



Modelling humanities data with TEI-XML

SCHOLARLY EDITING AND MANUSCRIPT CATALOGUING IN THE DIGITAL AGE

Dr Katarzyna Anna Kapitan
23 October 2024

Recap: TEI (basics)

```
<TEI xmlns='http://www.tei-c.org/ns/1.0'>  
  <teiHeader>  
    metadata  
  </teiHeader>  
  <text>  
    textual content  
  </text>  
</TEI>
```

Recap: TEI (basics)

► Week 3, Exercise 2 (Homework)

- ▶ In pairs, compare the teiHeaders of your files.
- ▶ Did you encode the same types of information?
- ▶ What are the main similarities and differences?

Recap: TEI (basics)

► Week 3, Exercise 2 (Homework)

- Compare your file with:
 - Ex2_FirstFolio_spoiler_KAK.xml
 - Ex2_FirstFolio_spoiler_Bod.xml
- What are the main similarities and differences?

<TEI>: How it works

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title> </title>
      </titleStmt>
      <publicationStmt>
        <p> </p>
      </publicationStmt>
      <sourceDesc>
        <msDesc> </msDesc>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <p> <p/>
    </body>
  </text>
</TEI>
```

<TEI>: How it *also* works

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      [...]
      <sourceDesc> </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <listBibl>
        <msDesc> </msDesc>
      </listBibl>
    </body>
  </text>
</TEI>
```

Cataloguing Manuscripts with TEI XML

Kapitan, Modelling humanities data with TEI-XML

Introduction to Cataloguing

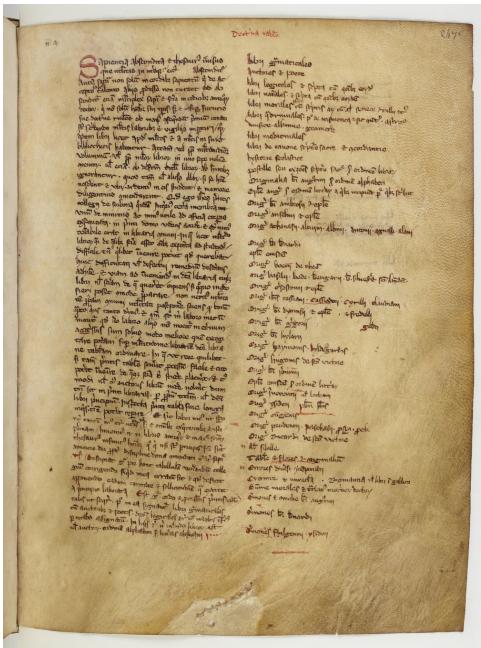


Clay tablet from the Nippur collection of the University of Pennsylvania Museum, 29-15-155.

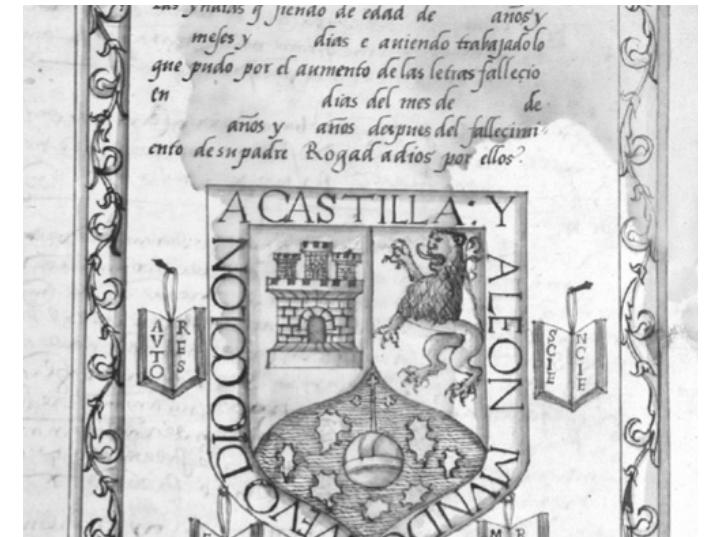
Source:

<https://www.penn.museum/collections/object/521562>

Kapitan, Modelling humanities data with TEI-XM



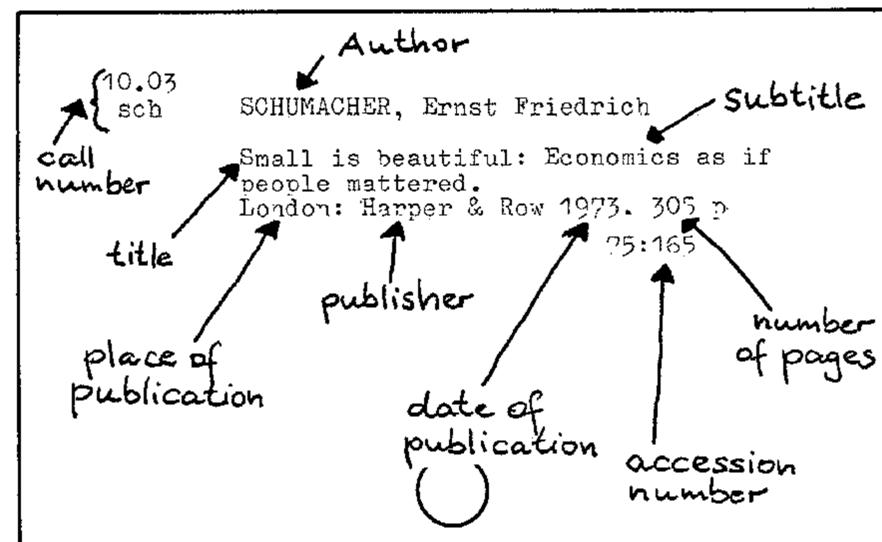
Paris, Bibliothèque nationale de France , NAL 99, p. 259
Source: <https://gallica.bnf.fr/>



Hernando Colón's coat of arms featuring the four inventories of his library:
Autores, Sciencie, Materee, Epitome.
Source: Archivo Histórico Provincial de Sevilla.

Introduction to Cataloguing

- ▶ Transferring information about book/document to a card according to set rules
(The New Zealand Digital Library)
- ▶ The function of the catalogue is to **identify all the items in a collection and to group like items together**
(Encyclopedia Britanica)



Main entry card/unit card/identity card.
Source: The New Zealand Digital Library, <http://www.nzdl.org/>

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019

Introduction to Cataloguing

by repository / collection

by subject

Script

Language

Theme

Author

Text

Period...



Copenhagen, Den Arnamagnæanske Samling, AM 71 8vo.
Photo: Suzanne Reitz.

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019

Kapitan, Modelling humanities data with TEI-XML

Introduction to Cataloguing

by repository / collection:

- Guillaume de Villefroy, *Catalogus codicum manuscriptorum Bibliothecae Regiae* (Paris, 1739–1744).
- Falconer Madan, *List of Additions to the Manuscripts in the British Museum in the Years 1841-1845* (London, 1850)
- Kristian Kålund, *Katalog over den Arnamagnæanske Håndskriftsamling* (Copenhagen, 1889–1894)

Introduction to Cataloguing

by language:

- N. R. Ker, Catalogue of Manuscripts Containing **Anglo-Saxon** (Oxford, 1957)

by subject/theme:

- Dorothea Waley, Catalogue of **Latin and Vernacular Alchemical Manuscripts** in Great Britain and Ireland, Dating from before the XVI Century, 3 vols. (Brussels, 1928-31).
- Dieter Studer-Joho, A Catalogue of Manuscripts Known to Contain Old **English Dry-Point** Glosses (Tübingen, 2017)

by author:

- Georgius Lacombe et al., **Aristoteles latinus.** Codices, Corpus philosophorum Medii Aevi (Bruges, 1955)

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
Kapitan, Modelling humanities data with TEI-XML

TABLE OF CONTENTS.

CLASSICAL ROMANCES.

Cycle of Troy :—

							PAGE
Heroica, by Flavius Philostratus	1
Iliaca, by Joannes Tzetzes	2
Dictys Cretensis	9
Dares Phrygius	12
Latin Poems, by Simon Chèvre d'Or, etc.	27
Roman de Troie, by Benoit de Sainte-More	35
Historia Trojana, by Guido delle Colonne	40
Trójumanna Saga	62
Rímur af Trójumönnum, by Jón Jónasson	63
Recueil, by Raoul Le Fèvre	64
Latin History of Troy	66
Filostrato, by Boccaccio	67
Troilus and Chryseyde, by Chaucer	70
Troy Book, by Lydgate	75
Roman d'Eneas	82
Romance of Troy	84
Romance of Thebes, by Lydgate	87
Romance of Jason, from Raoul Le Fèvre (Dutch)	92

Additional 4869. ff. 2-66.

Paper; A.D. 1679. Folio; ff. 65, having 37 to 39 lines to a page. With initials in red and green. Forming the first of five sagas, written by Jón Thóðarson, at the expense of Magnús Jónasson of Vigur, an island in Isafjörður, in the north-western peninsula of Iceland, in the years 1679, 1680. The whole volume contains:

- | | |
|-------------------------------|-------------------------------|
| 1. The present article. f. 2. | 4. Octaviani saga. f. 161. |
| 2. Magus Jarls saga. f. 67. | 5. Addonii saga. ff. 206-235. |
| 3. Orkneyings saga. f. 102. | |

TRÓJUMANNA SAGA: translated from a Danish abridgment of the Historia Trojana of Guido delle Colonne. Divided by the ornamental initials into 22 sections. *Icelandic*

Heading: ‘

til Enda og so h
unninn og innat
aar þá Abdon v
hefur kongur h

Additional 11,113. ff. 171-281.

Paper; xviiith cent. Folio; ff. 111, each page having from 17 to 21 lines, in which the verses are written like prose.

The whole volume contains:

- | | |
|--|--------------------------------------|
| 1. Vemundar saga og Víga-Skutu.
f. 1. | 3. Brandkrossa pátr. f. 84. |
| 2. Gunnars pátr Þórðarhábana.
f. 71. | 4. Gull-Póris saga. f. 92. |
| | 5. Kormaks saga. f. 127. |
| | 6. The present article. ff. 171-281. |

LAYS UPON THE STORY OF TROY: versified from the Trójumanna Saga by Jón Jónasson of Berunes in East Iceland. In 30 Rímur (lays). Imperfect at the end. *Icelandic*.

Title: “Rymur af Troio-Mónnum kveðnar af Jone Jons syne a Berunese.” Begins:

“Froðar meistarar fyrr um heim
foru og löndin vyða,
liðugt neitte Laungum þeim
Lioða vess að smyða.”

Ward H.L.D., Catalogue of **Romances** in the Department of Manuscripts
in the **British Museum**, vol. 1 (London 1883)

Introduction to Cataloguing

- ▶ Digital Catalogues
 - ▶ Repository Based Catalogues
 - ▶ National Projects
 - ▶ Smaller/Individual Projects

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019

Kapitan, [Modelling humanities data with TEI-XML](#)

Repository Based Digital Catalogues

The image displays two side-by-side screenshots of digital library websites. On the left is the website for the Digital Library of the Biblioteca Apostolica Vaticana (DigiVatLib). The header features the logo 'DVL DIGIVATLIB' and links for 'HOME', 'DIGITAL COLLECTIONS', 'ONLINE CATALOGUES', and 'ABOUT'. Below the header is a large image of a medieval manuscript page showing a scene with figures and musical notation. To the right of the image is a search interface titled 'Vatican Library Search and discover collections' with a search bar and a 'SEARCH' button. Below the search interface are two more manuscript images with labels 'Capp.Sist.160'. A section titled 'Selected Manuscripts' shows four thumbnail images of manuscripts. On the right is the website for Digital Bodleian. The header features the logo 'DIGITAL BODLEIAN' and a link to 'https://digital.bodleian.ox.ac.uk'. Below the header is a large image of a medieval manuscript page featuring a detailed illustration of various animals and figures. Text below the image reads 'DIGITAL.BODLEIAN MAKES THESE EXTRAORDINARY LIBRARY COLLECTIONS AVAILABLE ONLINE FOR THE VERY FIRST TIME...'. There is a search bar with the placeholder 'Enter new search term' and a 'SEE HOW IT WORKS' button. Below the search bar is a section titled 'THE COLLECTIONS' with a 'Filter collections' dropdown and two thumbnail images of manuscripts.

**The Digital Library of the Biblioteca Apostolica Vaticana
(DigiVatLib)**
<https://digi.vatlib.it>

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
Kapitan, Modelling humanities data with TEI-XML

Digital Bodleian
<http://digital.bodleian.ox.ac.uk/>

National Projects

The image shows two side-by-side screenshots of digital manuscript libraries. On the left is the BnF Gallica website (<https://gallica.bnf.fr/accueil/en/content/accueil-en?mode=desktop>). The header features the BnF logo and 'Gallica'. It has sections for 'NEWS' (from the digital library, 5 001 350 documents online), 'THE MOOD' (with a small image of a manuscript page), 'THE WEATHER' (with a small image of a manuscript page), and 'Discover 800 medieval manuscripts from France and England' (with a small image of a manuscript page). On the right is the e-codices website (<https://www.e-codices.ch/en>). The header features the e-codices logo and navigation links for Person Index, Annotations, About, Browse & Search, and language selection (de, en, fr, it). The main content area shows a large image of a medieval manuscript page with decorative initials and text, followed by text about the e-codices project: 'e-codices - Virtual Manuscript Library of Switzerland', 'The goal of e-codices is to provide free access to all medieval and a selection of modern manuscripts of Switzerland by means of a virtual library.', and 'At the moment, the virtual library contains 2263 manuscripts from 92 different collections. The virtual library will be continuously updated and extended.' Below this are sections for 'Collections' (Swiss Collections, Location, Library / Collection), 'Documents', and 'News' (Latest Manuscripts: 06/13/2019).

Gallica (FR) Digital Catalogue of Bibliothèque nationale de France & partner institutions
<https://gallica.bnf.fr>

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
Kapitan, Modelling humanities data with TEI-XML

Smaller Projects

The screenshot shows a dark-themed website for a research project. At the top, a navigation bar includes links for 'home', 'texts', 'manuscripts', 'bibliography', 'from the editors', and 'about the project'. Below this is a green header bar with the text 'Home'. The main content area features the title 'STORIES FOR ALL TIME: THE ICELANDIC FORNALDARSÖGUR' in large white letters. Underneath, a section titled 'Stories for all time' contains a detailed text about the project's aims and funding, mentioning the Arnamagnæan Institute and the Velux Foundation. A small image of a medieval manuscript page is visible on the left.

Stories for all time, Catalogues of Fornaldarsögur Norðurlanda
<http://fasnl.net/>

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
Kapitan, Modelling humanities data with TEI-XML

The screenshot shows a website for the Virtual Monastic Library of Lorsch. The header features the text 'Bibliotheca Laureshamensis digital' and 'Virtual Monastic Library of Lorsch' with a 'codex' logo. To the right is a large image of a medieval manuscript showing several figures in a scriptorium or library setting. The sidebar on the left contains links for 'Home', 'Monastic Library of Lorsch' (with sub-links for 'Library and Scriptorium', 'Selected Manuscripts', 'Manuscript Illumination', 'Literature', 'Virtual Library', 'Search', 'Manuscripts', and 'Decoration'), 'Image Gallery', 'Glossary', 'Press Coverage', 'Terms of Use', 'Project Partners', and 'Project Description'. The main content area has a heading 'Bibliotheca Laureshamensis – digital: Virtual Monastic Library of Lorsch' and a detailed text about the project's aim to virtually reconstruct the monastic library of UNESCO-World Heritage Site Abbey Lorsch. It mentions the reconstruction of more than 300 medieval manuscripts and their distribution worldwide. The sidebar on the right lists recent blog posts in German:

- 14.11.2016: [Beschreibungen der Handschriften nun auch im Druck erschienen](#) — 0 Kommentare
- 29.08.2016: [Tafelwerk zur Lorscher Schriftgeschichte kürzlich erschienen](#) — 0 Kommentare
- 22.01.2016: [Vortrag zur Lorscher Torhalle](#) — 0 Kommentare
- 02.12.2015: [Letzte Lorscher Handschrift online](#)

Virtual Monastic Library of Lorsch
<http://bibliotheca-laureshamensis-digital.de/>

Digital & Digitised Catalogues

AM 415 fol.

There are currently no images available for this manuscript.

HEIMSKRINGLA — NOREGS KONUNGA SÖGUR; DANMARK?, 1700-1799

CONTENTS

Heimskringla — Noregs konunga sögur

Note En afskrift, forsynet med varianter og forberedt til trykken, af den islandske tekst fra begyndelsen til og med kong Magnús Erlingsson

Language of Text Icelandic

PHYSICAL DESCRIPTION

Support Papir.

No. of leaves 630. Bl. 171-72 og 366-67 er ubeskrevne. 335 mm x 214 mm.

HISTORY

Origin Har tilhørt Det Kongelige Biblioteks Addit. 11 fol. og Addit. 15 b fol.

ADDITIONAL

573. AM. 415, fol. (forhen Addit 11 og 15 b, fol.). Pap. 33,5 × 21,4 cm. 630 bl. 18. årh. Bl. 171—72 og 366—67 ubeskrevne.

Heimskringla eðr Noregs konunga sögur.

En med varianter forsynet og til trykken forberedt afskrift af den islandske tekst forfra til kong Magnús Erlingsson (incl.) — svarende til folioudgaven af Heimskringla tom. I—III.

K. Kålund, *Katalog over den Arnamagnæanske Håndskriftsamling* (Copenhagen 1889), vol. 1, p. 313.

Handrit (2019) <https://handrit.is/en/manuscript/view/da/AM02-0415>

Den Arnamagnæanske Samling, AM 415 fol.
Digital description from 1889

Digital & Digitised Catalogues

ALL latin 7906 ADVANCED SEARCH | SELECTIONS | BLOG

Full notice

Title : 1.^o Publili Terentii , Afri, comoediae sex ; decimo tertio saeculo exaratae. – 2.^o D. Juvenalis satyrarum quarta, cuius initium desideratur, quinta, sexta, septima et pars octavae ; decimo saeculo exaratae. – 3.^o Publili Virgilii Maronis aeneidos libri tertii finis, quartus et quintus integri ; nono saeculo exarati. – 4.^o Daretis Phrygii de excidio Trojano historia : interprete Cornelio Nepote , cuius praemittitur epistola ad Crispum Sallustium ; ea pars codicis saeculo eodem exarata. – 5.^o Gregorii , Episcopi Turonensis, historiae Francorum liber primus ; eodem saeculo exaratus : finis desideratur. – 6.^o Anonymi carmen cuius titulus : disputatio mundi et religionis ; decimo quarto saeculo exarata.

Publication date : 201-0900

Type : manuscript

Type : manuscrit

Language : latin

Format : Parchemin

Description : Content : F. 59-88. Vergilius, Aeneis

Description : Colbertinus

Description : Lieu de copie : Allemagne ouest

Rights : domaine public

Rights : public domain

Identifier : ark:/12148/btv1b85721963

Source : Bibliothèque nationale de France, Département des manuscrits, **Latin 7906**

Relationship : <http://archivesetmanuscrits.bnf.fr/ark:/12148/cc67162m>

Provenance : Bibliothèque nationale de France

Date of online availability : 03/06/2013

(BnF)

Gallica (2019):
<https://gallica.bnf.fr/ark:/12148/btv1b85721963.r=Latin%207906?rk=21459;2>

vii M C M V I.

Codex membranaceus, olim Colbertinus.

Ibi continentur :

1.^o *Publili Terentii*, Afri, comœdiæ sex ; decimo tertio saeculo exaratae.

2.^o *D. Juvenalis* satyrarum quarta, cuius initium desideratur, quinta, sexta, septima & pars octavae ; decimo saeculo exaratae.

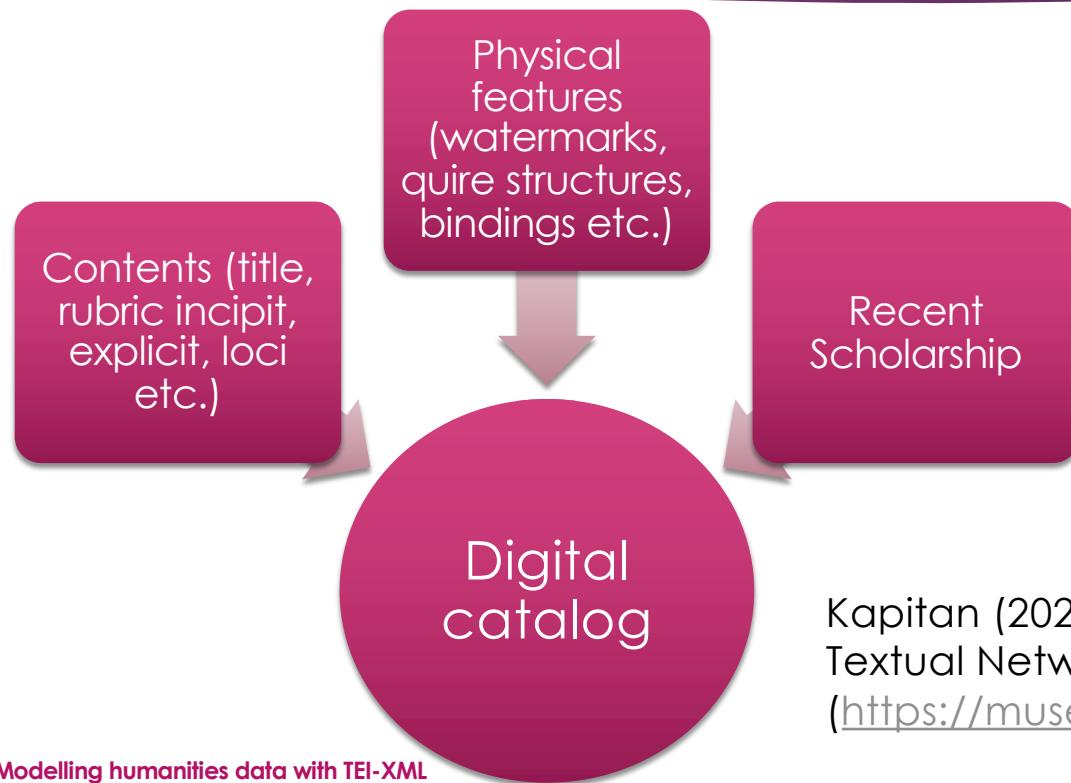
3.^o *Publili Virgilii Maronis* aeneidos libri tertii finis, quartus & quintus integri ; nono saeculo exarati.

4.^o *Daretis Phrygii* de excidio Trojano historia : interprete *Cornelio Nepote*, cuius Fff ii

Guillaume de Villefroy, *Catalogus codicum manuscriptorum Bibliothecae Regiae* (Paris, 1744), vol. 4, p. 411.

Bibliothèque nationale de France, Lat. 7906
Digital description from 1744

Digital & Digitised Catalogues



Kapitan (2021). Perspectives on Digital Catalogs and Textual Networks of Old Norse Literature
(<https://muse.jhu.edu/article/794855>)

TEI Manuscript Description Module

- ▶ The **TEI manuscript description module** provides tools for a detailed description of manuscripts, including:
 - the current holding institution
 - the intellectual content
 - codicological and palaeographical features
 - text layout and illumination/decoration
 - origin and provenance
 - primary and secondary bibliographical references

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
Kapitan, Modelling humanities data with TEI-XML

TEI Manuscript Description Module

msdescription

The screenshot shows the TEI Guidelines website with a dark blue header. The header includes the TEI logo, a navigation bar with links to Guidelines, Activities, Tools, Membership, Support, About, and News, and a search bar. The main content area has a white background. At the top of the content area, it says "TEI: Guidelines for Electronic Text Encoding and Interchange" and "P5 Version 4.8.0. Last updated on 2nd September 2024, revision 4e6e41b0b". On the left, there is a sidebar with a "Table of contents" section listing various chapters under "11 Manuscript Description". Below this are links to "10 Dictionaries" and "12 Representation of Primary Sources", followed by a "Home" link. The main content area starts with "11 Manuscript Description" and "11.1 Overview". It contains text about the msdescription module and its purpose. There is also a note about the diversity of needs and a link to "11.2 The Manuscript Description Element". A sidebar on the right contains links to "11.1 Overview", "11.3 Phrase-level Elements", and "Home".

TEI Guidelines:

<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>

TEI Manuscript Description Module

msdescription

- ▶ The **TEI manuscript description module (msdescription)** is intended for several different kinds of applications:
 - electronic catalogue record or finding aid
 - metadata component within a digital edition
 - tool for quantitative codicology

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019
[Kapitan, Modelling humanities data with TEI-XML](#)

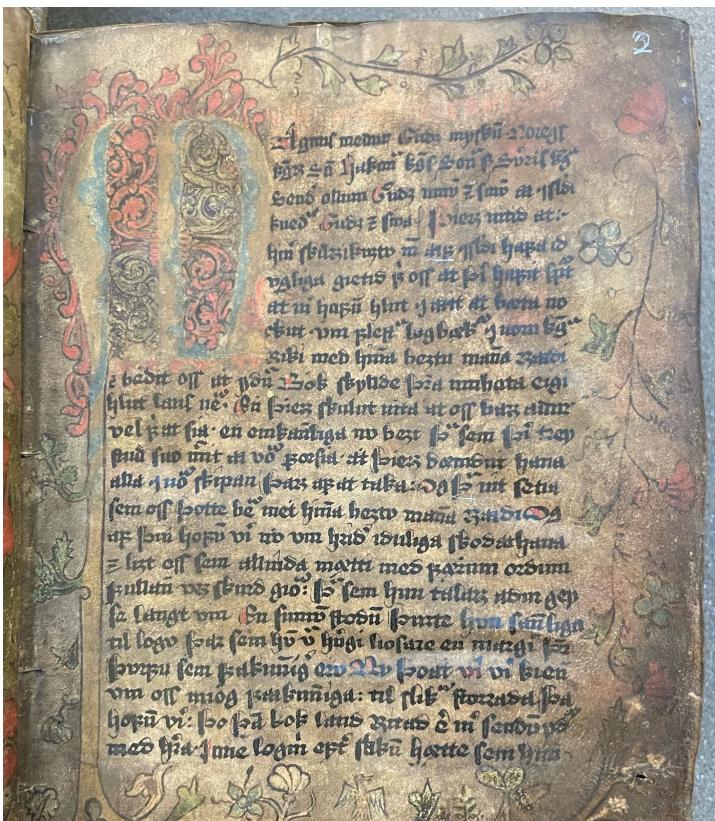


Photo of Oxford, Bodleian Library, MS. Icel. e. 1
taken by KAK on 31 January 2023
(with permission from the library).

Kapitan, Modelling humanities data with TEI-XML

msDesc in digital catalogues



MS. Icel. e. 1

Summary Catalogue no.: 31081

Contents

Jónsbók (code of laws)

Language(s): Old Norse

Physical Description

Form: codex

Support: parchment

Decoration

Fine borders, initials. (Pächt and Alexander iii. 1291, pl. CXIX)

History

Origin: 15th century first half : Icelandic and German (?)

Record Sources

Source: Description of MS. Icel. e. 1
https://medieval.bodleian.ox.ac.uk/catalog/manuscript_6284

The screenshot shows a GitHub repository page for 'bodleian / medieval-mss'. The commit 'paul-storey Add ARKs (Intellectual Entity and Catalogue Record)' has been pushed 9019fb last year. The file 'MS_Icel_e_1.xml' is selected, showing its XML code. The XML describes an entity with a title 'MS. Icel. e. 1 (Title)', a collection 'MS. Icel. (Icelandic)', and a person 'Matthew Halford' as a contributor.

```
<xml-model href="https://raw.githubusercontent.com/bodleian/consolidated-tei-schema/master/msdesc.rng" type="application/xml" select="1" />
<teiHeader>
  <fileDesc>
    <titleMS>
      <title>MS. Icel. e. 1</title>
      <title>MS. Icel. (Icelandic)</title>
    </titleMS>
    <respStmt>
      <respSummary description="responsiblePerson">
        <persname>Elizabeth Solopava</persname>
        <persname>Matthew Halford</persname>
      </respSummary>
      <responsible>
        <orgName>Special Collections</orgName>
        <orgName type="unit">Bodleian Libraries</orgName>
        <orgName type="institution">University of Oxford</orgName>
        <email>specialcollections.enquiries@bodleian.ox.ac.uk</email>
      </responsible>
    </respStmt>
    <idno type="MS. Icel. e. 1"></idno>
    <idno type="collection">Icelandic</idno>
    <idno type="catalogue">Western</idno>
  </fileDesc>
</teiHeader>
<publicationStmt>
  <source>MS. Icel. e. 1</source>
  <sourceIdentifier>MS_Icel_e_1</sourceIdentifier>
  <sourceIdentifier>urn:uuid:4e0a2a2d-0a2c-43f2-8a2a-0a2c4e0a2a2d</sourceIdentifier>
</publicationStmt>
```

Source: Description of MS. Icel. e. 1
https://github.com/bodleian/medieval-mss/blob/master/collections/Icel/MS_Icel_e_1.xml

msDesc in digital editions



[CLARIN](#) [Menota](#) [Corpora](#) [Text](#) [Search](#) [Dictionary](#) [Documentation](#) [User guides](#)

AM 242 fol - Codex Wormianus

Page: |< < 2 > >|

[Dipl

[mc. fol. 2]
1 A lmatigr guð skaplað i upphafi himin ok iðoð ok alla þa
2 lut er þeim rýgla ok lífarr með tu er atter eru frá komnar
3 adam ok evo ok riolagðar þeira kynlóði ok drepður um **heim**
4 allan. Enn er þram lidið funder *ja* vísindar mannríkis
5 voru suner goðir réttir trúðar enni miklu pléri færur þa epten
6 gíraðum heiminför ok væxtri gráðir vörðr Þórir jí ðækti
7 gud heiminför i hvarar gangi ok óllum kynkvændum heimins nem
8 með nos. Eptir þora lofti því, menið er heimins byggðu
9 ok ward enn fyrst átta er a þolimontok og byggð verðollan þa
10 kínfer er eldkaki agimi riar ok metubær en aþraekv guði hýjinn
11 at jví er tilr villus ægír neppa grub, en huert munði þa *ra lega* s
12 fluormerkum suna kom at þvír týndu grub napni ok úthol um weiss
13 ðælili kynni að kapara finum, en ægír ña feittu guð þeim íarsleg
14 þer fylkis við ura i heiminslæti, miðlaði haorn ok spektina su
15 ti ok allar græmer jar er fari matti loptzin ok barðaness, þat hu
16 dunduðu enn málum gegna er iðkinn ok örðrin af ruglarar hóp
17 fumum litum ok bæli skáld at hærri, bær viri wið, at iðsinum ór
18 fumum litum ok bæli skáld at hærri, bær viri wið, at iðsinum ór

Photo of Copenhagen, Den
Arnamagnæanske Samling AM 242 fol.
Taken from Handrit.org
<https://handrit.is/manuscript/view/da/AM242folM02-0242/>

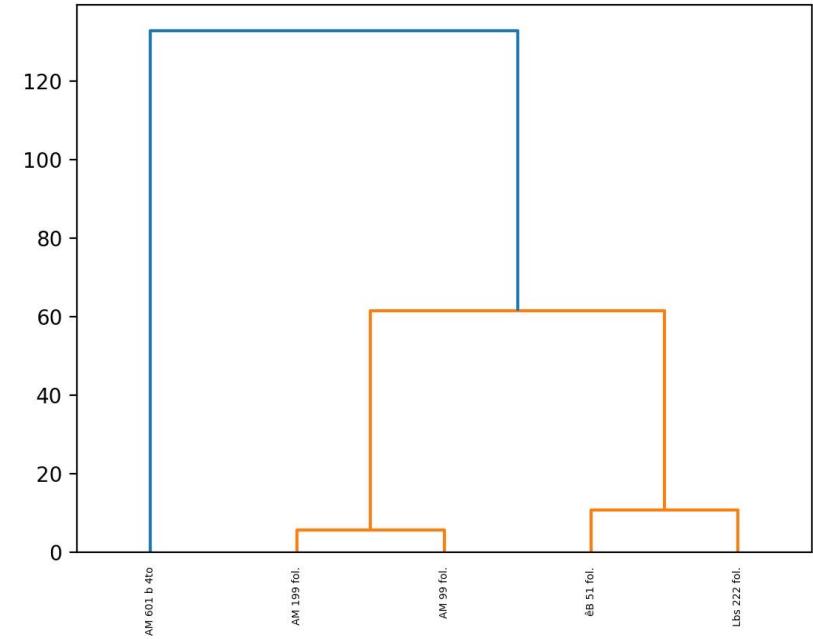
The MENOTA edition of Codex Wormianus by Karl Gunnar Johansson (2016)
<https://clarino.uib.no/menota/text/menota/AM-242-fol>

XML file behind the MENOTA edition of Codex Wormianus by Karl Gunnar Johansson (2016)

msDesc for quantitative codicology

**Five late 17th-cent. manuscripts,
classified based on:**

- Material
- Watermarks
- Leaf height & width
- Written area height & width
- Number of written lines per page
- Number of words per line
- ~~Number of columns~~



AM 601 b 4to – scribe 1,
AM 199 fol. – scribe 2, AM 99 fol. – scribe 2,
IB 51 fol. – scribe 3, Lbs 222 fol. – scribe 3

<msDesc>

- ▶ <msDesc> (manuscript description) contains a description of a single identifiable manuscript or other text-bearing object such as an early printed book.

The screenshot shows the TEI Guidelines website with a dark purple header and sidebar. The main content area displays the msDesc module page. At the top, there's a navigation bar with links to Guidelines, Activities, Tools, Membership, Support, About, News, and a search bar. Below the navigation is the title "TEI: Guidelines for Electronic Text Encoding and Interchange" and a note about the version being P5 Version 4.8.0, last updated on 2nd September 2024. The msDesc module page includes a brief overview, a table of attributes, and sections for members and containment.

Module	msdescription — Manuscript Description
Attributes	<ul style="list-style-type: none">att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:spaceatt.global.rendition: @rend, @style, @renditionatt.global.linking: @corresp, @synch, @sameAs, @copyOf, @next, @prev, @exclude, @selectatt.global.analytic: @anaatt.global.facs: @facsatt.global.change: @changeatt.global.responsibility: @cert, @respatt.global.source: @source <ul style="list-style-type: none">att.sortable: @sortKeyatt.type: @type, @subtypeatt.declaring: @declsatt.docStatus: @status
Member of	model.biblLike
Contained by	cmc: post core: add cit corr del desc emph head hi item listBibl meeting note orig p q quote ref reg relatedItem said sic stage textLang title unclear

Source:

<https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-msDesc.html>

TEI Manuscript Description Module

<msDesc>

<msDesc>

<msIdentifier> Info about the current and former shelfmarks </msIdentifier>

<msContents> Description of *textual contents* </msContents>

<physDesc> Description of physical features </physDesc>

<history> Info about the origin and provenance of the manuscript </history>

<additional> Administrative information </additional>

</msDesc>

<msDesc>

- ▶ <**msIdentifier**> contains the information required to identify the manuscript, such as holding institution and shelfmark.
- ▶ <**msContents**> describes the intellectual content of a manuscript either as a series of paragraphs or as an itemised list, with name of author and title, transcriptions of rubrics, incipits, explicits etc., language(s), as well as primary bibliographic references.

TEI Manuscript Description Module

<msDesc>

- ▶ **<physDesc>** groups information concerning all physical aspects of the manuscript, its material, size, format, collation (quire structure), script, decoration, binding, marginalia etc.
- ▶ **<history>** provides information on the history of the manuscript, its origin, provenance and acquisition by its current holding institution.
- ▶ **<additional>** groups other information about the manuscript, in particular administrative information relating to its availability, custodial history, surrogates etc.

<msIdentifier>

<**country**> contains the name of a geo-political unit, such as a nation or a country.

<**settlement**> contains the name of a settlement such as a city, town or village.

<**repository**> contains the name of a repository within which manuscripts or other objects are stored, possibly forming part of an institution.

<**idno**> supplies any form of identifier used to identify some object, such as a bibliographic item, a person, a title, an organization, etc. in a standardized way.

<msIdentifier>

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno type="Bod">MS Poet. Rawl. D. 169.</idno>
  </msIdentifier>
```

Source: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/examples-msIdentifier.html>

<msContents>

<**summary**> contains an overview of the available information concerning the intellectual content of a manuscript.

<**msItem**> describes an individual work or item within the intellectual content of a manuscript.

<**p**> marks paragraphs in prose.

<msContents>

```
<msDesc>
  <msIdentifier>
    <settlement>Oxford</settlement>
    <repository>Bodleian Library</repository>
    <idno>MS. Add. A. 61</idno>
    <altIdentifier type="SC">
      <idno>28843</idno>
    </altIdentifier>
  </msIdentifier>
  <msContents>
    <p>
      <quote>Hic incipit Brutus Anglie,</quote> the <title>De origine et gestis
      Regum Angliae</title> of Geoffrey of Monmouth (Galfridus Monumetensis): beg.
      <quote>Cum tecum multa & de multis.</quote> In Latin.</p>
  </msContents>
```

Source: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/examples-msContents.html>

<msContents>

```
<msContents>
  <msItem n="1" xml:id="MS_Add_A_61-item1">
    <author key="person_89028232">Geoffrey of Monmouth</author>
    <title key="work_1677">Historia regum Britanniae</title>
    <textLang mainLang="la">Latin</textLang>
  </msItem>
</msContents>
```

Source: https://github.com/bodleian/medieval-mss/blob/master/collections/Add_A/MS_Add_A_61.xml

<msItem>

<**locus**> defines a location within a manuscript typically as a sequence of folio references.

<**author**> contains the name(s) of an author, personal or corporate, of a work.

<**title**> contains a title for any kind of work.

<**textLang**> describes the languages identified within the work being described, rather than its description.

- ▶ Using the TEI Guidelines (<https://www.tei-c.org/>), find two other elements that might appear in msItem?

<msItem>

```
<msContents>
  <msItem n="1" defective="false"
    class="#law">
    <locus from="1v" to="71v">1v-71v</locus>
    <title type="uniform">Jónsbók</title>
    <incipit>Magnus m<ex>ed</ex> guds miskun Noregs k<ex>onungu</ex>r</incipit>
    <explicit>en<ex>n</ex> u<ex>ir</ex>da þo t<ex>il</ex> fullra aura</explicit>
  </msItem>
</msContents>
```

Source: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/examples-msItem.html>

@xml:id

@xml:id (identifier) provides a unique identifier for the element bearing the attribute.

```
<msDesc xml:id="Ms-Paris-BNF-Lat-5691">
```

@xml:lang

@xml:lang (language) indicates the language of the element content

```
<teiHeader xml:lang="en">
```

- ▶ **NB: @xml:lang does not indicate the language of the manuscript text; it is the language of the description!**

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019

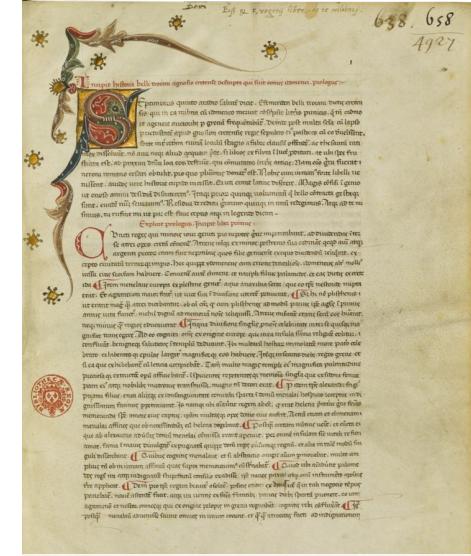
Kapitan, Modelling humanities data with TEI-XML

Exercise 1.1: Paris, BnF, Latin 5691

- ▶ Ex1_Paris_BnF_Latin_5691_Description_NKY.xml
- ▶ Ex1_Paris_BnF_Latin_5691_Facsimile_Gallica.pdf
- ▶ Answer the following questions:
 - ▶ How many texts are in this manuscript?
 - ▶ Why the first `<msItem>` has two `<author>` elements and why one of them has `@xml:lang` and the other doesn't?
 - ▶ What values the attribute `@type` has on `<title>` and what do they mean?

Source: Kapitan & Yavuz, Manuscripts in the Digital Age: XML-Based Catalogues and Editions, ESU DH, Leipzig 2019

Kapitan, Modelling humanities data with TEI-XML



Paris, Bibliothèque nationale de France,
Latin 5691, 1r. Photo: Gallica.

Ex2: Cod. Haen. 3518

- ▶ Using
 - ▶ Ex2_Manuscript_Description_Template.xml
 - ▶ Ex2_Leipzig_Cod.Haen._3518_Description_ManuscriptaMedievalia.pdf
 - ▶ Digital Facsimile on Handschriftenportal.de
(<https://resolver.staatsbibliothek-berlin.de/HSP000522D400000000>)
- ▶ Create a TEI-XML manuscript description of Cod. Haen. 3518 and encode
 - ▶ msIdentifier
 - ▶ msContents

