Ninja Blog,
30.10.2022. = (Lang 2022)

Editorial Schools and Workflows

What is a text?

Patrick Sahle's text wheel



Edition workflows and schools i

The following slides are adapted from Georg Vogeler's slides on digital editing.

Traditional Workflow of Philological Editing

- **Heuristics:** Find your textual witnesses
- **Transcription:** transfer the text into your preferred alphabet (from the original/from a photo)
- **Collation:** Compare the textual witnesses
- **Recensio:** Evaluate the variants and create a stemma (commenting)
- Write your introduction
- Typesetting
- Create an index (referring to pages)
- Print and distribution

In the digital realm:

- edition
- beta version
- just TEI data publication
- hybrid (web and print) publication
- minimal edition (adaptation from minimal computing)

A digital edition is social, iterative, a process...

Edition workflows and schools ii

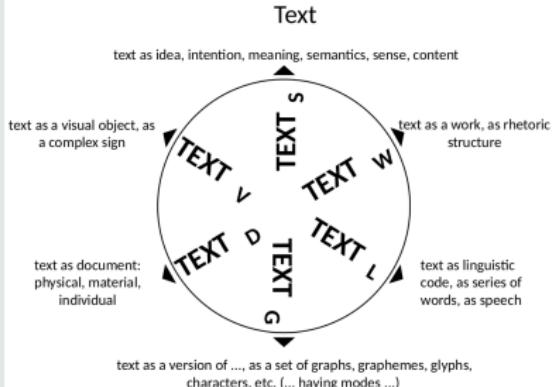
„Edition“

- a particular version of a book
- a particular version of a product
- all the copies of a book that are printed or published at one time

„Historical-critical“ Edition

- Documentation of the history of the text transmission in the “apparatus of variants”
- Critical evaluation of the textual transmission

Patrick Sahle's text wheel



Karl Lachmann (1793–1851)

Editorial intervention: Philological knowledge (linguistics, style) allows us to emend fragmentary textual transmission.

Editorial Schools

- Lachmann ('Historisch-Kritische Ausgabe', stemmatology)
- Reading text vs. critical edition = reader-oriented pragmatism / modernised edition / 'reading text'
- Last authorized edition
- First edition / *editio princeps*
- Main manuscript (Bédier 1937, 'Leithandschrift')
- 'diplomatic edition / transcription'
- Typographic/photographic facsimile
- 'documentary editing' (Tanselle)
- variorum edition
- genetic edition / *critique génétique* (Gabler)
- Copy-Text-Theory (Greg 1950/3, Bowers 1976, 1978)
- New Philology (Stackmann 1964, 1993, 1994, 1999; Ruh 1978, Cerquiglini 1989)

If you wanted to practice...

Write a RIDE review on a digital edition (suggestions for review)
according to their review criteria!

(If you actually plan on submitting, contact the editors first. 2000 words minimum,
project from your field. Peer-review can spot things you're unsure about.)

Digital Archives

Digital long-term archiving

- ensuring the authentic and sustainable availability of digital resources on the level of the bitstream and on a semantic level
- integral principle to every form of sustainable data storage
- begins with the data production in a sustainable data format (and ideally, following a recognized data standard)
- requires standardization of data formats and archiving workflows
- serves both the dissemination as well as the preservation of digital content
- not just about technical solutions but also institutional stability and policies

A digital archive

- is more than a mere collection of scanned book pages or digitized images, etc.
- if offers metadata, norm data and controlled vocabularies

Digital Archives

Digital Archive

- organized collection of digital objects (text, images, audio, video and multimedia streams)
- digital objects are described by standards both in terms of contents (e.g. TEI) and bibliographically (e.g. Dublin Core)
- published sustainably using interfaces, services and APIs (e.g. OAI-PMH)
- digital objects have unique, persistent and citable identifiers (e.g. DOI, URN, PURL, PID)
- authenticity of objects is checked by means of digital signatures or checksums ('Is the number of bytes in the object still the same it used to be?')

Trustworthy Digital Archives

as defined by the Research Libraries Group (RLG)

- secure organizational structure and legal status
- financial sustainability
- technological and procedural aptness
- ensuring data and system security
- documentation and transparency
- conformity with the OAIS standard

→ retro-digitizing objects but also standards for new (born digital) resources.

Digital Archives vs. Digital Editions (Sahle 2007)

Digital documents

Material objects are the target of digitization but digitization doesn't reproduce them – it represents them in a digital format. A digital document is a view on the original material object.

Archive

Traditionally an archive is an ordered collection of documents with the goal of documenting them, preserving them in the long-term and making them accessible. → this function isn't only carried out by actual archives but also by museums, libraries and other cultural heritage institutions. Traditional archive material doesn't need to be represented because its physical objects can be accessed directly.

(we have already def'd editions)

Editions are often based in archival material. The edition isn't a storage device, it is a publication of a historical source and the editorial work done on it. While the edition contains lots of editorial work and enrichment, archival documents are usually original and largely unprocessed. A non-digital archive isn't a form of publication in and of itself – a digital archive *is* a form of publication, too, like a digital edition → blurs the lines a bit.

We could say that **the difference lies in the depth of data enrichment and editorial work.** → once a presentation form is provided, data from a digital archive becomes a digital edition.

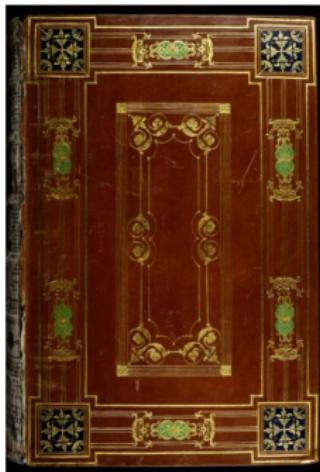
Digital Archive

A main goal of a digital archive is to preserve and publish a specific choice of documents as a collection (ideally, representative and well-balanced). The digital objects aren't necessarily direct representations but may have undergone editorial intervention (e.g. normalized orthography). Its documents should be uniform in how they are encoded and processed. In the case of the single source principle, the edition is generated dynamically ('on the fly') from the archived source data. (If any edition is provided).

Library catalog data

Theuerdank (Graz Sondersammlungen – Rara 1 III 11723)

Pfintzing, Melchior, et al. Die geuerlicheiten und einsteils der geschichten des loblichen streytparen und hochberümten helds und Ritters herr Tewrdannckhs. 1517.



Bibliographic Metadata

Title	Die geuerlicheiten und einsteils der geschichten des loblichen streytparen und hochberümten helds und Ritters herr Tewrdannckhs
Author	Pfintzing, Melchior G W
Participants	Maximilian <I., Heiliges Römisches Reich, Kaiser> G W
Published	Nürnberg : durch den Eltern Hannsen Schönsperger, 1517
Description	[290] BL ; 36 cm : zahlr. Ill. (Holzschn.)
Annotation	In Fraktur
Annotation	Vorlageform des Erscheinungsvermerks [Kolophon]: Gedruckt in der Kayserlichen Stat Nürnberg durch den Eltern Hannsen Schönsperger Burger zü Augspurg
Bibl. Reference	(VLID)3871082
Keywords	Maximilian <I., Heiliges Römisches Reich, Kaiser> G W / Maximilian <I., Heiliges Römisches Reich, Kaiser, Teuerdank> G W
URN	urn:nbn:at:at-ubg:2-22673
VD16	M 1649

Links

Download	 PDF
Reference	 Universitätsbibliothek Graz
Availability	 In My Library
Archive	 METS (OAI-PMH)
IIIF	 IIIF Manifest

MARC XML example i

Library of Congress MARC XML for Theuerdank (simplified for demonstration purposes): <https://www.loc.gov/item/2021667794>

```
<record xmlns="http://www.loc.gov/MARC21/slim"
         xmlns:zs="http://docs.oasis-open.org/ns/search-ws/sruResponse">
<leader>02892nam a22004093i 4500</leader>
<controlfield tag="001">22061662</controlfield>

<datafield ind1=" " ind2=" " tag="500">
    <subfield code="a">"BSB Shelfmark: Rar. 325 a"--Note extracted
        from World Digital Library.</subfield>
</datafield>

<datafield ind1=" " ind2=" " tag="500">
    <subfield code="a">Original resource extent: 290 unnumbered
        sheets : illustrations.</subfield>
</datafield>
[...]

</record>
```

MARC XML example ii

```
<record xmlns="http://www.loc.gov/MARC21/slim"
         xmlns:zs="http://docs.oasis-open.org/ns/search-ws/sruResponse">
    [...]
    <datafield ind1=" " ind2=" " tag="520">
        <subfield code="a">
            Among the many endeavors undertaken by the Holy Roman Emperor Maximilian I (1459--1519) to further his legacy was his plan of an epic retelling of his own life story in the form of several works.
            [...]
            Johann Schönsperger, a printer in Nuremberg, did the first, very small print run in 1517, to be delivered to other princes and sovereigns after the Emperor's death.
            [...]
            Each of the 118 chapters is decorated by a xylograph (wood engraving). The preparatory drawings for the xylographs were created by the artists Leonhard Beck, Hans Schäufelein, and Hans Burgkmair the Elder. The black-letter type of the Theuerdank, designed by calligrapher Vinzenz Rockner, was to become very influential for the development of German typography.</subfield>
    </datafield>
</record>
```

Annotating with XML markup

XML: eXtensible Markup Language

- W3Schools Tutorial
- paradigm of the separation of form and content
- XML is a metalanguage

.XML

- RSS, SOAP, XAML
- MathML, GraphML
- XHTML
- RDF
- KML
- Scalable Vector Graphics (SVG)

“ Extensible Markup Language (XML) is a **markup language** and file format for storing, transmitting, and reconstructing arbitrary data. It defines a **set of rules for encoding documents** in a format that is **both human-readable and machine-readable**. (Wikipedia) ”

XML rules

XML can be checked for **validity** (validation if it complies with a standard) and **well-formedness** (following the rules of XML) → will only be parsed if well-formed. Thus: **Heed thy error messages!**

There are rules on how elements can be named (you can look them up if relevant or will get informed by an error message).

`<key>value</key>` . XML as a key value notation

Rules

- Hierarchical nesting below the root
- exactly one root element, i.e. one out-most russion doll
- start and end tag
- tag names are case-sensitive (!)
- empty elements allowed (& can be shortened)

Minimal example

```
<?xml version="1.0" ?>
<root>
  <element attribute="value">
    content
  </element>
  <!-- comment -->
</root>
```

XML rules i

Prolog

```
<xml version="1.0" encoding="utf-8"> ..... XML declaration  
<?xsl-stylesheet type="text/xsl" href="my.xsl"?> . processing  
instructions (optional)
```

you can include document models (optional)
DTD, XML Schema, RelaxNG, Schematron

entities ‘protected’ characters that have a meta meaning in XML like:

<	<
>	>
&	&

XML family and vocabularies

- XML** structured description of data
- XPath** navigating xml documents
- XML Schema** strict data model
- XSL** Extensible Stylesheet Language
- XSLT** XSL-Transformations, i.e. transforming XML documents
- XSL-FO** formatted output (e.g. print)
- XQuery** query language for XML databases
- and more**

- **(X)HTML** Hypertext Markup Language
- **EAD** Encoded Archival Description
- **TEI** Text Encoding Initiative
- **CEI** Charters Encoding Initiative
- **MEI** Music Encoding Initiative
- **LIDO** Lightweight Information Describing Objects (describing museum or collection objects)
- **SVG** Scalable Vector Graphics
- **KML** Keyhole Markup Language (geography)
- **MathML**
- **CML** Chemical Markup Language,
- ...

Practice!

Open a new XML document in your editor (Oxygen).
Create new elements and find 3 ways to break it so that you get an
error.
Then fix the error.

Teaser: Digital editions – some examples i

GAMS ufas

Urfehdebücher der Stadt Basel – digitale Edition

Universität Basel

Ratsbücher O10, Urfehdebuch X
(1563-1569)

Ziterverschlag:
Urfehdebuch X der Stadt Basel (1563-1569), in: Die Urfehdebücher der Stadt Basel – digitale Edition, hg. v. Susanna Burghartz, Sonja Calm und Georg Vogeler (Basel/Graz 2016. (Gesetzt verändert am 31.1.2017);
hd:11471/1010.2.1.

Zeitraum:
03.01.1563 bis 04.12.1569

fol. 1v

[Aeno d[em]ni 1563]

Beginn Undatiert 1563 1564 1565 1566 1567 1568 1569 Ende

Datum: 1564-02-25

Person
Berner Däbler ein welischer von wegen das er sich mitt win überleidet, und einer denphale gesondert halb in unser geistreich [eigentl.] gefangenschaft inns Touzusell gefangen hafft auf den 25 tag Hornungs ussgelossen iusavit ut moria

Tatbestand
absichtlich (Durchgräbung)

Straße
Orte

Schlagwort
[aufgegraben]

Beginn Undatiert 1563 1564 1565 1566 1567 1568 1569 Ende

Teaser: Digital editions – some examples ii

CoReMA → Look at: <http://gams.uni-graz.at/corema>

The screenshot shows the CoReMA digital edition interface. At the top, there are navigation icons and a URL bar showing http://gams.uni-graz.at/corema/pat_recipes. Below the URL bar is a horizontal menu with links: Home, Recipes, Manuscripts, Analysis, Models, Project, and Todos. The main content area displays a medieval manuscript page with text in Gothic script. A search bar is located at the top right of the manuscript view. Below the search bar are buttons for Ingredient, Dish, and Tool. Underneath these buttons are tabs for Name and Ref. To the left of the manuscript, there is a sidebar with a question mark icon and a link to the change log. At the bottom of the page, there is a navigation bar with links for Description, Recipes, Transcription, Full-Width, and Metadata.

This is the most recent version of this document. See change log

[269r]

Ein guet kuechenpueech

Fish as collared pork (Nr. I) [cooking recipe] ¶

Wlldw machen ein schwein koppff (pig head) von vischen (fish) Sonym cherpfen (carp) oder schleyen (tench) oder perben (common barbel) vnd schueb sye schon vnd schlach sy auf vnd mach sy zw stuecken vnd thu sy In ein pfan (pan) vnd gewß dar an halbs wasser (water) halbs wein (wine) vnd daz dy pruee (water wine broth) nuer ains vin gers hach gee uber dy visch

Search:

Ingredient Dish Tool

Name Ref.

?

Description Recipes Transcription Full-Width Metadata

Teaser: Digital editions – some examples iii

Furnace and Fugue (music in MEI, text and images)

The screenshot shows a digital edition interface for 'Furnace and Fugue'. At the top, there's a navigation bar with links like 'Home', 'About', 'Search', 'Help', and 'Log In'. Below it is a secondary navigation bar with 'DIGITAL EDITION', 'ESSAYS', 'IMAGES', and 'CONTENTS'. The main content area has a red header 'f&f' and a dropdown menu 'Normalized English'. On the left, there's a 'Comparative View' section with a large image of a classical figure and Latin text. The right side features a musical score with three voices labeled 'Voice 1', 'Voice 2', and 'Voice 3'. Below the score is a section titled 'the first Epigram' with a poem in English. At the bottom, there are search and zoom icons.

This screenshot shows a search results page from the 'Furnace and Fugue' digital edition. The top navigation bar is identical to the previous screenshot. The main area is titled 'RESULTS' and contains two items: 'Title page' and 'Emblem 14'. Each result item includes a thumbnail image, a title, and a 'VIEW' button. Below the results, there's a 'TAGS' section with terms like 'Aegeus, Apple, Arithusa, Alas, Classical clothing, Column, Dragon, Earth, Embracing, Female, Fence,' and a 'Emblem Collections' section.

From TEI to Edition

What did all those examples have in common?

- all are digital editions
- all are based on data in TEI-XML

That's what we want. – How do we get there?

Remember

- XML is great for capturing & long-term archiving data in its whole complexity
 - ...but also ends up chaotic with a lot of detail
 - theoretically human-readable
 - in practice, it's better to create a new representation of the data to read
- **Desired result:** view/presentation, visualization and interaction with the data as a website (i.e. HTML data)
- → we need to transform the data

TEI Publication Tools

There is a plethora of out-of-the-box solutions to make editions out of TEI data. → [TEI List of Resources for Publishing and Delivery Tools](#), for example:

1. teiPublisher
2. EVT
3. TEICHI
4. Versioning Machine
5. Boilerplate
6. Oxgarage
7. GAMS ...

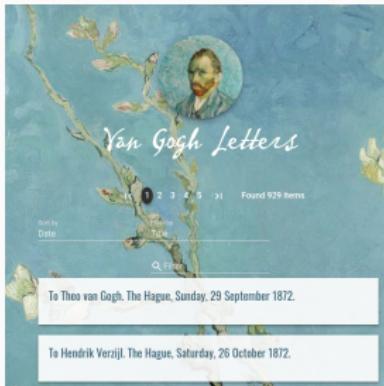
TEI Publisher: The Instant Publishing Toolbox

- <https://teipublisher.com/index.html>
- https://wiki.tei-c.org/index.php/TEI_Publisher
- <https://wiki.tei-c.org/index.php/TeiPublisher>
- <http://hisoma.huma-num.fr/exist/apps/tei-publisher/index.html>
- “Testimonial”: Florian Chiffolleau (2020), Blog Post: *Publication of my digital edition – Working with TEI Publisher.*



TEI Publisher: The Instant Publishing Toolbox ii

Demos: Van Gogh Letters ● Early English Books Online (EEBO)



This screenshot shows a detailed view of a letter from Theo van Gogh to his father, dated Sunday, 29 September 1872. The interface includes a header with the title "To Theo van Gogh, The Hague, Sunday, 29 September 1872.", a search bar, and a navigation menu. The main content area displays the letter's text in two columns: "Original text" and "Translation". The "Original text" column shows the Dutch text, and the "Translation" column shows the English translation. A "View Notes" button is also present. The letter discusses the racing meet at the "Rennbahn" and the author's visit to his parents' house. A note explains the meaning of the word "oportunitas". The EEBO logo is visible at the bottom right.



Arcus recessum loc est Hierophantika Agrippa Graecis ergo modis cogitare, ad Exercitaciones librorum apud tempus descriptis, deinceps latentes, secretos & occultos, quae sunt in natura, corpora, ex via figurativa, et sensu, et novem auctoritate medicorum, philosophorum, fabrorum, theologicis, philosophicis, politicis, Chymicis & medicis perulgatos authores ediderim ac in publicam mundi lucem nunc tandem sparserim, while que impensis de tot praeiunctis ad diplomandam inscriperim, postquam rem ipsam, mea studia ac mestrum, vestrum que canadore simul respserit, miris abe que dubio doctiss. Inter Christianos, ut spinos, nullus est ad o' rerum omnium impensis, si modo sensu communicalat & prima literarum rudimentatenet, qui non aliquando libros Eutheniorum (quorum ingle-

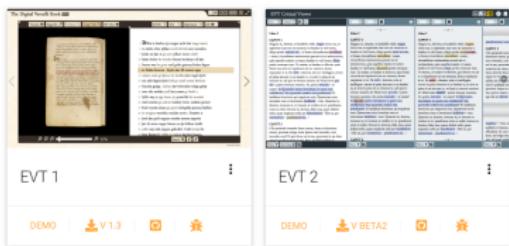
MICHAELIS MAIERI COMITIS PAL.
CAES. •RAE•ATIO AD LECTORES.

*S*i miremni lectori leversoli, quid Case* milit exierit car buse tractu ARC ANISSIMORVM nomine insignitus post tot & tantos de Hierophantis, algorij, fabris, theologicis, philosophicis, politicis, Chymicis & medicis perulgatos authores ediderim ac in publicam mundi lucem nunc tandem sparserim, while que impensis de tot praeiunctis ad diplomandam inscriperim, postquam rem ipsam, mea studia ac mestrum, vestrum que canadore simul respserit, miris abe que dubio doctiss. Inter Christianos, ut spinos, nullus est ad o' rerum omnium impensis, si modo sensu communicalat & prima literarum rudimentatenet, qui non aliquando libros Eutheniorum (quorum ingle-

Edition Visualization Technology (EVT)

- <http://evt.labcd.unipi.it/> - <http://evt-project.sourceforge.net/>
- <https://visualizationtechnology.wordpress.com/>
- Roberto Rosselli Del Turco, Giancarlo Buomprisco, Chiara Di Pietro, Julia Kenny, Raffaele Masotti and Jacopo Pugliese: *Edition Visualization Technology: A Simple Tool to Visualize TEI-based Digital Editions*, Journal of the Text Encoding Initiative 8 (2014/15): Selected Papers from the 2013 TEI Conference “TEI Processing: Workflows and Tools”; <https://doi.org/10.4000/jtei.1077>
- Del Turco, Roberto Rosselli. “Designing an Advanced Software Tool for Digital Scholarly Editions: The Inception and Development of EVT (Edition Visualization Technology).” *Textual Cultures*, vol. 12, no. 2, 2019, pp. 91–111. JSTOR, www.jstor.org/stable/26821538.

A light-weight, open source tool specifically designed to create digital editions from XML-encoded texts, freeing the scholar from the burden of web programming and enabling the final user to browse, explore and study digital editions by means of a user-friendly interface.



TeICHI – Bringing TEI Lite to Drupal

- TEICHI: a TEI lite integration into Drupal (<http://teichi.org>)
- Sebastian Pape, Christof Schöch and Lutz Wegner: *TEICHI and the Tools Paradox. Developing a Publishing Framework for Digital Editions*, Journal of the Text Encoding Initiative 2 (2012): Selected Papers from the 2010 TEI Conference; <https://doi.org/10.4000/jtei.432>
- TEICHI wiki

Accueil Introduction Essai sur le récit Dossier critique Aide

Essai sur le récit

:: Introduction
L'introduction présente Bérardier de Bataut et l'*Essai sur le récit* ainsi que la présente édition électronique.

:: L'Essai sur le récit
L'édition électronique annotée de l'*Essai sur le récit* de Bérardier de Bataut permet l'affichage alternatif d'un texte de lecture ou d'une transcription linéaire.

:: Dossier critique
Le dossier critique donne des informations bio-bibliographiques et fournit la documentation éditoriale et technique du projet d'édition.

:: Accès à l'aide
L'aide fournit des renseignements sur les possibilités de navigation et les fonctionnalités du site.

Référence bibliographique
François-Joseph Bérardier de Bataut, *Essai sur le récit, ou Entretiens sur la manière de raconter* (Paris : Charles-Pierre Berton, 1776). Édition électronique sous la direction de Christof Schöch, 2010. URL : <http://www.berardier.org>. (Version 0.6, 12/2010.)



The title page features the title 'ESSAI SUR LE RÉCIT, OU ENTRETIENS SUR LA MANIÈRE DE RACONTER' in large, decorative capital letters. Below it, the author's name 'PAR M. F. J. BERARDIER DE BATAUT, Essai en Théâtre, Récit de Notre-Dame de Sionne & autres Projets d'Éditions de l'Université de Paris.' is printed. At the bottom, it says 'A PARIS, Chez CHARLES-PIERRE BERTON, Librairie de l'Université de Paris, 1776. M. DCC. LXXVI. Aute Approbation & Privilege du Roi.' with a date '1776' written vertically.

The Versioning Machine

- Versioning Machine (<http://v-machine.org/>)
- The TEI Versioning Machine Versioning Machine (Duke)

"Faith is a fine invention"

7 Total Versions ▾ Line numbers Bibliographic panel

Bibliographic Information	Version a660: A 660, verse embedded in letter to	Version h201: H 201, fascicle version of poem.
<p>"Faith is a fine invention"</p> <p>by Emily Dickinson</p> <p>Original Source</p> <p>See Witness List.</p> <p>Witness List</p> <ul style="list-style-type: none">• Witness a660: A 660, verse embedded in letter to Samuel Bowles.• Witness h201: H 201, fascicle version of poem.• Witness h72: H 72, fascicle version of poem.• Witness p1891: Published as poem XXX in the second volume of Todd and Higginson's Poems of Emily Dickinson.• Witness 11894: Letter to Samuel Bowles published in Todd's edition of Dickinson's letters.• Witness cp32: Published as poem LVI in Martha Dickinson Bianchi's Complete Poems of Emily Dickinson.	<p>1 "Faith" is a fine invention</p> <p>2 When Gentlemen can see -</p> <p>3 But Microscopes are prudent</p> <p>4 In an Emergency .</p>	<p>1 Faith is a fine invention</p> <p>2 For Gentlemen who see -</p> <p>3 But Microscopes are prudent</p> <p>4 In an Emergency !</p>

TEI Boilerplate

- <http://teiboilerplate.org/>
- Boilerplate (<http://dcl.slis.indiana.edu/teibp/>)
- Demo
- TEI Boilerplate Wiki
- Slides on TEI Boilerplate
- Exercise on TEI Boilerplate
- Poster on TEI Boilerplate

TEI Boilerplate Examples & Features

Nested <div> elements and headings

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam gravida, urna eget venenatis mollis, ipsum tellus hendrerit tellus, sed vestibulum urna quam eget enim. Pellentesque odio lacus, laoreet quis commodo eu, vehicula congue purus.

Nested <div> elements and headings

Vivamus rhoncus erat mi. Vestibulum sapien sapien, pulvinar id vehicula in, porta tristique ligula. Sed commodo diam sit amet dui posuere luctus. Morbi tempus sem adipiscing nisl molestie pharetra.

Nested <div> elements and headings

Sed fermentum dui vel enim consectetur malesuada. Sed a mi felis. Curabitur sit amet mi augue, ac volutpat leo. Maecenas vitae condimentum orci. Morbi non quam sed neque tincidunt tempus posuere sit amet purus.

Nested <div> elements and headings

Nested <div> elements and headings

Nested <div> elements and headings



Oxgarage

- REST-based transformation web service for TEI documents (<https://oxgarage.tei-c.org/>)
 - allows you to transform a number of (markup-based) formats to TEI or create them from TEI (`.html`, `.tex`, `.docx`, ...)
 - Based on TEI base stylesheets (<http://www.tei-c.org/Tools/Stylesheets/>)
- ✓ very useful tool
- ✗ not all TEI elements taken into account, not customizable

Oxgarage

...offers a set of standard stylesheets to convert between TEI-encoded XML documents and other documents, mostly markup-based, such as `.docx`, `.html` or `.tex`.

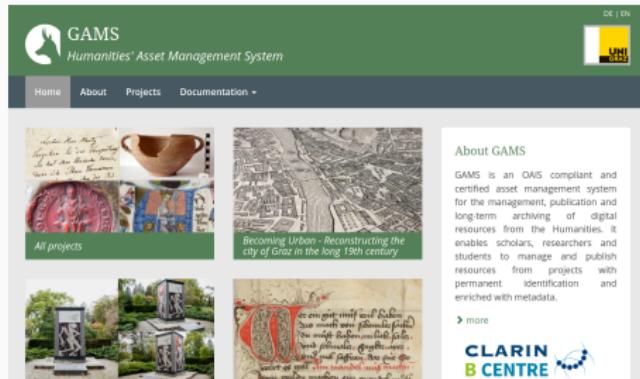
However, not all elements are covered and you have no control over how exactly they are processed.

Tools make things easier superficially but can come at a cost: potential bugs, less than ideal usability, lack of control and customizability:

→ to customize we need to write our own transformation using XSL stylesheets

GAMS (Humanities Asset Management System) i

- GAMS
(Geisteswissenschaftliches Asset Management System = AMS for the Humanities)
- based on **FEDORA** (*Flexible Extensible Digital Object Repository Architecture*) = infrastructure dedicated to the persistent archival and management of resources considered to be worthy of long-term preservation.



<https://gams.uni-graz.at>

GAMS (Humanities Asset Management System) ii

- user access provided through the **Cirilo** client
- **functionalities:** object creation and management, versioning, normalization & standards, choice of data formats.
- offers a plethora of **pre-defined content models** for data such as TEI, MEI, LIDO, SKOS, ontologies, R code and story lines
- → offers publication pipelines but is highly customizable
- more info:
<http://gams.uni-graz.at/doku>



Long-term archiving heritage data i

Long-term preservation

denotes the process of maintaining, curating and keeping data usable over a long period of time (10+ years).

Key functionalities of long-term archiving architectures include

- persistent identification,
- versioning,
- support of different data formats,
- management of associated metadata,
- data export and retrieval,
- security and scalability.

Special emphasis is placed on

- sustainability,
- citability &
- guarantee of long-term access to the contained resources.

XML is great for long-term archiving!

Consequently, data formats and software used for preservation should follow **open source & non-proprietary standards**; data is ideally encoded in an **unicode XML format**:

- plain text files are small
- human- & machine-readable
- recognized standard stable since 1998
- more on XML later...

Examples of data quality in GAMS i

The screenshot shows a search result for the year 1563. The top navigation bar includes links for Home, Projekt, Urfehdebuch, Kategorien, Datenkorb (0), Erweiterte Suche, Volletextsuche, and a search input field. The main content area displays a card for 'Ratsbücher O10, Urfehdenbuch X (1563-1569)'. The card includes a thumbnail, a title, a 'Zitierschlag' section with a URL, and a 'Zeitraum:' section indicating the period from 03.01.1563 to 04.12.1569. Below this is a large preview image labeled 'fol. 1v'.

The screenshot shows a detailed view of a TEI XML object. It includes sections for 'Dataset', 'Object type: cm:TEI', 'Permalink: https://gams.uni-graz.at/o:graf.2387', 'Handle: No handle assigned', 'Publication date: 2020-07-22T16:42:03.702Z', 'Last modified: 2020-09-08T12:04:50.176Z', 'License: Resource licensed under Creative Commons BY-NC 4.0', 'Versions', 'View object (HTML | Fedora profile)', 'Export Metadata (Dublin Core)', and 'Files in this object (current version)' which lists 'TEI_SOURCE tei+xml' (29.1 KB). There are also download links for 'Download selected files (zip)' and 'Download all files (zip)'.

← <http://gams.uni-graz.at/o:ufbas.1563>

- **top image:** info on recommended citation, downloadable source data in XML & RDF
- **below:** ‘data view’ of the Dublin Core (DC) metadata for an XML object in GAMS (ensuring citability, etc.)

← <http://gams.uni-graz.at/archive/objects/o:graf.2387/methods/sdef:Object/getDC?>

Examples of data quality in GAMS ii

- GAMS isn't an out-of-the-box tool – it's something between a **Content Management System (CMS)**, a **repository** (long-term archiving) and a **publication platform**.
- The XSLT transformations applied to files in GAMS are custom but there are **wippets** (standard javascript functionalities → *widget + snippet*, example here) and **standard templates** which can be reused.
- as a publication platform: presentation and long-term archiving, allowing for persistent access to the data, versioning, etc.

Examples of data quality in GAMS iii

One more example: <https://gams.uni-graz.at/beurb>

Metadata view on XML data

Nicht sicher | gams.uni-graz.at/beurb.bei.1910

... PDF-Mit-Resen... Tools | About... Index | Smartphone-G... Gridview-G... Stellmärkte... Digital-Humanis... Stellmärkte... Jobs | EAON-TH...

Quellen • Raum-zeitliche Navigation • Analyse • Projekt •

Illustrierter Führer durch die steiermärkische Landeshauptstadt Graz

METS Source TEI SOURCE

Über den Reisebericht Faksimile & Text Karte & Text



Titel	Illustrierter Führer durch die steiermärkische Landeshauptstadt Graz
Autor	Walter von Semetkowsky
Datum	1910
Ort	Graz
Verlag	Uv. Mosers Buchhandlung (J. Meyerhoff) k.k. Hof-Buchhändler
Umfang	Seite 1-83
Permalink	https://gams.uni-graz.at/o/beurb.tel.1910

Map view linked to edited text

Nicht sicher | gams.uni-graz.at/beurb.bei.1910/beitrag/10/gesamt.html#text

... PDF-Mit-Resen... Tools | About... Index | Smartphone-G... Gridview-G... Stellmärkte... Digital-Humanis... Stellmärkte... Jobs | EAON-TH...

Quellen • Raum-zeitliche Navigation • Analyse • Projekt •

Über den Reisebericht Faksimile & Text Karte & Text


Legende:
- Rote Linie
- Orte
- Gelbe Linien
- Verkehrsinfrastruktur
- Staatsgebäude
- Parks
- Plätze
- Straßen
- Wasser

[1] Praktische Zeiteinteilung.
Wer einen Tag in Graz verbringen möchte, beginnt sich vormittags vor allem in den Schätzen der Elektrischen Seefahrt, i. V. Kaiser-Franz-Josef-Tal ab sieben S. 20, dessen Besuch kein Fremder unterlassen darf (Mindestzeitraumend 1 Stunde); diesen schließt sich ein Spaziergang durch den Stadtzaun (S. 27) und ein Rundgang durch die innere Stadt (S. 30). Nachmittags pflegt man einen Ausflug zum Hünstein (S. 78) zu machen; von der Höhe hat eine herrliche Aussicht auf die Stadt und ihre Umgebung; wenn es dem Fremden paßt, läßt sich in Anschluß daran mit Benützung der elektrischen Bahn auch Maria-Trost (S. 78) besuchen. Für den Abend sind vielleicht noch die Vorführungen des auch verworbenen Anforderungen entsprechenden Stadttheaters (S. 10) zu empfehlen.
Wer zwei oder mehr Tage bleibt, kann den nächsten Vormittag einer eingehenderen Beschreibung der inneren Stadt und der Museen, des Landhauses, Landes-Zeithauses und Joanneums widmen.

Both views are created ‘on-the-fly’ from the same XML file (=single source principle, more on that later...).

Practice!

Use <https://oxgarage.tei-c.org/> to transform some data to HTML. What do you notice? Do you like it or would you need to customize?

Text Encoding Initiative

TEI Primer

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>
```

Resources

- Learn TEI
- Teach Yourself
- P5 = 5. Proposal
- MEI for music
- CEI for charters
- <http://www.tei-c.org/>

TEI Header

fileDesc = bibliographical description of the contents of the document

encodingDesc = connection of electronic document to source (i.e. transcription rules, etc.)

```
<TEI> <!-- root element -->
  <teiHeader>
    <fileDesc> ... </fileDesc> <!-- obligatory -->
    <encodingDesc> <!-- optional -->
    <profileDesc> <!-- optional -->
    <revisionDesc> <!-- optional -->
  </teiHeader>
  <text> ... </text>
</TEI>
```

profileDesc = describes all non-bibliographical aspects of the text (i.e. creation, languages)

revisionDesc = tracks changes in the digital document

Using TEI i

Gentle Intro to XML

TEI Core

- **div** (division)
- **p** (paragraph)
- **head** (heading)
- **lb** (linebreak)
- **pb** (page break / beginning)
- **hi** (highlight)
- **l** (line)
- **lg** (line group)
- **list**

Attributes

- **@n** (label)
- **@type** (typing)
- **xml:id** (unique identifier)
- **xml:lang** (language)
- **@rend** (rendering)
- **@ana** (interpretation)

Using TEI ii

```
<foreign xml:lang="en">word</foreign>
<term type="homonym"/>
<date when="2009-04-27"/>
<time when="12:00:00"/>
<name type="person"/>
<persName n="Caesar" xml:id="#44BC">Caesaris</persName>
<!-- or -->
<persName key="ID.01.208"/>
<person/>
<emph/> <hi rend="italic">italic text</hi>
<seg/> <abbr type="acronym"/>
<placeName xml:id="#Whitby">Abbey</placeName>
```

Using TEI iii

Name spaces identified via URI

<prefix:name> e.g. `<tei:p>` ('I mean the `<p>` according to the TEI standard.')

declaration `<element xmlns="URI"> ...`

`<prefix:element xmlns:prefix="URI"> ...`

e.g.

`<tei:p xmlns:tei="http://www.tei-c.org/ns/1.0">...`

TEI is organized in modules

Acts of speech (reference) if speaker name is mentioned, otherwise TEIs 'said':

```
<sp who="#person">
    <speaker>1.</speaker> <p>Bla, bla, bla.</p>
</sp>

<said who="#Adolphe">- Alors, Albert, quoi de neuf?</said>
```

Letters in TEI (reference)

```
<div type="letter" n="14">
    <head>Letter XIV: Miss Clarissa Harlowe to Miss Howe</head>
    <opener>
        <dateline>Thursday evening, March 2.</dateline>
        <salute>Hallo,</salute>
    </opener>
    <p>On Hannah's depositing my long letter ...</p>
    <closer>
        <salute>Yours more than my own,</salute>
        <signed>Clarissa Harlowe</signed>
    </closer>
</div>
```

Names, Dates, Places

Named Entities & indirect reference

TEI 13: Names, Dates, People, Places

- persName for personal names, **<rs>** for *referring string* when mentioned indirectly ('he', 'the woman', etc.) → @key or @ref to specify who it is (reference).
- forename
- surname
- roleName (z.B. 'king')
- genName ('the Younger')
- addName
- nameLink ('von').

```
<name role="writer" type="person"
      ref="http://d-nb.info/gnd/118540238">
  Goethe</name>
<person>
  <addName type="Former">Murray</addName>
  <forename>Wilhelmina</forename>
  <addName type="nickname">Mina</addName>
</person>
```

Metadata in the TEI Header

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>
        <!-- title of the resource -->
      </title>
    </titleStmt>
    <publicationStmt>
      <p>
        <!-- Information about distribution of the resource -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
        <!-- Information about source from which the resource derives -->
      </p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Metadata in the TEI Header ii

The title and author in the `<titleStmt>` isn't the bibliographic data from the source! It describes the digital document and its authors or editors.

If you want to describe your source documents, you need elements like `<sourceDesc>` or `<msDesc>`:

```
<sourceDesc>
  <bibl>
    <title level="a">The Interesting story of the Children
      in the Wood</title>. In
    <author>Victor E Neuberg</author>, <title>The Penny Histories</title>.
    <publisher>OUP</publisher>
    <date>1968</date>. </bibl>
  </sourceDesc>
```

```
<sourceDesc>
  <p>Born digital: no previous source exists.</p>
</sourceDesc>
```

Metadata in the TEI Header iii

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Thomas Paine: Common sense, a
          machine-readable transcript</title>
      <respStmt>
        <resp>compiled by</resp>
        <name>Jon K Adams</name>
      </respStmt>
    </titleStmt>
    <publicationStmt>
      <distributor>Oxford Text Archive</distributor>
    </publicationStmt>
    <sourceDesc>
      <bibl>The complete writings of Thomas Paine, collected and edited
          by Phillip S. Foner (New York, Citadel Press, 1945)</bibl>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

<msDesc>

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno type="Bod">MS Poet. Rawl. D. 169.</idno>
  </msIdentifier>
  <msContents>
    <msItem>
      <author>Geoffrey Chaucer</author>
      <title>The Canterbury Tales</title>
    </msItem>
  </msContents>
  <physDesc>
    <objectDesc>
      <p>A parchment codex of 136 folios, measuring approx
         28 by 19 inches, and containing 24 quires.</p>
      <p>The pages are margined and ruled throughout.</p>
      <p>Four hands have been identified in the manuscript: the first 44
         folios being written in two cursive anglicana scripts, while the
         remainder is for the most part in a mixed secretary hand.</p>
    </objectDesc>
  </physDesc>
</msDesc>
```

<titlePage>

To describe a title page (e.g. early modern print copperplates, etc.),
use <titlePage>:

```
<titlePage>
<docTitle>
  <titlePart type="main">THOMAS OF Reading.</titlePart>
  <titlePart type="alt">OR, The sixe worthy yeomen of the West.</titlePart>
</docTitle>
<docEdition>Now the fourth time corrected and enlarged</docEdition>
<byline>By T.D.</byline>
<figure>
  <head>TP</head>
  <p>Thou shalt labor till thou returne to duste</p>
  <figDesc>Printers Ornament used by TP</figDesc>
</figure>
<docImprint>Printed at <name type="place">London</name> for <name>T.P.</name>
  <date>1612.</date>
</docImprint>
</titlePage>
```

<front>

You might also need **<front>** (front matter): contains any prefatory matter (headers, abstracts, title page, prefaces, dedications, etc.) found at the start of a document, before the main body.

```
<front>
  <epigraph>
    <quote>Nam Sibyllam quidem Cumis ego ipse oculis meis vidi in ampulla
      pendere, et cum illi pueri dicerent: <q xml:lang="grc">Σίβυλλα τι
        θέλεις</q>; respondebat illa: <q xml:lang="grc">ἀποθανεῖν θέλω.</q>
    </quote>
  </epigraph>
  <div type="dedication">
    <p>For Ezra Pound <q xml:lang="it">il miglior fabbro.</q>
    </p>
  </div>
</front>
```

Making the TEI your own i

How to find information on TEI elements

...and teach yourself how to use new elements:

- General TEI guidelines (XML Primer, Learn the TEI page, etc.)
- web-search TEI + (element you want to know about), i.e. "tei teiHeader" and you will get:
 1. definition page
 2. list of all examples for that element → directly over websearch or click 'show all' in the examples on the 'definitions page'
 3. sometimes even an module overview text for things as big as `<teiHeader>` (has its own module)

Relevant TEI modules

- all All modules
- 5 Characters, Glyphs, and Writing Modes,
- 10 Manuscript Description,
- 11 Representation of Primary Sources,
- 12 Critical Apparatus,
- 13 Names, Dates, People, and Places.

Also: The TEI guidelines are documentation and reference, not necessarily ideal teaching tools → overwhelming. Maybe try other tutorials like the TEI by example page, for example the tutorials on [Primary Sources](#) and [Critical Editing](#).

Making the TEI your own ii

Oxygen tricks

- If the TEI schema is linked to your document and you have internet, you can hover over elements and click to be redirected to the relevant info page.
- If you open a tag (by just typing '<'), the editor will suggest a list of elements currently allowed where you're standing (for example, <*teiHeader*> is very picky about the sequence).

TEI practice!

How can your bibliographical description be integrated into a
`<teiHeader>`?

Use websearch ('tei teiHeader') to learn how to use new elements
(overview plus examples view).

Read about the TEI Header Module.

Try TEI by example: The TEI Header

We will learn about `<msDesc>` later.

Transcriptions using Transkribus

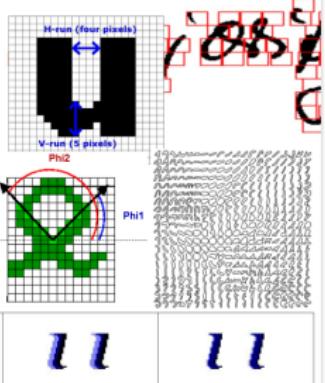
Transcription i

- OCR (Optical Character Recognition) – e.g. Transkribus (transcription support)
- Transkribus Keyword Spotting
- fuzzy search which should also find the word if it's mistranscribed
- Writer identification

converting images into text

measurable is e.g.

- density of pixels per area
- distance between edges
- angel between edges
- segments („Fraglets“)
- „automatic Overlap“
- ...



Processing steps

- Digitising
- Preprocessing
 - conversion into 2bit images
 - separation writing and background
 - edge detection
 - segmentation
- „feature“ extraction
- classification/clustering



$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

$$k \in \{ 'a' \dots 'z', '0' \dots '9' \}$$

Transcription ii

Typical phenomena

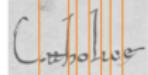
- “Special characters”
- Abbreviations
- damaged or unreadable text
- additions, deletions, substitutions, corrections
- editorial interventions (emendations and conjectures)
- editorial additions or omissions

Preprocessing: e.g. different segmentation methods

- Grid



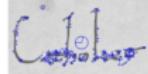
- Vertical Cuts / Seam Cuts



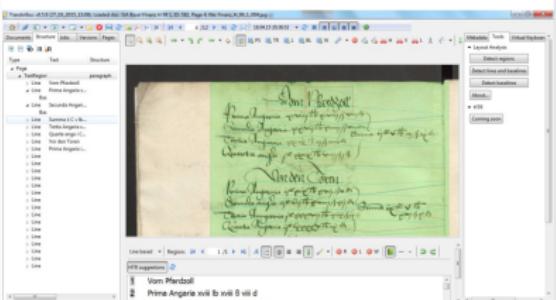
- Connected Components



- Keypoint-based



Transkribus



What is Transkribus?

Transkribus

“ ...is a comprehensive platform for the digitisation, AI-powered text recognition, transcription and searching of historical documents. (source) ”

- we're using the web version TranskribusLite:

<https://transkribus.eu/lite>

- for more complexity (which you might not need), download the software (Transkribus eXpert)
- Lite is easier to learn & has only the essentials.
- you need an account and to buy credits after you have used up your initial 200
- print & manuscript text recognition have different pricing
- there are stipends

Transkribus: Lite versus eXpert

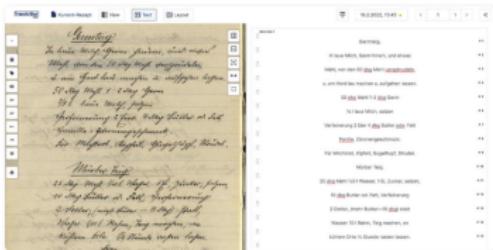


Transkribus ▾ ScanTent read&search The COOP ▾ Resources ▾ [Sign in](#) EN ▾

Transkribus Lite

[Try now](#) ↗

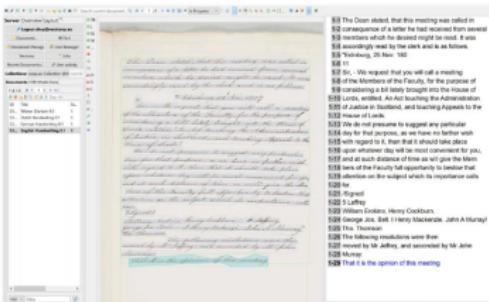
Transkribus Lite is the **browser version** of Transkribus. Automatically **transcribe**, comfortably **edit**, and easily **collaborate** on historical documents. And with Transkribus Lite you can train your own **AI models**. In your browser.



Transkribus eXpert

[Download](#) ↗

The Expert Client is the standalone version of Transkribus with the full power of the Transkribus platform: **digitisation**, **AI-powered recognition**, **transcription** and **searching** of historical documents.



Transkribus: How to guides

The screenshot shows a web browser with multiple tabs open, all related to transcription tools. The main content is the 'Transkrus' resources page from <https://readcoop.eu/transkrus/resources/>. The page has a header with the 'READ coop' logo and navigation links for 'Transkrus', 'ScanTent', 'read&search', 'The COOP', 'Resources', and 'Sign In'. A language selector 'EN' is also present. The main content area is divided into three columns: 'Getting started', 'How-To Guides', and 'Further Resources'. The 'Getting started' column includes links for 'Getting Started', 'Download and Installation', 'Uploading Documents', 'Layout Analysis', 'Handwritten Text Recognition', and 'Model Training', each with a corresponding icon. The 'How-To Guides' column features a large 'Transkrus in 10 Steps' icon, followed by 'How to transcribe Documents', 'How to Train & Apply Models', 'Transkrus Lite Introduction', 'How to use Structural Tagging', and a link to 'See all How-To Guides >'. The 'Further Resources' column includes 'Resource Center', 'Glossary', 'Events', 'Videos', 'Frequently asked Questions', 'Documentation for Developers', and an 'Insights Blog'. On the right side of the page, there is a sidebar with a 'Transkrus sation, ing of' heading, some descriptive text, and a small image of a document being transcribed. At the bottom, there is a footer with a 'Transkrus' logo and a copyright notice: '© 2023 READ Coop e.V. All rights reserved. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.' Below the footer, a series of small icons represent different file formats and types.

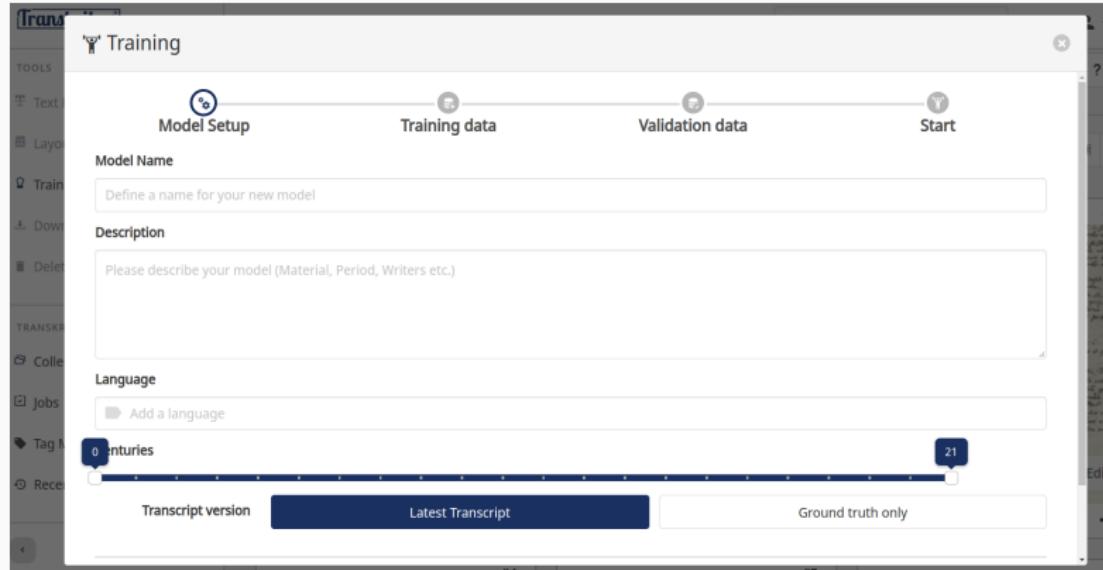
Transkribus: Creating transcriptions to train your own model

The screenshot shows the Transkribus interface. At the top, there's a logo for 'Transkribus lite', a menu item 'TRAINING_VALIDATION_SET_Alche...', and several buttons for 'View', 'Text', 'Layout', and a timestamp '17.3.2022, 22:42'. Below this is a photograph of a medieval manuscript page with Latin text written in two columns. To the right of the image is a transcription of the text. The transcription includes several small square icons with symbols: a plus sign, a minus sign, a double arrow, a single arrow pointing left, a double arrow pointing right, and a square with a diagonal line. The text in the transcription is as follows:

Impossible igitur est, aurum fieri posse, nisi ex materia
Auro proxima;
Quæ cum non sint metalla imperfecta,
nec proiectione alicuius pulvisculi, inde fiet aurum:
Quod si vero haec metalla imperfecta ac argentum vivum
vulgare redigantur in materiam auro proximam, tum
possibile est:

- first, transcribe a number of pages
 - recommended: 25–75
 - depending on print or handwritten
 - you can build on base models (should be similar)
 - you can speed up the process by creating a model, running it on new pages, correcting them and then repeating the process
 - correcting is usually still faster than transcribing from scratch
- be mindful to adhere to the transcription guidelines you want the model to learn
- train your model using the gold-standard transcriptions
- use the model on the pages you want transcribed
- there will probably be errors: fix them & use the extra training data

Transkribus: Configuring your model



Transkribus: further resources

- How to historical text recognition: A Transkribus Quickstart Guide, L^AT_EX-Ninja Blog, 10. November 2019, URL. → how to reuse existing models for print on the example of the Noscemus GM4
- Training my own Handwritten Text Recognition (HTR) model on Transkribus Lite, L^AT_EX-Ninja Blog, 22. March 2022, URL. → experiences training my own model

TEI for Digital Editing

TEI for Digital Editing i

TEI can describe the structure of a text, e.g.

- speaker, verse line, stage directives
- greeting, signature
- Visual aspects of the script
- special characters, new lines

Simple layout markup

- beginning of a new **line**: `<lb/>`
- beginning of a new **page**: `<pb/>` *@n* for an explicit numbering
- beginning of a new **column**: `<cb/>`
- **highlighted** text: `<hi>`
 - Attribute *@rend* to describe the appearance
 - Alternative encoding: , ,
- graphical elements in the text: `<figure>`
- `<fw>` (forme work) contains a running head (e.g. a header, footer), **catchword**, or similar material appearing on the current page.

```
<fw place="top-centre" type="head">Poëms.</fw>
<fw place="top-right" type="page-no">29</fw>
```

TEI for Digital Editing ii

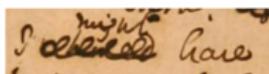
Documenting particularities of the writing surface

<damage> @agent, @degree,
@unit, @quantity,
@extent, @precision,
@scope

<unclear>

<gap> any omission in the transcription – @reason, e.g. sampling, inaudible, irrelevant, cancelled

e.g. unclear passage



```
<gap reason="wormhole" quantity="5" unit="character"/>  
<damage agent="coffee" quantity="3" unit="line"/>
```

Other important attributes

@cert(ainty) how certain you are about the suggested transcription?

@resp(onsibility) who did it?

@evidence where you got the clues from (internal, external, conjecture)?

```
I <subst>  
  <add place="above">might</add>  
  <del>  
    <unclear reason="overinking"  
          cert="medium" resp="#LDB">  
      should</unclear>  
    </del> </subst> have
```

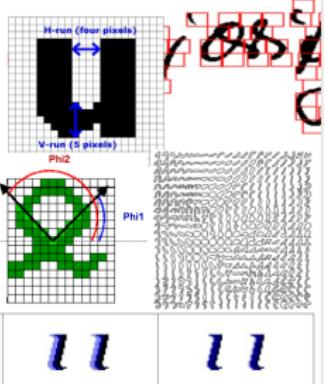
Transcription i

- OCR (Optical Character Recognition) – e.g. Transkribus (transcription support)
- also: Transkribus Keyword Spotting
- also: fuzzy search which should also find the word if it's mistranscribed
- Writer identification

converting images into text

measurable is e.g.

- density of pixels per area
- distance between edges
- angel between edges
- segments („Fraglets“)
- „automatic Overlap“
- ...



Processing steps

- Digitising
- Preprocessing
 - conversion into 2bit images
 - separation writing and background
 - edge detection
 - segmentation
- „feature“ extraction
- classification/clustering



$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

$$k \in \{ 'a' \dots 'z', '0' \dots '9' \}$$

Transcription ii

Typical phenomena

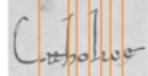
- “Special characters”
- Abbreviations
- damaged or unreadable text
- additions, deletions, substitutions, corrections
- editorial interventions (emendations and conjectures)
- editorial additions or omissions

Preprocessing: e.g. different segmentation methods

- Grid



- Vertical Cuts / Seam Cuts



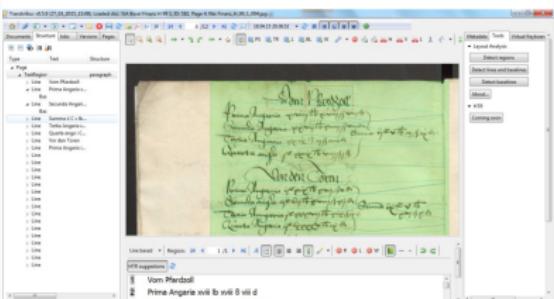
- Connected Components



- Keypoint-based



Transkribus



Transcription iii

Transcriptions contain:

- layout
- additions
- corrections
- modifications
- voids, space, holes, gaps ...
- alternative transcriptions
- editorial interventions
- Enhanced transcription

```
<pb>, <lb>, <cb>, <hi>, <g>,
<handShift/>
<add>, <addSpan>,
<corr>, <del>, <delSpan>, <sic>
<subst>
<gap>, <damage>
<choice>, <alt>
<unclear>, <supplied>, <reg>
<add>, <addSpan>, <corr>, <choice>,
<damage>, <del>, <delSpan>,
<restore>, <gap>, <sic>
```

Genetic Edition

Critique Génétique

Research interest: Reconstruct the writing process in the working manuscripts of the author

```
<zone>
<line>Alone
<seg type="alternative"
      xml:id="alt1"> before </seg>
<add place="above"
      type="alternative"
      xml:id="alt2">beside</add>
his native river-
</line>
<alt targets="#alt1 #alt2"
      mode="excl" weights="01"/>
</zone>
```

That's ~~xml:id="del1"~~ superfluous
<restore>
 <redo target="del2" />
 <del xml:id="del2">deleted
</restore>
text.

He sat <seg type="transposition"
xml:id="trans1"> at
his table</seg>
<seg type="transposition"
xml:id="trans2"> head on hands
</seg>.

```
<listTranspose>
<transpose>
  <ptr target="#trans2"/>
  <ptr target="#trans1"/>
</transpose>
</listTranspose>
```

The script: palaeography

- soft hyphen: *@break="no"*
- *@rend* / *@rendition*
 - *@rend*: verbal description; each word describes a single facet (*rend="indented:5cm"*)
 - *@rendition*: reference to description of the rendition in the *teiHeader//encodingProfile*
- „special characters“:

- *<g>*
- Does it exist in Unicode?
(<http://www.unicode.org>).
As an entity in XML:

```
      &x[hexadecimal code];  
      &#[decimal code];
```

- Search <https://unicode-table.com/>

Unicode for historical texts

- *Combining Diacritical Marks* (0300–036F) and *Supplement* (1DC0–1DFF): Superscripts, Subscripts
- *Latin Extended Additional* (1E00–1EFF): characters with diacritics
- *Latin Extended-D* (A720–A7FF): Ligatures, abbreviations, ...
- *General Punctuation* (2000–206F) and *Supplement* (2E00–2E7F)
- *Ancient Symbols* (10190–101CF): roman measurements, coins...
- Search unicode entities:
<https://unicode-table.com/>

Editorial interventions

- expansion of abbreviations
- Conjectures
- Normalisations
- *<abbr>, <expan>* plus
<am>, <ex>
- *<sic>, <corr>*
- *<orig>, <reg>*

All these can be paired:

- General for all editorial interventions: *<choice>*.
- Explicitly for substitutions: *<subst>*.
- *<supplied>* for additions by the editor
- *<unclear>* for unreadable text (*@reason, @agent, @hand*)

Abbreviations i

In Western MS, we usually distinguish:

- **Suspensions:** the first letter or letters of the word are written, generally followed by a point : for example 'e.g.' for 'exempla gratia'
- **Contractions:** both first and last letters are written, generally with some mark of abbreviation such as superscript strokes, or points : e.g. 'Mr.' for 'Mister'
- **Brevigraphs:** Special signs such as the Tironian nota used for 'et', the letter p with a barred tail used for 'per', the letter c with a circumflex used for 'cum'/'con' etc.
- **Superscripts:** Superscript letters (vowels or consonants) used to indicate various kinds of contraction: e.g. 'w' followed by superscript 'ch' for 'which'.

- **<expn>** The element content is considered as the expansion of an abbreviation. In the text: USA → transcription:

```
<expn>United States of America</expn>
```

- **<abbr>** The element content is an abbreviation

```
<abbr>USA</abbr>
```

- **<ex>** (expansion) and **<am>** (abbreviation mark) for the omitted part of the abbreviation, e.g.

```
e<ex>xempla</ex> g<ex>ratia</ex>  
e<am>.</am> g<am>.</am>
```

Abbreviations ii

Abbreviations can also be considered as alternatives: *<choice>*, e.g. 'Zum Beispiel' and 'z.B.':

```
<choice>
  <expan>Zum Beispiel</expan>
  <abbr>Z.B.</abbr>
</choice>
```

Or respectively:

```
Z<choice>
  <am>.</am>
  <ex>um</ex>
</choice>

B<choice>
  <am>.</am>
  <ex>eispiel</ex>
</choice>
```

Transcription = Interpretation

'geminination dash' – possible solutions:

- Uncommented expansion
- Unicode m with "combining macron"
- Encoding as an XML-Entity
- <g> referring to <charDecl>
- Only <am/> for the stroke
- Only <ex> for the expansion
- As a <choice> with <abbr> and <expan>, the first incl. a abbreviation mark <am> and the second the expansion <ex>

Geminimation dash

horizontal, bended or curved stroke above a nasal letter indicating the omission of a further instance of the same letter. (source)

Modifications

- addition, deletion, substitution, transpose, or:
- modification (represents any kind of general modification without interpretation)

Changing writer

<handShift /> @new : the hand which writes from this place onward

<handDesc> (part of *msDesc*)

<handNotes> (part of *profileDesc*)

<handNote> for a particular description

@xml:id an identifier for the hand

Text and images i

Images of a text are encoded in a *facsimile* – structure parallel to *teiHeader* and *text*:

```
<tei>
  <teiHeader>...</teiHeader>
  <facsimile> ...</facsimile>
  <text>...</text>
</tei>
```

<facsimile>

- <*surface*> = something meant to be seen
 - *@uly*, *@ulx*; *@lrx*, *@lry* =upper left x/y- and lower right y/x coordinates
 - coordinates form a grid, which can be referred → *@ulx* and *@uly* are usually 0
- <*graphic*>: image, *@url* : image file
- <*zone*> = an area on the surface.
Coordinates refer to the grid defined in *@uly*, *@ulx*; *@lrx*, *@lry* of the <*surface*>.

Text and images ii

Example

surface

zone

@ulx, @uly

@lrx, @lry

graphic = <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/Images/faes-fig1.png>



<facsimile>

<surface

ulx="0" uly="0" lrx="200" lry="300">

<graphic url="Bovelles-49i.png" />

<zone

ulx="25" uly="25" lrx="180" lry="60">

</zone>

<zone

ulx="28" uly="75" lrx="175" lry="178">

</zone>

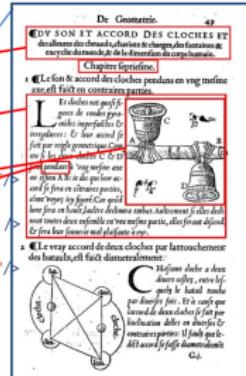
ulx="105" uly="76" lrx="175" lry="160">

</zone>

ulx="45" uly="125" lrx="60" lry="130"/>

</surface>

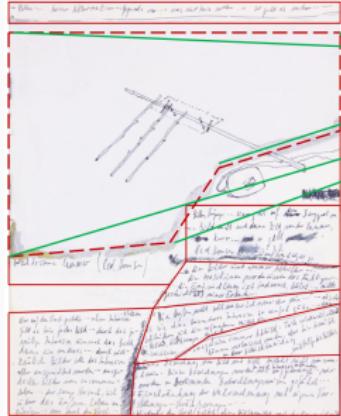
</facsimile>



<zone>

@points:

List of coordinates (pairs of numbers), which combined by lines enclose a region on the surface.



<zone>

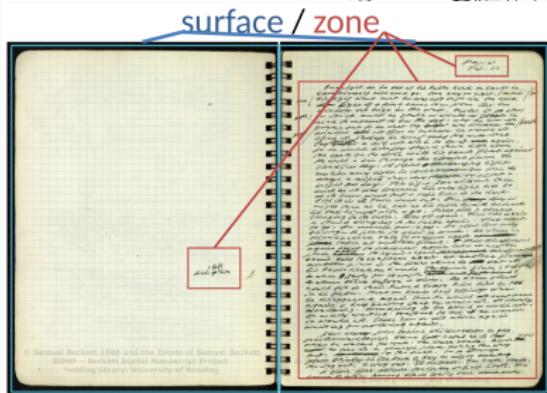
points="0,29
534,20 536,215
334,282 259,376
0,409"/>

Text and images iii

Linking text and image

- @facs, content corresponding with a `xml:id` in the facsimile structure:

```
<surface xml:id="p49">  
  <zone xml:id="p49z2" />  
  <graphic url="test.png" />  
</surface>  
  
<text><body><div>  
  <pb n="49" facs="#p49"/>...  
  <head facs="#p49z2">  
    Chapitre septiesme </head>  
</div></body></text>
```



Tools for text-image linking

1. Image markup tool (Martin Holmes, http://www.tapor.uvic.ca/~mholmes/image_markup/index.php)
2. TextGridLab: <http://www.textgridlab.de>
3. T-PEN (<http://www.t-pen.org>)
4. <http://imagecoordinates.com>

Text and images iv

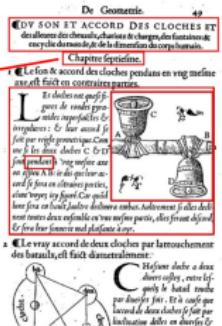
Embedded transcription

- "Embedded transcription": Text directly in `<surface>`
- *Relevant elements:*
`<sourceDoc>`, `<surface>`,
`<zone>`, `<line>`, `@rotate`

```
<sourceDoc>
  <surfaceGrp n="leaf1">
    <surface facs="page1.png"> <zone>All the writing on page 1</zone> </surface>
    <surface>
      <graphic url="page2-highRes.png"/>
      <zone> <line>A line of writing on page 2</line> </zone>
    </surface>
  </surfaceGrp>
</sourceDoc>
```

Embedded transcription

```
<sourceDoc ulx="0" uly="0"
  lrx="..." ...>
  <surface>
    <zone ulx=".." ...>
      <line>Chapitre  
septiesme</line>
    </zone>
    <graphic url="test.png" />
  </surface>
  </facsimile>
```



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

Critical Apparatus ii

The apparatus can be located anywhere as a *<listApp>*:

- in the *<body>* of the document
- in the *<back>* in other documents
- is referenced by *@loc*

→ *<rdgGrp>* aggregates several readings of a common type.

Witness list

@wit refers to descriptions in the header: *<listWit>* = list of *<witness>*-elements identified by *@xml:id* each describing a textual witness e.g. by *<bibl>*- or *<msDesc>* elements in *teiHeader/sourceDesc*

```
<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>
```

Critical Apparatus iii

There are different options:

Location referenced + external

```
<p><pb n="f13"/><lb n="f13-z1" />  
Dis ouentürlich buoch bewiset wye  
von einer Frowen ge<lb/>nannt  
Melusina ...</p>  
<!-- ... -->  
<listApp><app loc="f13-z1">  
    <lem wit="#BR1">ouentürlich</lem>  
    <rdg wit="#SK1">ouentuorlich</rdg>  
    <rdg wit="#AS1">abenteürlich</rdg>  
</app></listApp>
```

Double endpoint + external

The apparatus is integrated into the `<body>` linked to identifiers for its beginning and its end (e.g. with `<anchor>`) but encoded anywhere (e.g. in the place where it was located in the printed source); referenced by `@from` & `@to`.

```
<p>Dis <anchor xml:id="A1"/>  
ouentürlich<anchor xml:id="A2"/> buoch  
bewiset wye von einer  
Frowen ge<lb/>nannt Melusina  
...</p>  
<!-- ... -->  
<app from="#A1" to="#A2">  
    <rdg wit="#SK1">ouentuorlich</rdg>  
    <rdg wit="#AS1">abenteürlich</rdg>  
</app>
```

Critical Apparatus iv

Location referenced + inline

```
<p n="p1">
  Dis ouentúrlich
  <app loc="p1">
    <rdg wit="#SK1">
      ouentuorlich</rdg>
    <rdg wit="#AS1">
      abenteürlich</rdg>
  </app>
  buoch bewiset wye
  von einer Frowen
  ge<lb/>nannt Melusina ...
</p>
```

Double endpoint + internal

The apparatus is integrated into the `<body>` after the referenced passage and linked an identifiers for its beginning; referenced by `@from`.

```
<p n="1">
  Dis <anchor xml:id="a"/>
  ouentúrlich<app from="#a">
    <rdg wit="#SK1">
      ouentuorlich</rdg>
    <rdg wit="#AS1">
      abenteürlich</rdg>
  </app>
  buoch bewiset wye von einer
  Frowen ge<lb/>nannt Melusina
  ...
</p>
```

Critical Apparatus v

Last but not least...

Parallel segmentation

Encoding the „base text“ as the *lem* in the *app*-element. Can be done only inline. Possibility to nest variants.

```
<p n="1">Dis
  <app>
    <lem wit="#BR1">
      ouentúrlich</lem>
    <rdg wit="#SK1">
      ouentuorlich</rdg>
    <rdg wit="#AS1">
      abenteürlich</rdg>
  </app>
  buoch bewiset wye von einer
  Frowen ge<lb/>nannt Melusina
  ...
</p>
```

Which one to choose?

1. **Referenced** imitates the classical print version, is relatively fast to create but can be imprecise in referencing.
2. **Double-Endpoint** is relatively complex to encode and to process, but exact and the only form to handle overlapping structures.
3. **Parallel Segmentation** can be easily processed with XSLT but not very flexible in documenting complex changes and overlapping structures.

Further suggestions

Suggestions for typical problems in editing

Missing text (om.) `<rdg>` remains empty; `@cause` can contain a controlled term to describe the situation (e.g. omisit)

Additions (add.) `<lem>` remains empty

Corrections (corr. ex ...) `<rdg>` contains the complete encoding

`<subst>...<add>...</add></subst>`

Tools for collation

Juxta Commons Texts are reduced to flat text. Variants are encoded in the parallel segmentation method.

CollateX Creates a graph. Compares every version with the existing graph and searches for gaps.

Stemmatology in TEI

`<eTree>` each part of the tree which can have descendants

`<eLeaf>` each part of the tree, which has only ancestors

`@type` e.g. hypothetical, extant, lost ...

`<label>` for the short names („Sigla“)

`<ptr>` for „contaminations“ i.e. texts influenced by other manuscript traditions

TEI Critical Apparatus Toolbox

- <http://teicat.huma-num.fr/>
- by Marjorie Burghart
- Check encoding: consistency etc.
- Display parallel versions.
- Print an edition of a TEI XML edition, with a TEI-to-L^AT_EX and PDF transformation (*reledmac!* → XSL is here).
- Annotate images: lets you easily trace zones on an image to prepare a documentary edition (sometimes kind of buggy) → create your *<facsimile>*.
- Get statistics on the XML tags used in different parts of your edition plus word counts.



The screenshot shows the homepage of the TEI Critical Apparatus Toolbox. At the top left is a logo featuring a white cat with a yellow 'TEI' ribbon around its neck. To the right of the logo, the text 'Critical Apparatus Toolbox' is written in a dark, sans-serif font. The main title 'TEI Critical Apparatus Toolbox' is centered below the logo in a large, bold, dark font. At the bottom of the page is a navigation bar with several links: 'Home', 'Check your encoding', 'Display parallel versions', 'Print an edition', 'Annotate an image', 'Get statistics', 'Help', 'Download', and 'Credits'. The 'Home' link is currently highlighted with a light gray background.

TEI practice!

Fill out the `teiHeader` and encode a `<titlePage>`

Use websearch ('tei titlePage') to learn how to use new elements
(overview plus examples view).

References

- [1] Gaston Bachelard. *Le nouvel esprit scientifique*. 10th ed. (Original work published 1934). Paris: Les Presses universitaires de France, 1968.
- [2] Johanna Drucker. "Humanities Approaches to Graphical Display". In: *Digital Humanities Quarterly* 5/1 (2011).
- [3] Sarah Lang. "What you really need to know about Digital Scholarly Editing". In: *The L^AT_EX Ninja Blog* (2022). URL:
<https://latex-ninja.com/2022/10/30/what-you-really-need-to-know-about-digital-scholarly-editing/>.
- [4] Patrick Sahle. "Digitales Archiv und Digitale Edition. Anmerkungen zur Begriffsklärung". In: *Literatur und Literaturwissenschaft auf dem Weg zu den neuen Medien*. Ed. by Michael Stolz. Zürich, 2007, pp. 64–84.
- [5] Patrick Sahle. "What is a Scholarly Digital Edition?" In: *Digital Scholarly Editing: Theories and Practices*. Ed. by Matthew J. Driscoll and Elena Pierazzo. Cambridge: Open Book Publisher, 2016, pp. 19–39. URL:
<https://books.openedition.org/obp/3381>.

- [6] Herbert Stachowiak. *Allgemeine Modelltheorie*. Wien, 1973.