

# Modelling humanities data with TEI-XML

SCHOLARLY EDITING AND MANUSCRIPT CATALOGUING IN THE DIGITAL AGE

Dr Katarzyna Anna Kapitan  
22 October 2025

# Recap: TEI (basics)

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

```
    <teiHeader>
```

metadata

```
    </teiHeader>
```

```
    <text>
```

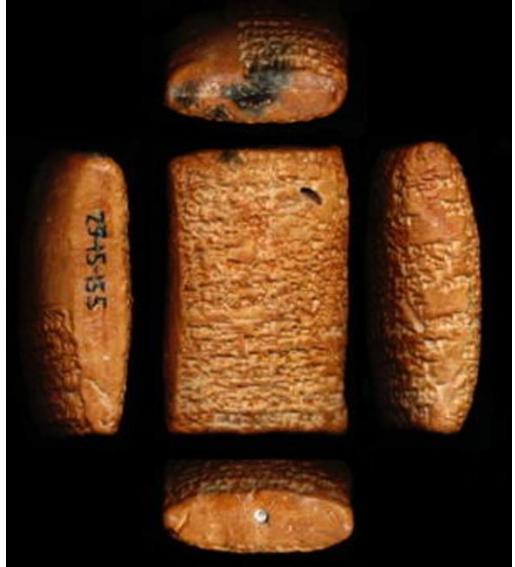
textual content

```
    </text>
```

```
</TEI>
```

# Cataloguing Manuscripts with TEI XML

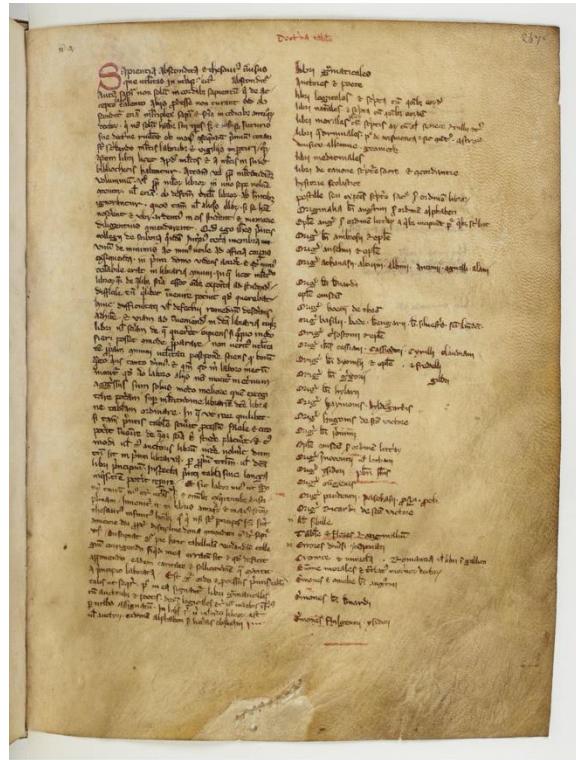
# Introduction to Cataloguing



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

Source:

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



Paris, Bibliothèque nationale de France , NAL 99, p. 259

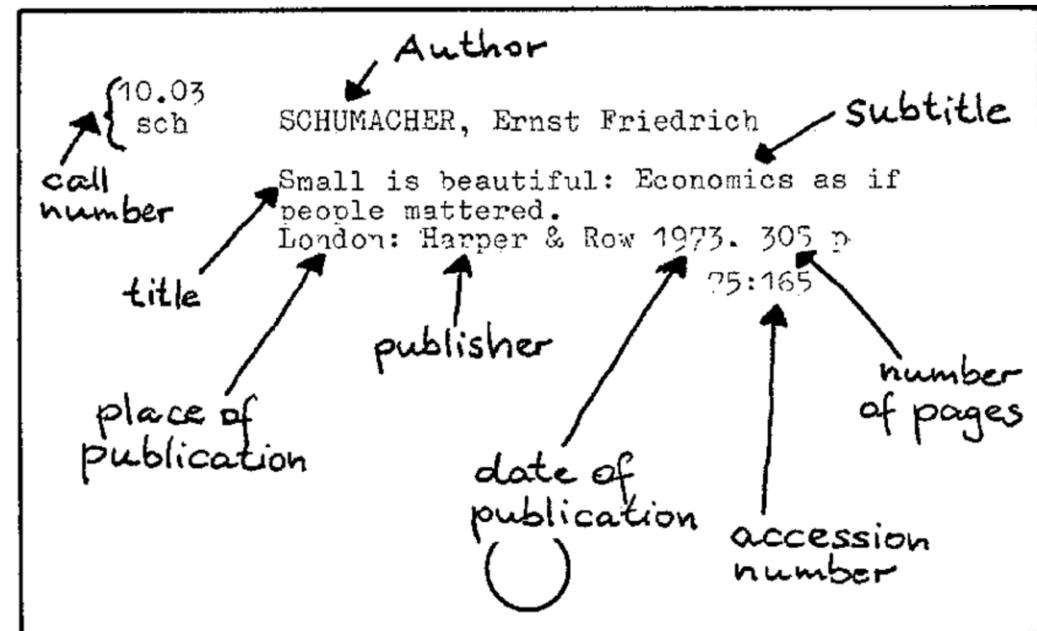
Source: <https://gallica.bnf.fr/>



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 Britannica)*



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)

## Additional 4869. ff. 2-66.

# 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 Trójumönum, 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

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 Thórðarson, at the expense of Magnús Jónasson of Vigur, an island in Isafjörðr, in the north-western peninsula of Iceland, in the years 1679, 1680. The whole volume contains:

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

TRÓJUMANNA SAGA: translated from a Danish abridgment of the Historia Trojana of Guido delle Colonne. Divided by the ornamental initials into 98 sections. *Icelandic*.

Heading :

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

#### Additional 11,113. ff. 171-281.

Paper; xviii<sup>th</sup> 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.<br>2. Gunnars pátr Þiðrandahána. f. 71. | 3. Brandkrossa pátr. f. 84.<br>4. Gull-Þóris saga. f. 92.<br>5. Kormaks saga. f. 127.<br>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-Mönum 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

# 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 English, Vatican Library, Support Digitization, HOME, DIGITAL COLLECTIONS, ONLINE CATALOGUES, and ABOUT. Below the header are two manuscript images: one showing a scene with figures and animals, and another showing musical notation. A search bar with the placeholder 'Search and discover collections' and a 'Simple' link is visible. Below the images are sections for 'Selected Manuscripts' featuring four small manuscript snippets and a 'Capp. Sist. 160' label. On the right is the website for Digital Bodleian. The header includes the logo 'DIGITAL BODLEIAN' and a search bar with the URL 'https://digital.bodleian.ox.ac.uk'. Below the header is a large image of a medieval manuscript page depicting various animals and figures. Text on the page reads 'DIGITALBODLEIAN MAKES THESE EXTRAORDINARY LIBRARY COLLECTIONS AVAILABLE ONLINE FOR THE VERY FIRST TIME...'. Below the image are sections for 'Enter new search term' and 'SEE HOW IT WORKS'. At the bottom are sections for 'THE COLLECTIONS' and 'Filter collections', with images of historical documents and 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 displays two side-by-side screenshots of digital manuscript libraries. On the left is the Gallica website (https://gallica.bnf.fr), operated by the Bibliothèque nationale de France (BnF). The interface features a header with the BnF logo and 'Gallica'. Below the header are sections for 'NEWS' and 'SELECTION'. A large image of a medieval manuscript initial is shown. Other news items include 'Discover 800 medieval manuscripts from France and England' and 'Andalusian music from Maghreb'. On the right is the e-codices website (https://www.e-codices.ch/en), operated by the Swiss National Library. The header includes the e-codices logo and links for 'Person Index', 'Annotations', 'About', 'Browse & Search', and language options (de, en, fr, it). The main content area features a large image of a medieval manuscript page with decorative initials. Text on the page reads: 'Knowing I lov'd my books, he furnish'd me, From mine own library with volumes that I prize above my dukedom.' - William Shakespeare, *The Temepest* (1611). Below the image are sections for 'Collections', 'News', and 'Documents'.

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

**STORIES FOR ALL TIME: THE ICELANDIC FORNALDARSÖGUR**

home texts manuscripts bibliography from the editors about the project

Home

### Stories for all time

*Stories for all time* is a research project based at The Arnamagnæan Institute, a research centre within the [Department of Nordic Studies and Linguistics](#) at the [University of Copenhagen](#). Principal funding for the project was provided by the [Velux Foundation](#) in the period 2011–2015. The project's aim is to survey the entire transmission history of the *Fornaldarsögur Norðurlanda* (generally known in English as 'mythical-heroic' or 'legendary' sagas), some 35 late medieval Icelandic prose narratives dealing with the early history of mainland Scandinavia, before the unification of Norway under Haraldr hárfagri and the settlement of Iceland in the late 9th century.

**Bibliotheca Laureshamensis digital**

Virtual Monastic Library of Lorsch

Home

Monastic Library of Lorsch

- > Library and Scriptorium
- > Selected Manuscripts
- > Manuscript Illumination
- > Literature

Virtual Library

Search

- > Manuscripts
- > Decoration

Image Gallery

Glossary

Press Coverage

Terms of Use

Project Partners

Project Description

**Bibliotheca Laureshamensis – digital: Virtual Monastic Library of Lorsch**

The project's aim was to virtually reconstruct the monastic library of UNESCO-World Heritage Site [Abbey Lorsch](#). More than 300 medieval manuscripts have survived and today they are spread over 73 libraries worldwide. The unification of the Lorsch codices and fragments for the first time makes the Abbey's intellectual foundation, and moreover the Carolingian world picture visible and researchable.

In comparison to Abbey Lorsch almost no other library from the Early Middle Ages that is reconstructable today, accomplished a comparable achievement in synthesis than Lorsch Abbey. It was one of the centres, in which ancient heritage was secured throughout all the following disruptions through transcriptions, comments and further editing. Furthermore, the amount of knowledge was increased in entire medieval culture. For the Humanistic Lorsch's monastic library

**Virtual Monastic Library of Lorsch**

<http://bibliotheca-laureshamensis-digital.de/>

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 & 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 Magús Erlingsson

Language of Text Icelandic

#### PHYSICAL DESCRIPTION

Support Papir.

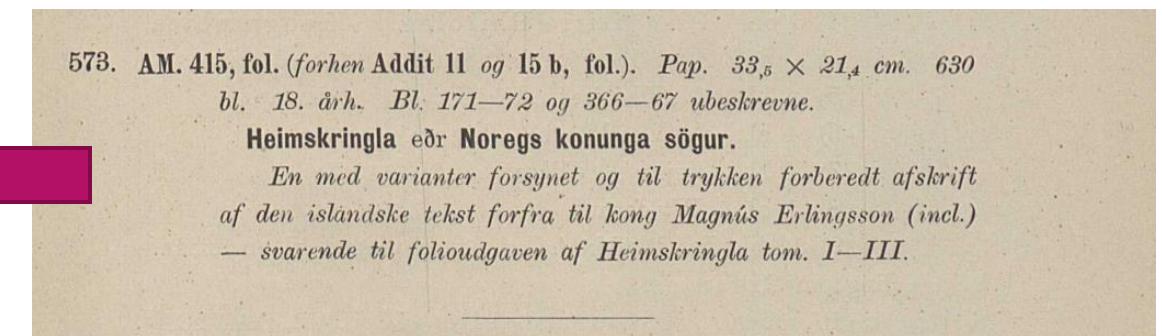
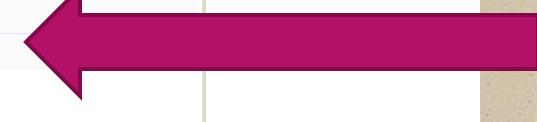
No. of leaves 630. Bl. 171-72 og 366-67 er ubeskrevne. 335 mm x 214 mm.

#### HISTORY

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

#### ADDITIONAL

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



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

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

# Digital & Digitised Catalogues

The screenshot shows a digital catalog entry for a manuscript. At the top, there's a search bar with 'latin 7906' and navigation links for 'ADVANCED SEARCH', 'SELECTIONS', and 'BLOG'. Below the search bar, the title 'Full notice' is displayed. The main content area contains detailed information about the manuscript, including its title, author, and various descriptive fields like 'Title', 'Publication date', 'Type', 'Language', 'Format', 'Description', 'Source', 'Relationship', 'Provenance', and 'Date of online availability'. A large red arrow points from the right side of the catalog entry towards the Latin text on the right.

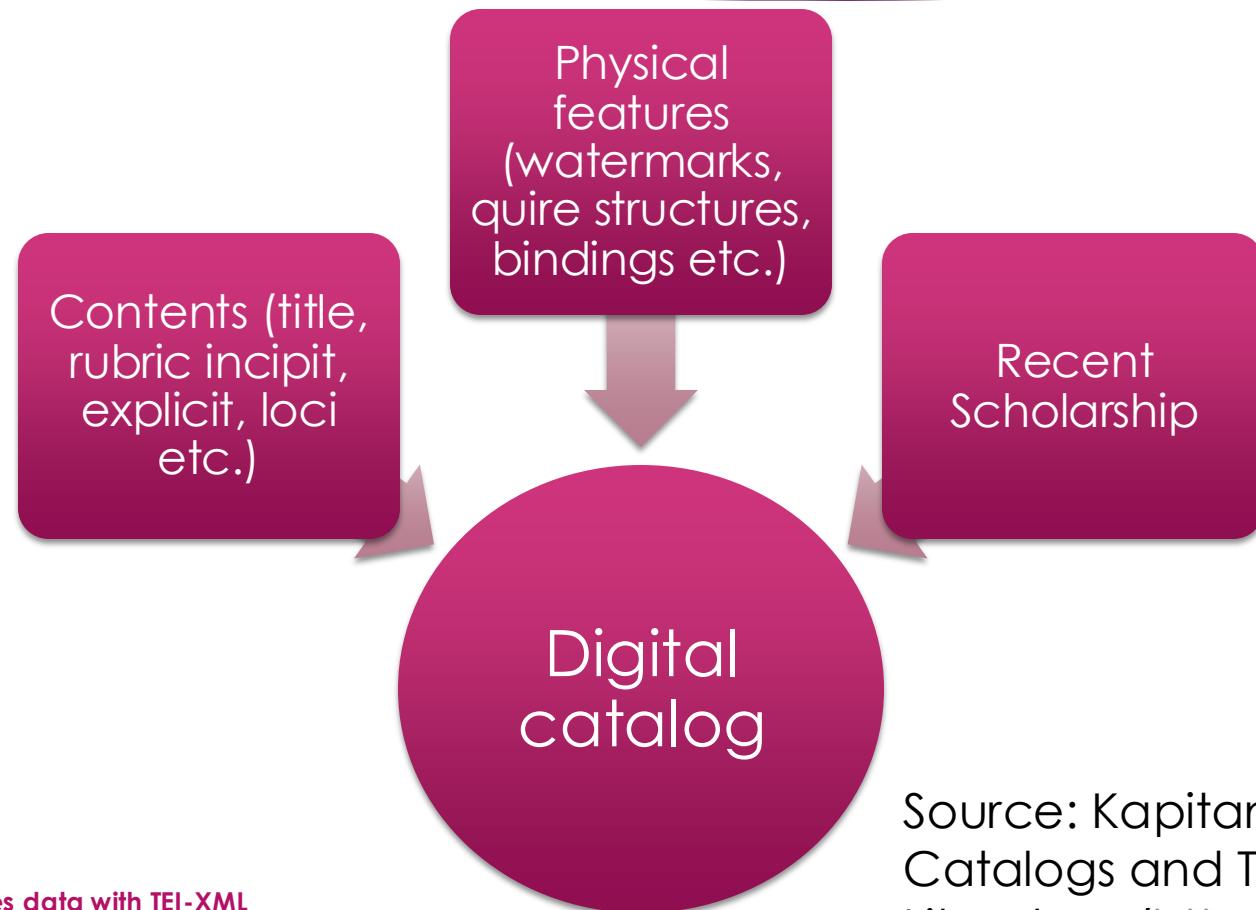
Gallica (2019):  
<https://gallica.bnf.fr/ark:/12148/btv1b85721963.r=Latin%207906?rk=21459;2>

vii M C M V I.  
Codex membranaceus, olim Colbertinus.  
Ibi continentur:  
1.º *Publili Terentii*, Afri, comœdiæ sex;  
decimo tertio sœculo exaratæ.  
2.º *D. Juvenalis* satyrarum quarta, cuius  
initium desideratur, quinta, sexta, septima  
& pars octavæ; decimo sœculo exaratæ.  
3.º *Publili Virgilii Maronis* æneïdos libri  
tertii finis, quartus & quintus integri; nono  
sœculo exarati.  
4.º *Daretis Phrygii* de excidio Trojano  
historia : interprete *Cornelio Nepote*, cuius  
F f f ij  
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



Source: 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

# TEI Manuscript Description Module

## msdescription

TEI Guidelines Activities Tools Membership Support About News Search Search ▾

TEI: Guidelines for Electronic Text Encoding and Interchange  
P5 Version 4.8.0. Last updated on 2nd September 2024, revision 4e6e41b0b

<b>Table of contents</b> <a href="#">11.1 Overview</a> <a href="#">11.2 The Manuscript Description Element</a> <a href="#">11.3 Phrase-level Elements</a> <a href="#">11.4 The Manuscript Identifier</a> <a href="#">11.5 The Manuscript Heading</a> <a href="#">11.6 Intellectual Content</a> <a href="#">11.7 Physical Description</a> <a href="#">11.8 History</a> <a href="#">11.9 Additional Information</a> <a href="#">11.10 Manuscript Parts</a> <a href="#">11.11 Manuscript Fragments</a> <a href="#">11.12 Module for Manuscript Description</a>  <a href="#">« 10 Dictionaries</a> <a href="#">» 12 Representation of Primary Sources</a> <a href="#">Home</a>	<h2>11 Manuscript Description</h2> <h3>11.1 Overview <small>42</small></h3> <p>The <b>msdescription</b> module— defines a special purpose element which can be used to provide detailed descriptive information about handwritten primary sources and other text-bearing objects. Although originally developed to meet the needs of cataloguers and scholars working with medieval manuscripts in the European tradition, the scheme presented here is general enough that it can also be extended to other traditions and materials, and is potentially useful for any kind of text-bearing artefact. Where the textuality of an object is not the primary concern, encoders may wish to use the <a href="#">object</a> element which provides a very similar system of description (see <a href="#">14.3.6 Objects</a>).</p> <p>The scheme described here is also intended to accommodate the needs of many different classes of encoders. On the one hand, encoders may be engaged in <i>retrospective conversion</i> of existing detailed descriptions and catalogues into machine tractable form; on the other, they may be engaged in cataloguing <i>ex nihilo</i>, that is, creating new detailed descriptions for materials never before catalogued. Some may be primarily concerned to represent accurately the description itself, as opposed to the ideas and interpretations the description represents; others may have entirely opposite priorities. At one extreme, a project may simply wish to capture an existing catalogue in a form that can be displayed on the Web, and which can be searched for literal strings, or for such features such as titles, authors and dates; at the other, a project may wish to create, in highly structured and encoded form, a detailed database of information about the physical characteristics, history, interpretation, etc. of the material, able to support practitioners of <i>quantitative codicology</i> as well as librarians.</p> <p>To cater for this diversity, here as elsewhere, these Guidelines propose a flexible strategy, in which encoders must choose for themselves the approach appropriate to their needs, and are provided with a choice of encoding mechanisms to support those differing degrees.</p> <h3>11.2 The Manuscript Description Element</h3> <p>The <a href="#">msDesc</a> element will normally appear within the <a href="#">sourceDesc</a> element of the header of a TEI-conformant document, where the document being encoded is a digital representation of some manuscript original, whether as an encoded transcription, as a collection of digital images (as described in <a href="#">12.1 Digital Facsimiles</a>), or as some combination of the two. However, in cases where the document being encoded is essentially a collection of manuscript descriptions, the</p> <p><a href="#">« 11.1 Overview</a> <a href="#">» 11.3 Phrase-level Elements</a> <a href="#">Home</a></p>
--	---

## 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

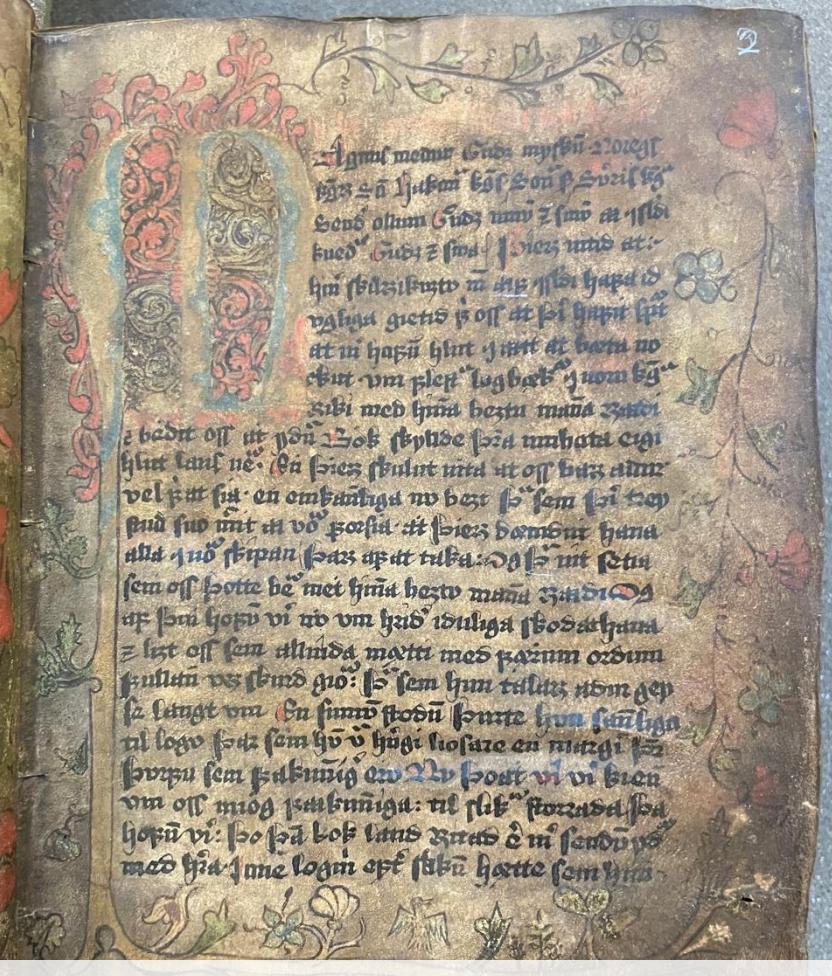


Photo of Oxford, Bodleian Library, MS. Icel. e. 1  
taken by KAK on 31 January 2023  
(with permission from the library).

# msDesc in digital catalogues

MS. Icel. e. 1

Summary Catalogue no.: 31082

## Contents

## *Jónsbók (code of laws)*

**Language(s):** Old Norse

## **Physical Description**

## **Form: codex**

**Support:** parchment

## Decoration

Fine borders, initials. (Pächt and Alexander iii. 1290, pl. CXVII). Fine miniature, coat of arms (very rubbed) added on (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/cat/a/g/o/a/manuscript\\_6284](https://medieval.bodleian.ox.ac.uk/cat/a/g/o/a/manuscript_6284)

The screenshot shows a GitHub repository named 'bodleian / medieval-mss'. The 'Code' tab is selected, displaying a pull request from 'paul-storey' titled 'Add ARKs (Intellectual Entity and Catalogue Record)'. The file being viewed is 'MS\_Icel\_e\_1.xml' under the 'collections / icel' directory. The code is a TEI XML document with the following structure:

```
<?xml-model href="https://raw.githubusercontent.com/bodleian/consolidated-tei-schema/master/msdesc.rng" type="application/xml" sc="1"?>
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>MS. Icel. e. 1</title>
      <title type="collection">MSS. Icel. (Icelandic)</title>
      <respStmt>
        <respSummary description=></resp>
        <persName>Elizabeth Solopova</persName>
        <persName>Matthew Holford</persName>
      </respStmt>
    </titleStmt>
    <publicationStmt>
      <publisher>
        <orgName type="department">Special Collections</orgName>
        <orgName type="unit">Bodleian Libraries</orgName>
        <orgName type="institution">University of Oxford</orgName>
        <email>specialcollections.enquiries@bodleian.ox.ac.uk</email>
      </publisher>
      <idno type="msID">MS_Icel_e_1</idno>
      <idno type="collection">Icelandic</idno>
      <idno type="catalogue">Western</idno>
    </publicationStmt>
    <sourceDesc>
      <msDesc xml:id="MS_Icel_e_1" xml:lang="en">
        <msIdentifier>
```

 Lady\_Margaret\_Hall 27 <country>United Kingdom</country>

[https://github.com/bodleian/medieval-mss/blob/master/collections/Icel/MS\\_Icel\\_e\\_1.xml](https://github.com/bodleian/medieval-mss/blob/master/collections/Icel/MS_Icel_e_1.xml)

# msDesc in digital editions



Photo of Copenhagen, Den  
Arnamagnæanske Samling AM 242 fol.  
Taken from Handrit.org  
(<https://handrit.is/manuscript/view/da/AM02-0242/>)

**CLARINO** Menota Corpora Text Search Dictionary Documentation User guides

AM 242 fol – Codex Wormianus

Page: |< < 2 > >|

[ Facs ]

[ms: fol. 2]

[ Dipl ]

[ms: fol. 2]

3 lut er þeim fylgia ok síðarft menu tua er ætter eru fra kommar  
3 adam ok euo ok fiolegaþ þeir kynfólk ok dærepur um **heim**  
4 allan. Enir er fram lífu stunder þa viðnaði mannfólkid  
5 voru sunnar goðer ok rett trúðer enni miklu pleiri finnur þa epten  
6 girndum heiminsfín ok vraktv guði bodoðr ok eyrir þi ðekti  
7 guð heiminiðum i fóvar ganki ollum kykkvenum heiminsfín nem  
8 með noa. Eptera noa plod liþum .viii. með þeir er heiminið nem  
9 ok varð enni fém ryktur at þa er polimentz ok byggð veröldlinn þa var  
10 klenf er elskafi agírnar piar ok metmáðr en aþraktvu guði hlyðni  
11 at því at þeir vildu ægja neppu guð, en huurr mættu þa fræga se  
12 flosmekrum sua kom at þeir týndu guðs napni ok uðaz um ureloð  
13 ðælli kynnt a skapara finum. En ægír at félit veitt guð heim i ardlegrum  
14 þeir fikkiliðu vel uera i heiminiðum. miðlaði hann ok spektina sua at  
15 ti ok allar græiner þær er fia matti loptzinfi ok iardbærenar, þa hu  
16 dundu hui þar munði gegna þær ok dyrin ok fuglarmer heiminiðum  
17 fumnum lutum ok þo ulik at hætti. þar var eitt eðli at iordin var gr

# The MENOTA edition of Codex Wormianus by Karl Gunnar Johansson (2016)

<https://clarino.uib.no/menota/text/menota/AM-242-fo>

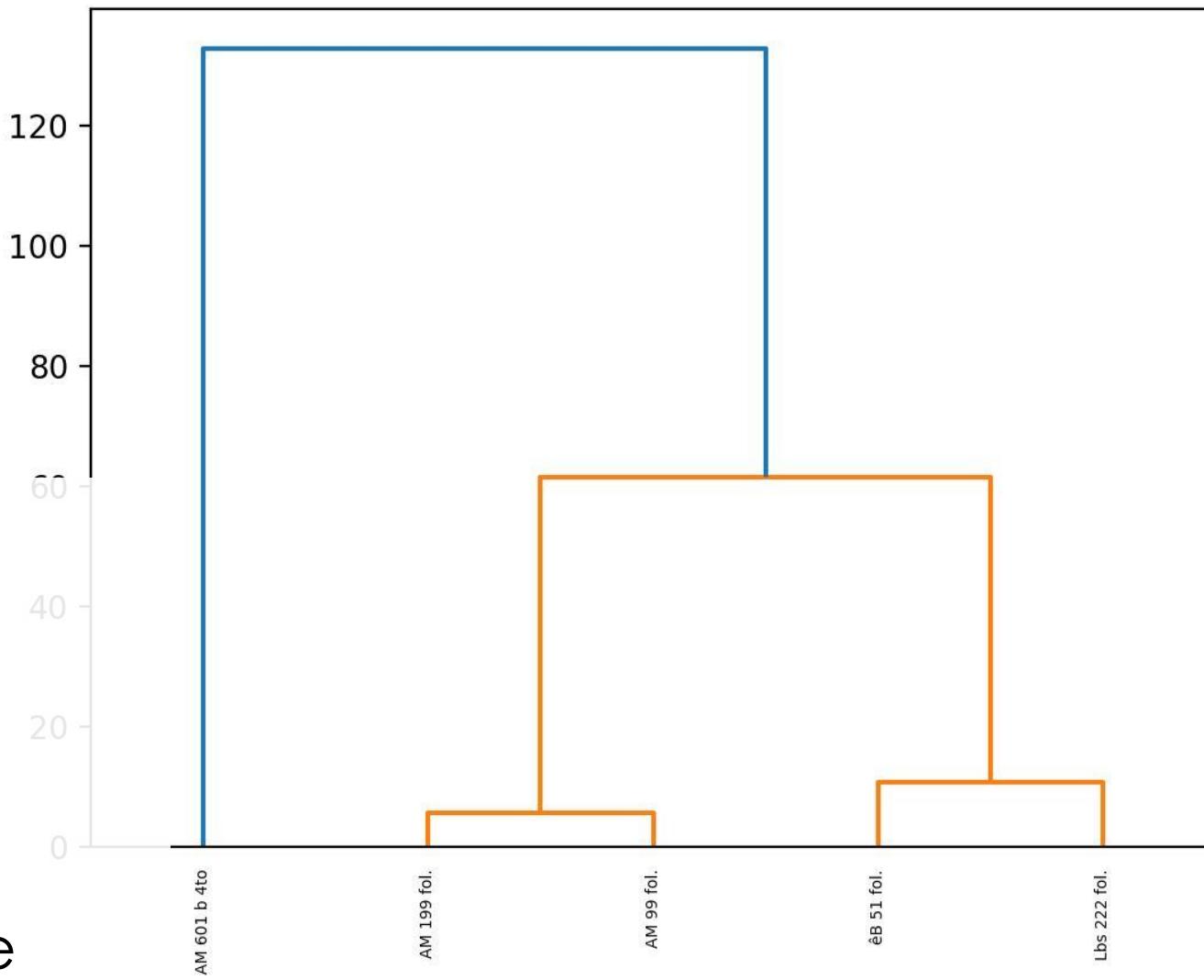
Kapitan, Modelling humanities data with TEI-XML

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

# quantitative codicology

**Five late 17<sup>th</sup>-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 blue header bar containing the TEI logo, a search bar, and navigation links for Guidelines, Activities, Tools, Membership, Support, About, and News. Below the header, a banner for "TEI: Guidelines for Electronic Text Encoding and Interchange P5 Version 4.8.0. Last updated on 2nd September 2024, revision 4e6e41b0b" is displayed. The main content area has a light gray background and features a table with the following rows:

<msDesc>	
Module	msdescription — Manuscript Description
Attributes	<ul style="list-style-type: none"><li>• att.global: @xml:id, @n, @xml:lang, @xml:base, @xml:space<ul style="list-style-type: none"><li>◦ att.global.rendition: @rend, @style, @rendition</li><li>◦ att.global.linking: @corresp, @syncrh, @sameAs, @copyOf, @next, @prev, @exclude, @select</li><li>◦ att.global.analytic: @ana</li><li>◦ att.global.facs: @facs</li><li>◦ att.global.change: @change</li><li>◦ att.global.responsibility: @cert, @resp</li><li>◦ att.global.source: @source</li></ul></li><li>• att.sortable: @sortKey</li><li>• att.typed: @type, @subtype</li><li>• att.declaring: @decls</li><li>• att.docStatus: @status</li></ul>
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>: 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>
```

# 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> in TEI Guidelines

- ▶ Work in Pairs
- ▶ Using the TEI Guidelines (<https://www.tei-c.org/>), find two other elements (besides <summary>, <msItem>, <p>) that may appear in msContents.
- ▶ To which modules of TEI do they belong?

# <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](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.

<b>May contain</b>	<p><b>analysis:</b> <a href="#">interp</a> <a href="#">interpGrp</a> <a href="#">span</a> <a href="#">spanGrp</a></p> <p><b>certainty:</b> <a href="#">certainty</a> <a href="#">precision</a> <a href="#">respons</a></p> <p><b>core:</b> <a href="#">author</a> <a href="#">bibl</a> <a href="#">biblStruct</a> <a href="#">binaryObject</a> <a href="#">cb</a> <a href="#">cit</a> <a href="#">editor</a> <a href="#">ellipsis</a> <a href="#">gap</a> <a href="#">gb</a> <a href="#">graphic</a> <a href="#">index</a> <a href="#">lb</a> <a href="#">listBibl</a> <a href="#">meeting</a> <a href="#">milestone</a> <a href="#">note</a> <a href="#">noteGrp</a> <a href="#">p</a> <a href="#">pb</a> <a href="#">quote</a> <a href="#">respStmt</a> <a href="#">textLang</a> <a href="#">title</a></p> <p><b>figures:</b> <a href="#">figure</a> <a href="#">notatedMusic</a></p> <p><b>header:</b> <a href="#">biblFull</a> <a href="#">funder</a> <a href="#">idno</a> <a href="#">principal</a> <a href="#">sponsor</a></p> <p><b>iso-fs:</b> <a href="#">fLib</a> <a href="#">fs</a> <a href="#">fvLib</a></p> <p><b>linking:</b> <a href="#">ab</a> <a href="#">alt</a> <a href="#">altGrp</a> <a href="#">anchor</a> <a href="#">join</a> <a href="#">joinGrp</a> <a href="#">link</a> <a href="#">linkGrp</a> <a href="#">timeline</a></p> <p><b>msdescription:</b> <a href="#">colophon</a> <a href="#">decoNote</a> <a href="#">explicit</a> <a href="#">filiation</a> <a href="#">finalRubric</a> <a href="#">incipit</a> <a href="#">locus</a> <a href="#">locusGrp</a> <a href="#">msDesc</a> <a href="#">msItem</a> <a href="#">msItemStruct</a> <a href="#">rubric</a></p> <p><b>spoken:</b> <a href="#">incident</a> <a href="#">kinesic</a> <a href="#">pause</a> <a href="#">shift</a> <a href="#">vocal</a> <a href="#">writing</a></p> <p><b>textcrit:</b> <a href="#">app</a> <a href="#">witDetail</a></p> <p><b>textstructure:</b> <a href="#">argument</a> <a href="#">byline</a> <a href="#">docAuthor</a> <a href="#">docDate</a> <a href="#">docEdition</a> <a href="#">docImprint</a> <a href="#">docTitle</a> <a href="#">epigraph</a> <a href="#">imprimatur</a> <a href="#">titlePart</a></p> <p><b>transcr:</b> <a href="#">addSpan</a> <a href="#">damageSpan</a> <a href="#">delSpan</a> <a href="#">fw</a> <a href="#">listTranspose</a> <a href="#">metamark</a> <a href="#">space</a> <a href="#">substJoin</a></p>
--------------------	---

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

# <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

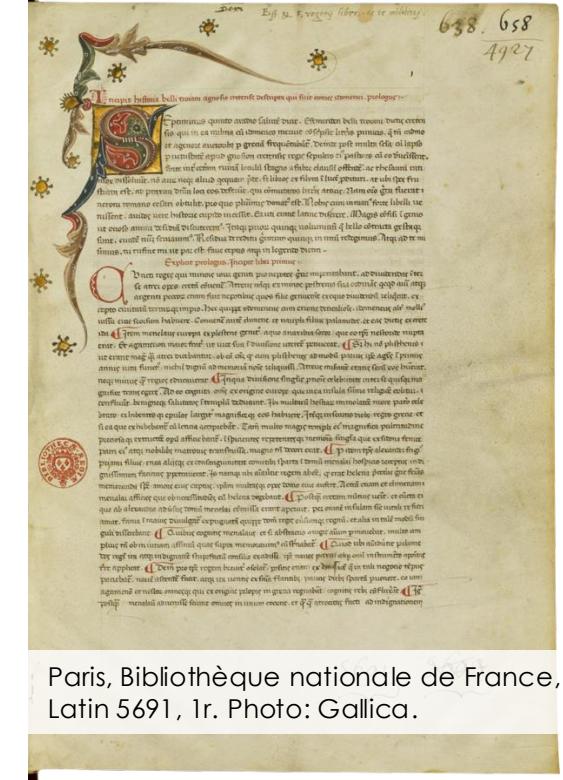
# Ex1: Paris, BnF, Latin 5691

## ► Using

- 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?



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

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

# Ex2: Cod. Haen. 3518



## Using

- ▶ Ex2\_Manuscript\_Description\_Template.xml
- ▶ Ex2\_Lipzig\_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**

