Modelling humanities data with TEI-XML

SCHOLARLY EDITING AND MANUSCRIPT CATALOGUING IN THE DIGITAL AGE

Dr Katarzyna Anna Kapitan
1 October 2025

Introductions

- **Teacher**:
  - Dr Katarzyna Anna Kapitan
  - katarzyna.kapitan [at] chartes.psl.eu
- **Students**:
  - Name
  - Background and Study Programme
  - Research Interests
  - Experience with Markup Languages (which ones, what experience)

Course overview

- **Classes**:
  - 10 x 2 hours between 1 October and 10 December
  - Wednesdays, 15:15-17:15
    - Besides 8 October -> **Time Change: 10 October @ 10:00 on Zoom (the session will be recorded for those who cannot make it)**

Course overview

- **Course prerequisites:**
  - Laptop with installed
    - **Oxygen XML Editor** (https://www.oxygenxml.com/xml_editor/download_oxygenxml_editor.html)
    - **Sublime Text** (https://www.sublimetext.com/download)

Objectives

- Understand the XML structure and its customisation
- Use the TEI Guidelines for text editing and manuscript cataloguing
- Create XML templates and validation rules
- Document encoding choices
- Apply XML to individual projects

Schedule

Session 1	Markup Languages and Text Encoding
Session 2	Extensible Markup Language (XML) and Document Type Definition (DTD)
Session 3	Text Encoding Initiative (TEI) Guidelines
Session 4	Describing Primary Sources
Session 5	Transcribing Primary Sources
Session 6	Editing Primary Sources
Session 7	Annotating Primary Sources
Session 8	XPath
Session 9	Customisation and Documentation
Session 10	Customisation and Validation

Assessment method

- **Assessment:**
  - Attendance and Participation (20%)
  - Encoding Project Portfolio:
    - Assignment 1 (due 20 October): 15%
    - Assignment 2 (due 17 November): 25%
    - Assignment 3 (due 22 December): 40%

Course Materials

- **GitHub Repo:** https://github.com/KAKDH/TNAH_XML2025

Text Encoding & Markup Languages

Markup Languages

- **What is a markup language?**
  - Computer language
  - Uses tags to define elements within a document
  - Human-readable

- **Examples**:
  - HTML (Hypertext Markup Language)
  - Markdown
  - TeX & LaTeX
  - **XML (Extensible Markup Language)**
  - Scribe, GML (Generalized Markup Language) & SGML (Standard Generalized Markup Language)

Markup

	Semantic	Presentational
LaTeX	*\selectlanguage{latin}{ad hoc}*	*\emph{ad hoc}*
HTML	*<i lang="la">ad hoc</i>*	*<i>ad hoc</i>*
XML	*<foreign xml:lang="la">ad hoc</foreign>*	*<hi rend="i">ad hoc</hi>*

Exercises

- **Sync the GitHub repo (https://github.com/KAKDH/TNAH_XML2025) and access the Exercises Folder: Week1/Exercises/Ex1**
- **GitHub Tutorials:**
- https://docs.github.com/fr/repositories/creating-and-managing-repositories/cloning-a-repository
- https://dev.to/warish/a-step-by-step-guide-to-cloning-a-github-repository-to-your-local-machine-4ngc
- https://dev.to/phrawzty/easily-keep-your-local-git-repo-up-to-date-l52

Exercise 1 (Week1/Exercises/Ex1)
Individual Work (10 min)

- Which of these files contain markup?
- What type of markup is used?
- How is the markup expressed?
- How is the document structured?
- What are these types of markup usually used for, give examples?
- List all markup elements used in the documents and explain what are they used for (what do you think they mean)?

Example 1:
AM_30_fol.txt
AM_30_fol.md
AM_30_fol.html
AM_30_fol.doc

Example 2:
AM_30_fol.txt
AM_30_fol.md
AM_30_fol.tex
AM_30_fol.doc

Group work (10 min)

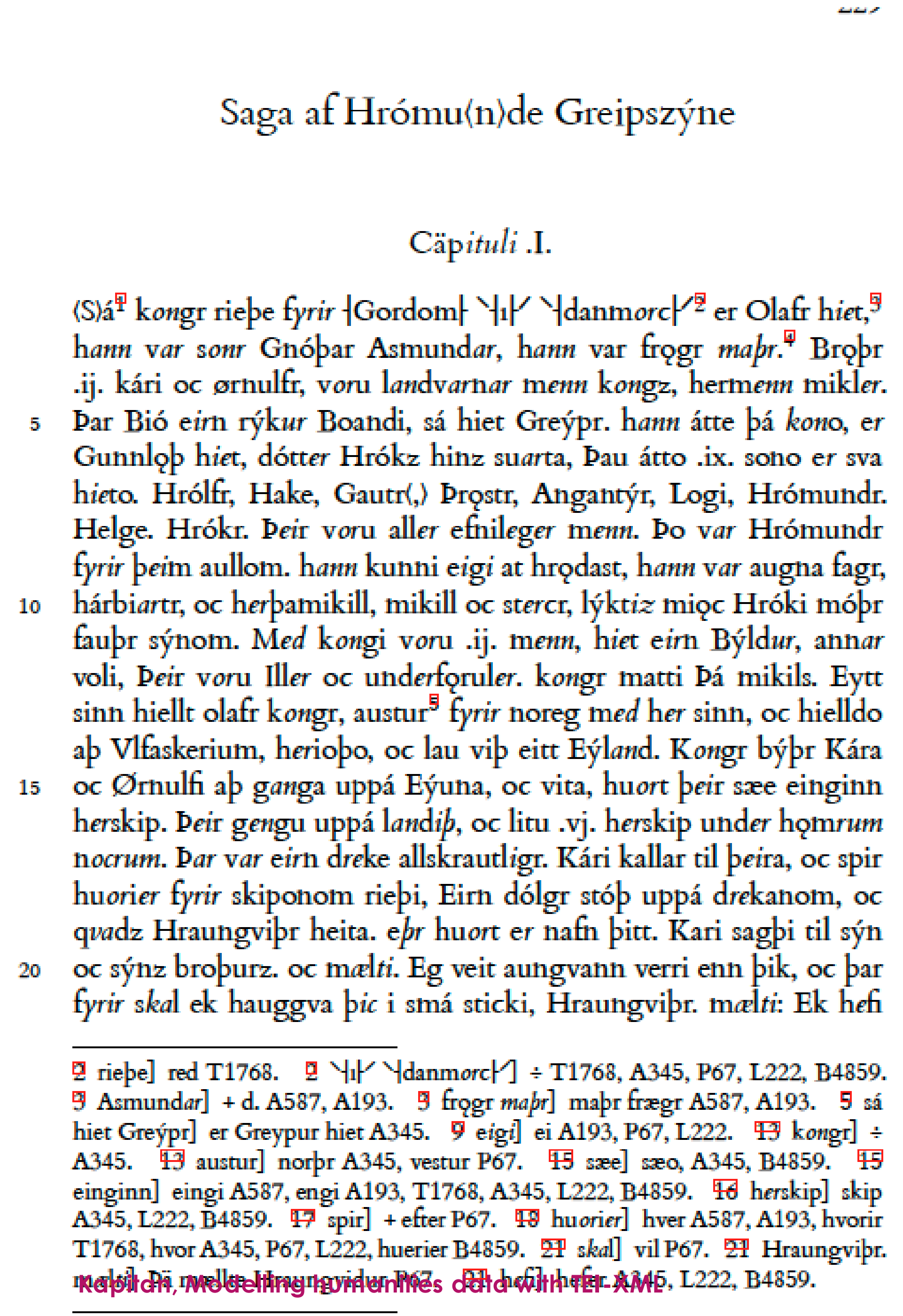
- *Create two groups. One group consisting of everyone who worked on example 1 and one group consisting of everyone who worked on example 2.*
- Compare your answers and prepare a presentation of your example, choose one person who will present your answers.
- **Group with example 1:** focus on HTML
- **Group with example 2:** focus on TEX

Plain Text | Formatted Text

- **Plain text** format contains no formatting information,
- **Formatted Text - Rich text** format includes formatting details such as font size, style, colour, etc.
  - **WYSIWYG** (**W**hat **Y**ou **S**ee **I**s **W**hat **Y**ou **G**et) - a formatted document as it will appear on screen or in print without showing the descriptive code.

Typesetting with TeX & LaTeX

- TeX is a typesetting system widely used in academia, especially in mathematics, computer science, engineering, linguistics, and many more, among them: **digital scholarly editing.**
- **WYSIWYM** (What You See Is What You **Mean**) is a paradigm for editing a structured document separating presentation from content.
- TeX commands commonly start with a backslash and are grouped with curly braces.
  - **\myCommand{ My Content goes here }**



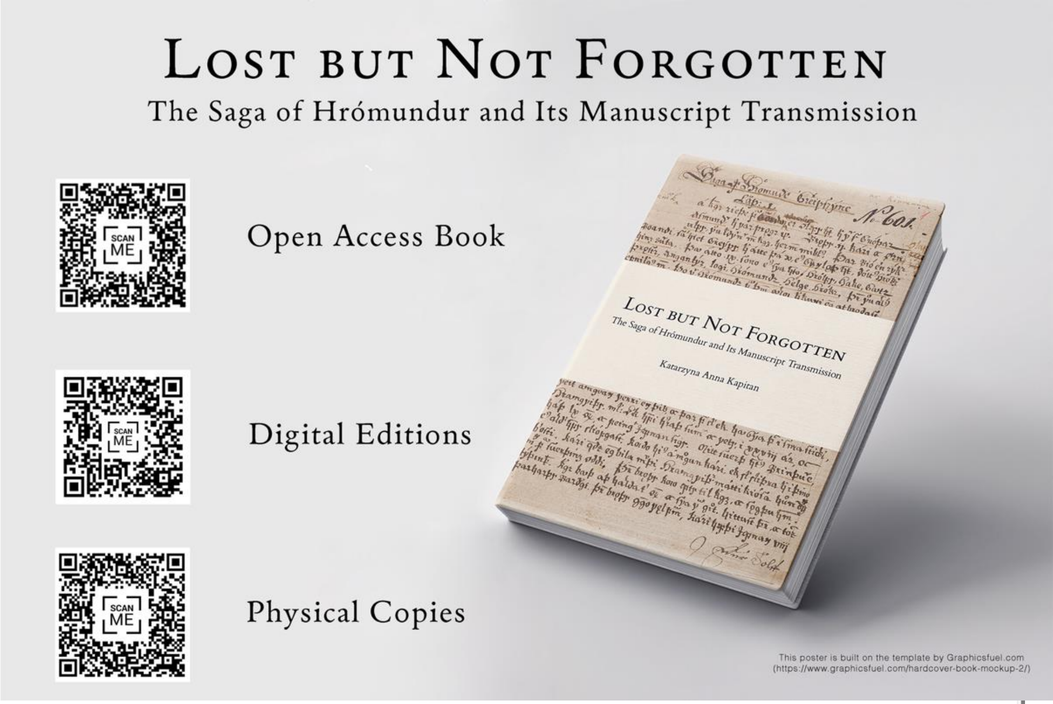
Saga af Hrómu(n)de Greipszýne

Cäpituli .I.

1r

(S)á[1] kongr rieþe fyrir ⊦Gordom⊦ `⊦ı⊦´ `⊦danmorc⊦´[2] er Olafr hiet,[3] hann var sonr Gnóþar Asmundar, hann var frǫgr maþr.[4] Brǫþr .ij. kári oc ørnulfr, voru landvarnar menn kongz, hermenn mikler. Þar Bió eirn rýkur Boandi, sá hiet Greýpr. hann átte þá kono, er Gunnlǫþ hiet, dótter Hrókz hinz suarta, Þau átto .ix. sono er sva hieto. Hrólfr, Hake, Gautr(,) Þrǫstr, Angantýr, Logi, Hrómundr. Helge. Hrókr. Þeir voru aller efnileger menn. Þo var Hrómundr fyrir þeim aullom. hann kunni eigi at hrǫdast, hann var augna fagr, hárbiartr, oc herþamikill, mikill oc stercr, lýktiz miǫc Hróki móþr fauþr sýnom. Med kongi voru .ij. menn, hiet eirn Býldur, annar voli, Þeir voru Iller oc underfǫruler. kongr matti Þá mikils. Eytt sinn hiellt olafr kongr, austur[5] fyrir noreg med her sinn, oc hielldo aþ Vlfaskerium, herioþo, oc lau viþ eitt Eýland. Kongr býþr Kára oc Ørnulfi aþ ganga uppá Eýuna, oc vita, huort þeir sæe einginn herskip. Þeir gengu uppá landiþ, oc litu .vj. herskip under hǫmrum nocrum. Þar var eirn dreke allskrautligr. Kári kallar til þeira, oc spir huorier fyrir skiponom rieþi, Eirn dólgr stóþ uppá drekanom, oc qvadz Hraungviþr heita. eþr huort er nafn þitt. Kari sagþi til sýn oc sýnz broþurz. oc mælti. Eg veit aungvann verri enn þik, oc þar fyrir skal ek hauggva þic i smá sticki, Hraungviþr. mælti: Ek hefi

2 rieþe] red T1768. 2 `⊦ı⊦´ `⊦danmorc⊦´] ÷ T1768, A345, P67, L222, B4859. 3 Asmundar] + d. A587, A193. 3 frǫgr maþr] maþr frægr A587, A193. 5 sá hiet Greýpr] er Greypur hiet A345. 9 eigi] ei A193, P67, L222. 13 kongr] ÷ A345. 13 austur] norþr A345, vestur P67. 15 sæe] sæo, A345, B4859. 15 einginn] eingi A587, engi A193, T1768, A345, L222, B4859. 16 herskip] skip A345, L222, B4859. 17 spir] + efter P67. 18 huorier] hver A587, A193, hvorir T1768, hvor A345, P67, L222, huerier B4859. 21 skal] vil P67. 21 Hraungviþr. mælti] Þá mællte Hraungvidur P67. 21 hefi] hefer A345, L222, B4859.



Lost but Not Forgotten

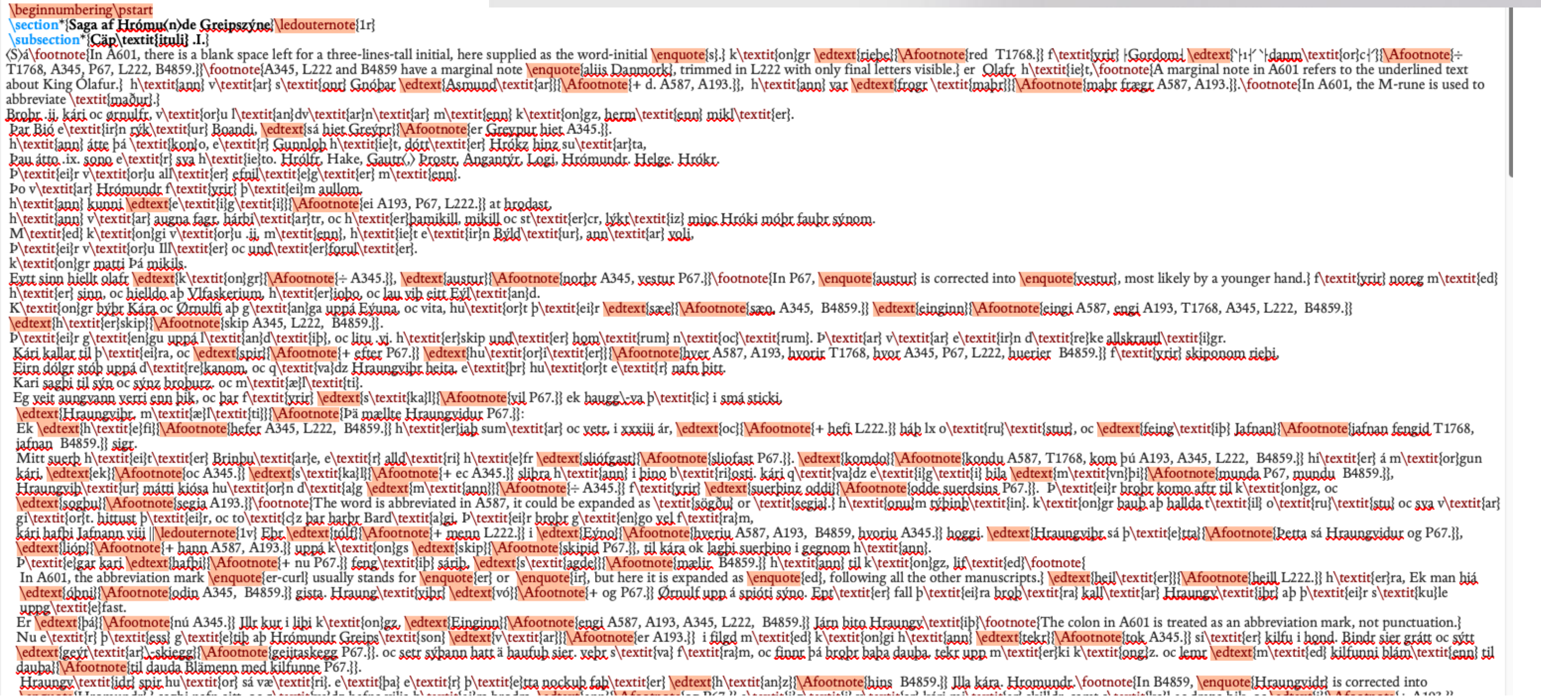
The Saga of Hrómundur and Its Manuscript Transmission

Open Access Book

Digital Editions

Physical Copies

This poster is built on the template by Graphicsfuel.com (https://www.graphicsfuel.com/hardcover-book-mockup-2/)



```
\beginnumbering\pstart
\section*{Saga af Hrómu(n)de Greipszýne}\ledouternote{1r}
\subsection*{Cäp\textit{ituli} .I.}
(S)á\footnote{In A601, there is a blank space left for a three-lines-tall initial, here supplied as the word-initial \enquote{s}.} k\textit{on}gr \edtext{rieþe}{\Afootnote{red  T1768.}} f\textit{yrir} ⊦Gordom⊦ \edtext{`⊦ı⊦´ `⊦danm\textit{or}c⊦´}{\Afootnote{÷
T1768, A345, P67, L222, B4859.}}\footnote{A345, L222 and B4859 have a marginal note \enquote{aliis Danmork}, trimmed in L222 with only final letters visible.} er  Olafr  h\textit{ie}t,\footnote{A marginal note in A601 refers to the underlined text
about King Ólafur.}  h\textit{ann} v\textit{ar} s\textit{onr} Gnóþar \edtext{Asmund\textit{ar}}{\Afootnote{+ d. A587, A193.}},  h\textit{ann} var \edtext{frogr \textit{maþr}}{\Afootnote{maþr frægr A587, A193.}}.\footnote{In A601, the M-rune is used to
abbreviate \textit{maður}.}
Broþr .ij. kári oc ørnulfr, v\textit{or}u l\textit{an}dv\textit{ar}n\textit{ar} m\textit{enn} k\textit{on}gz, herm\textit{enn} mikl\textit{er}.
 Þar Bió e\textit{ir}n rýk\textit{ur} Boandi, \edtext{sá hiet Greýpr}{\Afootnote{er Greypur hiet A345.}}.
 h\textit{ann} átte þá \textit{kon}o, e\textit{r} Gunnloþ h\textit{ie}t, dótt\textit{er} Hrókz hinz su\textit{ar}ta,
 Þau átto .ix. sono e\textit{r} sva h\textit{ie}to. Hrólfr, Hake, Gautr(,) Þrostr, Angantýr, Logi, Hrómundr. Helge. Hrókr.
 Þ\textit{ei}r v\textit{or}u all\textit{er} efnil\textit{e}g\textit{er} m\textit{enn}.
 Þo v\textit{ar} Hrómundr f\textit{yrir} þ\textit{ei}m aullom.
 h\textit{ann} kunni \edtext{e\textit{i}g\textit{i}}{\Afootnote{ei A193, P67, L222.}} at hrodast,
 h\textit{ann} v\textit{ar} augna fagr, hárbi\textit{ar}tr, oc h\textit{er}þamikill, mikill oc st\textit{er}cr, lýkt\textit{iz} mioc Hróki móþr fauþr sýnom.
 M\textit{ed} k\textit{on}gi v\textit{or}u .ij. m\textit{enn}, h\textit{ie}t e\textit{ir}n Býld\textit{ur}, ann\textit{ar} voli,
 Þ\textit{ei}r v\textit{or}u Ill\textit{er} oc und\textit{er}foru\textit{er}.
 k\textit{on}gr matti Þá mikils.
 Eytt sinn hiellt olafr \edtext{k\textit{on}gr}{\Afootnote{÷ A345.}}, \edtext{austur}{\Afootnote{norþr A345, vestur P67.}}\footnote{In P67, \enquote{austur} is corrected into \enquote{vestur}, most likely by a younger hand.} f\textit{yrir} noreg m\textit{ed}
 h\textit{er} sinn, oc hielldo aþ Vlfaskerium, h\textit{er}ioþo, oc lau viþ eitt Eýl\textit{an}d.
 K\textit{on}gr býþr Kára oc Ørnulfi aþ g\textit{an}ga uppá Eýuna, oc vita, hu\textit{or}t þ\textit{ei}r \edtext{sæe}{\Afootnote{sæo, A345,  B4859.}} \edtext{einginn}{\Afootnote{eingi A587, engi A193, T1768, A345, L222,  B4859.}}
 \edtext{h\textit{er}skip}{\Afootnote{skip A345, L222,  B4859.}}.
 Þ\textit{ei}r g\textit{en}gu uppá l\textit{an}d\textit{iþ}, oc litu .vj. h\textit{er}skip und\textit{er} hom\textit{rum} n\textit{oc}\textit{rum}. Þ\textit{ar} v\textit{ar} e\textit{ir}n d\textit{re}ke allskrautl\textit{i}gr.
  Kári kallar til þ\textit{ei}ra, oc \edtext{spir}{\Afootnote{+ efter P67.}} \edtext{hu\textit{or}i\textit{er}}{\Afootnote{hver A587, A193, hvorir T1768, hvor A345, P67, L222, huerier  B4859.}} f\textit{yrir} skiponom rieþi,
  Eirn dólgr stóþ uppá d\textit{re}kanom, oc q\textit{va}dz Hraungviþr heita. e\textit{þr} hu\textit{or}t e\textit{r} nafn þitt.
  Kari sagþi til sýn oc sýnz broþurz. oc m\textit{æ}l\textit{ti}.
  Eg veit aungvann verri enn þik, oc þar f\textit{yrir} \edtext{s\textit{ka}l}{\Afootnote{vil P67.}} ek haugg\-va þ\textit{ic} i smá sticki,
  \edtext{Hraungviþr. m\textit{æ}l\textit{ti}}{\Afootnote{Þä mællte Hraungvidur P67.}}:
  Ek \edtext{h\textit{e}fi}{\Afootnote{hefer A345, L222,  B4859.}} h\textit{er}iaþ sum\textit{ar} oc vetr, i xxxiij ár, \edtext{oc}{\Afootnote{+ hefi L222.}} háþ lx o\textit{ru}\textit{stur}, oc \edtext{feing\textit{iþ} Jafnan}{\Afootnote{jafnan fengid T1768,
  jafnan  B4859.}} sigr.
  Mitt suerþ h\textit{ei}t\textit{er} Brinþu\textit{ar}e, e\textit{r} alld\textit{ri} h\textit{e}fr \edtext{sliófgast}{\Afootnote{sliofast P67.}}. \edtext{komdo}{\Afootnote{kondu A587, T1768, kom þú A193, A345, L222,  B4859.}} hi\textit{er} á m\textit{or}gun
  kári, \edtext{ek}{\Afootnote{oc A345.}} \edtext{s\textit{ka}l}{\Afootnote{+ ec A345.}} slíþra h\textit{ann} i þino b\textit{ri}osti. kári q\textit{va}dz e\textit{i}g\textit{i} bila \edtext{m\textit{vn}þi}{\Afootnote{munda P67, mundu  B4859.}},
  Hraungvið\textit{ur} mátti kiósa hu\textit{or}n d\textit{a}g \edtext{m\textit{ann}}{\Afootnote{÷ A345.}} f\textit{yrir} \edtext{suerþinz oddi}{\Afootnote{odde suerdsins P67.}}.  Þ\textit{ei}r broþr komo aftr til k\textit{on}gz, oc
  \edtext{sogþu}{\Afootnote{segia A193.}\footnote{The word is abbreviated in A587, it could be expanded as \textit{sögðu} or \textit{segja}.} h\textit{onu}m týþinþ\textit{in}. k\textit{on}gr bauþ aþ hallda t\textit{il} o\textit{ru}\textit{stu} oc sva v\textit{ar}
  gi\textit{or}t. hittust þ\textit{ei}r, oc to\textit{c}z þar harþr Bard\textit{a}gi. Þ\textit{ei}r broþr g\textit{en}go vel f\textit{ra}m,
  kári hafþi Jafnann viij ||\ledouternote{1v} Ebr. \edtext{tólf}{\Afootnote{+ menn L222.}} i \edtext{Eýno}{\Afootnote{hveriu A587, A193,  B4859, hvoriu A345.}} hoggi. \edtext{Hraungviþr sá þ\textit{e}tta}{\Afootnote{Þetta sá Hraungvidur og P67.}},
  \edtext{lióp}{\Afootnote{+ hann A587, A193.}} uppá k\textit{on}gs \edtext{skip}{\Afootnote{skipid P67.}}, til kára ok lagþi suerþino i gegnom h\textit{ann}.
  Þ\textit{e}gar kari \edtext{hafþi}{\Afootnote{+ nu P67.}} feng\textit{iþ} sárib, \edtext{s\textit{agde}}{\Afootnote{mælir.  B4859.}} h\textit{ann} til k\textit{on}gz, lif\textit{ed}\footnote{
  In A601, the abbreviation mark \enquote{er-curl} usually stands for \enquote{er} or  \enquote{ir}, but here it is expanded as \enquote{ed}, following all the other manuscripts.} \edtext{heil\textit{er}}{\Afootnote{heill L222.}} h\textit{er}ra, Ek man hiá
  \edtext{óþni}{\Afootnote{odin A345,  B4859.}} gista. Hraung\textit{viþr} \edtext{vó}{\Afootnote{+ og P67.}} Ørnulf upp á spióti sýno. Ept\textit{er} fall þ\textit{ei}ra broþ\textit{ra} kall\textit{ar} Hraungv\textit{iþr} aþ þ\textit{ei}r s\textit{ku}le
  uppg\textit{e}fast.
 Er \edtext{þá}{\Afootnote{nú A345.}} Illr kur i liþi k\textit{on}gz, \edtext{Einginn}{\Afootnote{engi A587, A193, A345, L222,  B4859.}} Járn bito Hraungv\textit{iþ}\footnote{The colon in A601 is treated as an abbreviation mark, not punctuation.}
 Nu e\textit{r} þ\textit{ess} g\textit{e}tiþ aþ Hrómundr Greips\textit{son} \edtext{v\textit{ar}}{\Afootnote{er A193.}}  i filgd m\textit{ed} k\textit{on}gi h\textit{ann} \edtext{tekr}{\Afootnote{tok A345.}} sí\textit{er} kilfu i hond. Bindr sier grátt oc sýtt
  \edtext{geýt\textit{ar}\-skiegg}{\Afootnote{geitaskegg P67.}}. oc setr sýþann hatt ä haufuþ sier. veþr s\textit{va} f\textit{ra}m, oc finnr þá broþr baþa dauþa. tekr upp m\textit{er}ki k\textit{ong}z. oc lemr \edtext{m\textit{ed} kilfunni blám\textit{enn} til
  dauþa}{\Afootnote{til dauda Blämenn med kilfunne P67.}}.
  Hraungv\textit{idr} spir hu\textit{or} sá væ\textit{ri}. e\textit{þa} e\textit{r} þ\textit{e}tta nockuþ fаþ\textit{er} \edtext{h\textit{an}z}{\Afootnote{hins  B4859.}} Illa kára. Hromundr.\footnote{In B4859, \enquote{Hraungvidr} is corrected into
```

Kapitan, Modelling humanities data with TEI-XML

HTML (Hypertext Markup Language)

- HTML is the standard markup language for web pages.
- Markup:
  - **<tagname> Content goes here... </tagname>**
  - Start Tag --- Content --- End Tag

```
<!DOCTYPE html>
<html>
  <head>
    <title> This is my title</title>
  </head>
  <body>
    <p>This is a paragraph</p>
  </body>
</html>
```

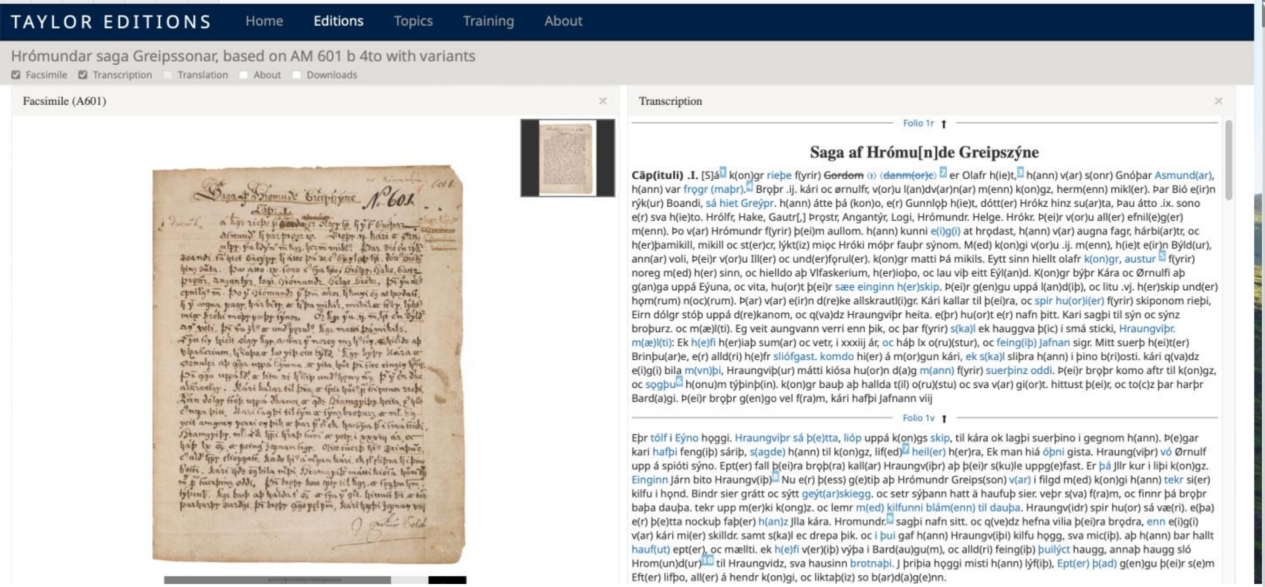
Basic elements:

- <html> – root element of an HTML page
- <head> – contains meta information about the document
- <body> – contains the visible page content

Other elements (examples):

- <div> - division/section of the page
- <table> - table
- <img> - image
- <ul> – list
- <h1> – heading
- <p> – paragraph

HTML is everywhere !!!



TAYLOR EDITIONS Home Editions Topics Training About

Hrómundar saga Greipssonar, based on AM 601 b 4to with variants

☑ Facsimile ☑ Transcription □ Translation □ About □ Downloads

Facsimile (A601)

Transcription

Folio 1r

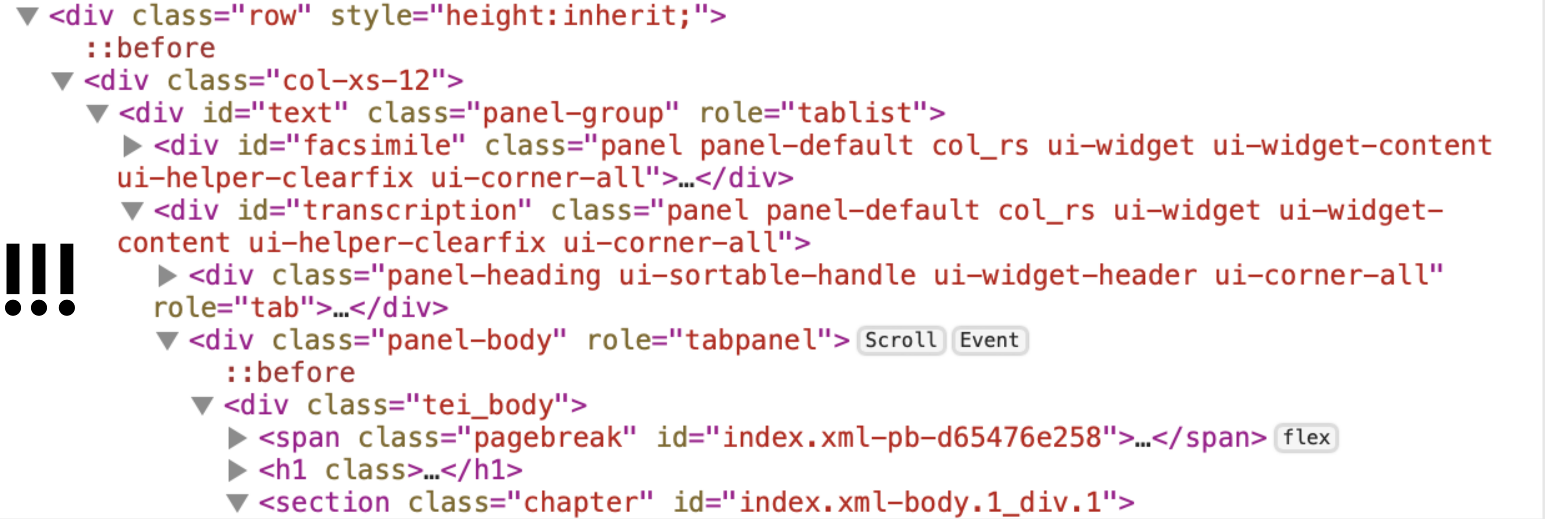
Saga af Hrómu[n]de Greipszýne

Cäp(ituli) .I. [S]á[1] k(on)gr rieþe f(yrir) ~~Gordom~~ ‹ı› ‹~~danm(or)c~~›[2] er Olafr h(ie)t,[3] h(ann) v(ar) s(onr) Gnóþar Asmund(ar), h(ann) var frọgr (maþr).[4] Brọþr .ij. kári oc ørnulfr, v(or)u l(an)dv(ar)n(ar) m(enn) k(on)gz, herm(enn) mikl(er). Þar Bió e(ir)n rýk(ur) Boandi, sá hiet Greýpr. h(ann) átte þá (kon)o, e(r) Gunnlọþ h(ie)t, dótt(er) Hrókz hinz su(ar)ta, Þau átto .ix. sono e(r) sva h(ie)to. Hrólfr, Hake, Gautr[,] Þrọstr, Angantýr, Logi, Hrómundr. Helge. Hrókr. Þ(ei)r v(or)u all(er) efnil(e)g(er) m(enn). Þo v(ar) Hrómundr f(yrir) þ(ei)m aullom. h(ann) kunni e(i)g(i) at hrọdast, h(ann) v(ar) augna fagr, hárbi(ar)tr, oc h(er)þamikill, mikill oc st(er)cr, lýkt(iz) miọc Hróki móþr fauþr sýnom. M(ed) k(on)gi v(or)u .ij. m(enn), h(ie)t e(ir)n Býld(ur), ann(ar) voli, Þ(ei)r v(or)u Ill(er) oc und(er)fọrul(er). k(on)gr matti Þá mikils. Eytt sinn hiellt olafr k(on)gr, austur[5] f(yrir) noreg m(ed) h(er) sinn, oc hielldo aþ Vlfaskerium, h(er)ioþo, oc lau viþ eitt Eýl(an)d. K(on)gr býþr Kára oc Ørnulfi aþ g(an)ga uppá Eýuna, oc vita, hu(or)t þ(ei)r sæe einginn h(er)skip. Þ(ei)r g(en)gu uppá l(an)d(iþ), oc litu .vj. h(er)skip und(er) họm(rum) n(oc)(rum). Þ(ar) v(ar) e(ir)n d(re)ke allskrautl(i)gr. Kári kallar til þ(ei)ra, oc spir hu(or)i(er) f(yrir) skiponom rieþi, Eirn dólgr stóþ uppá d(re)kanom, oc q(va)dz Hraungviþr heita. e(þr) hu(or)t e(r) nafn þitt. Kari sagþi til sýn oc sýnz broþurz. oc m(æ)l(ti). Eg veit aungvann verri enn þik, oc þar f(yrir) s(ka)l ek hauggva þ(ic) i smá sticki, Hraungviþr. m(æ)l(ti): Ek h(e)fi h(er)iaþ sum(ar) oc vetr, i xxxiij ár, oc háþ lx o(ru)(stur), oc feing(iþ) Jafnan sigr. Mitt suerþ h(ei)t(er) Brinþu(ar)e, e(r) alld(ri) h(e)fr sliófgast. komdo hi(er) á m(or)gun kári, ek s(ka)l sliþra h(ann) i þino b(ri)osti. kári q(va)dz e(i)g(i) bila m(vn)þi, Hraungviþ(ur) mátti kiósa hu(or)n d(a)g m(ann) f(yrir) suerþinz oddi. Þ(ei)r brọþr komo aftr til k(on)gz, oc sọgþu[6] h(onu)m týþinþ(in). k(on)gr bauþ aþ hallda t(il) o(ru)(stu) oc sva v(ar) gi(or)t. hittust þ(ei)r, oc to(c)z þar harþr Bard(a)gi. Þ(ei)r brọþr g(en)go vel f(ra)m, kári hafþi Jafnann viij

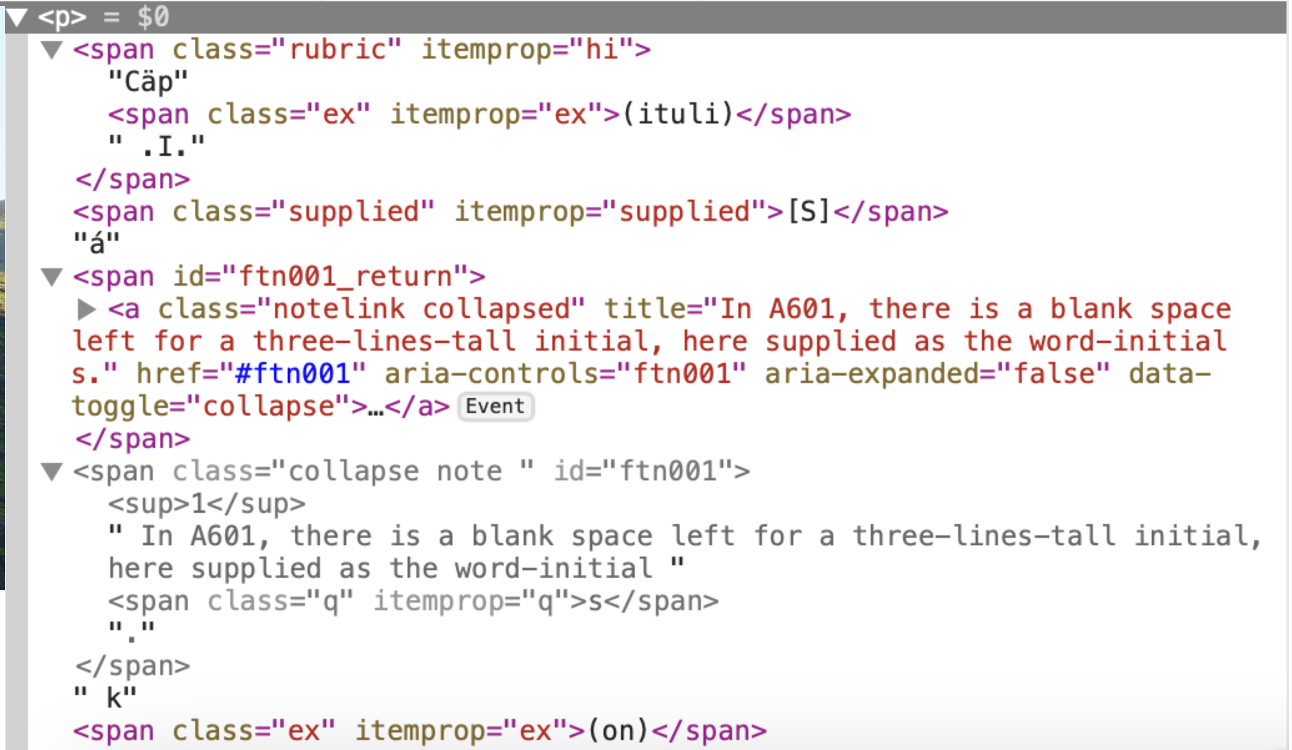
Folio 1v

Eþr tólf i Eýno họggi. Hraungviþr sá þ(e)tta, lióp uppá k(on)gs skip, til kára ok lagþi suerþino i gegnom h(ann). Þ(e)gar kari hafþi feng(iþ) sáriþ, s(agde) h(ann) til k(on)gz, lif(ed)[7] heil(er) h(er)ra, Ek man hiá óþni gista. Hraung(viþr) vó Ørnulf upp á spióti sýno. Ept(er) fall þ(ei)ra brọþ(ra) kall(ar) Hraungv(iþr) aþ þ(ei)r s(ku)le uppg(e)fast. Er þá Jllr kur i liþi k(on)gz. Einginn Járn bito Hraungv(iþ)[8] Nu e(r) þ(ess) g(e)tiþ aþ Hrómundr Greips(son) v(ar) i filgd m(ed) k(on)gi h(ann) tekr si(er) kilfu i họnd. Bindr sier grátt oc sýtt geýt(ar)skiegg. oc setr sýþann hatt ä haufuþ sier. veþr s(va) f(ra)m, oc finnr þá brọþr baþa dauþa. tekr upp m(er)ki k(ong)z. oc lemr m(ed) kilfunni blám(enn) til dauþa. Hraungv(idr) spir hu(or) sá væ(ri). e(þa) e(r) þ(e)tta nockuþ faþ(er) h(an)z Jlla kára. Hromundr.[9] sagþi nafn sitt. oc q(ve)dz hefna vilia þ(ei)ra brọdra, enn e(i)g(i) v(ar) kári mi(er) skilldr. samt s(ka)l ec drepa þik. oc i þui gaf h(ann) Hraungv(iþi) kilfu họgg, sva mic(iþ). aþ h(ann) bar hallt hauf(ut) ept(er), oc mællti. ek h(e)fi v(er)(iþ) výþa i Bard(au)gu(m), oc alld(ri) feing(iþ) þuilýct haugg, annaþ haugg sló Hrom(un)d(ur)[10] til Hraungvidz, sva hausinn brotnaþi. J þriþia họggi misti h(ann) lýf(iþ), Ept(er) þ(ad) g(en)gu þ(ei)r s(e)m Eft(er) lifþo, all(er) á hendr k(on)gi, oc liktaþ(iz) so b(ar)d(a)g(e)nn.

https://editions.mml.ox.ac.uk/editions/hromundar_A601/



```
▼ <div class="row" style="height:inherit;">
    ::before
  ▼ <div class="col-xs-12">
    ▼ <div id="text" class="panel-group" role="tablist">
      ▶ <div id="facsimile" class="panel panel-default col_rs ui-widget ui-widget-content
      ui-helper-clearfix ui-corner-all">…</div>
      ▼ <div id="transcription" class="panel panel-default col_rs ui-widget ui-widget-
      content ui-helper-clearfix ui-corner-all">
        ▶ <div class="panel-heading ui-sortable-handle ui-widget-header ui-corner-all"
        role="tab">…</div>
        ▼ <div class="panel-body" role="tabpanel"> Scroll Event
            ::before
          ▼ <div class="tei_body">
            ▶ <span class="pagebreak" id="index.xml-pb-d65476e258">…</span> flex
            ▶ <h1 class>…</h1>
            ▼ <section class="chapter" id="index.xml-body.1_div.1">
```



```
▼ <p> = $0
  ▼ <span class="rubric" itemprop="hi">
      "Cäp"
      <span class="ex" itemprop="ex">(ituli)</span>
      " .I."
    </span>
    <span class="supplied" itemprop="supplied">[S]</span>
    "á"
  ▼ <span id="ftn001_return">
    ▶ <a class="notelink collapsed" title="In A601, there is a blank space
    left for a three-lines-tall initial, here supplied as the word-initial
    s." href="#ftn001" aria-controls="ftn001" aria-expanded="false" data-
    toggle="collapse">…</a> Event
    </span>
  ▼ <span class="collapse note " id="ftn001">
      <sup>1</sup>
      " In A601, there is a blank space left for a three-lines-tall initial,
      here supplied as the word-initial "
      <span class="q" itemprop="q">s</span>
      "."
    </span>
    " k"
    <span class="ex" itemprop="ex">(on)</span>
```

Clear separation of content and presentation

```
<div>
    <h2>Short Description</h2>
    <p>AM 30 fol. is a paper manuscript in folio format comprising 56 leaves
gathered into five quires of six conjoint leaves each, with the exception of
the last quire, with consists of four conjoint leaves only. The whole codex is
made of one type of relatively thick laid paper with a "bull's head"
watermark.
    </p>
</div>
```

Clear separation of content and presentation

HTML + CSS (Cascading Style Sheets)

- HTML

```
<h1> My Text </h1>
```

- CSS

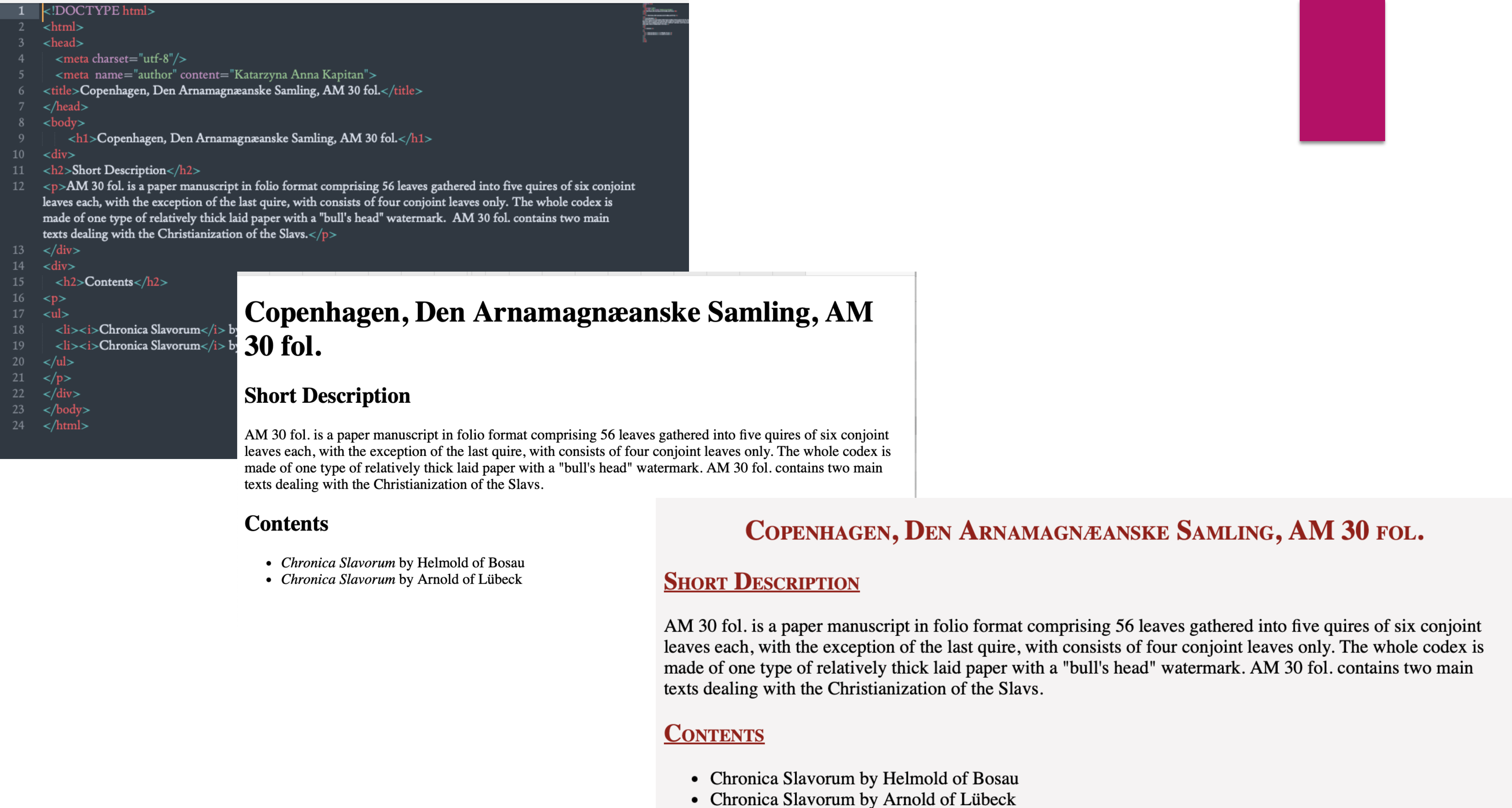
```
h1 {
    text-decoration: underline;
    text-align: center;
}
```

Exercise 2 (Week1/Exercises/Ex2)

- Open AM_30_fol.html with your browser – take a screenshot of the view
- Open AM_30_fol.html with Sublime Text
- Associate style.css with your html file by adding within <head> the element <link>:

```
<link rel="stylesheet" type="text/css" href="style.css">
```

- Save the file.
- Open AM_30_fol.html in a browser – take a screenshot of the view.



```
<!DOCTYPE html>
<html>
<head>
  <meta charset="utf-8"/>
  <meta  name="author" content="Katarzyna Anna Kapitan">
<title>Copenhagen, Den Arnamagnæanske Samling, AM 30 fol.</title>
</head>
<body>
    <h1>Copenhagen, Den Arnamagnæanske Samling, AM 30 fol.</h1>
<div>
<h2>Short Description</h2>
<p>AM 30 fol. is a paper manuscript in folio format comprising 56 leaves gathered into five quires of six conjoint leaves each, with the exception of the last quire, with consists of four conjoint leaves only. The whole codex is made of one type of relatively thick laid paper with a "bull's head" watermark.  AM 30 fol. contains two main texts dealing with the Christianization of the Slavs.</p>
</div>
<div>
  <h2>Contents</h2>
<p>
<ul>
  <li><i>Chronica Slavorum</i> by
  <li><i>Chronica Slavorum</i> by
</ul>
</p>
</div>
</body>
</html>
```

Copenhagen, Den Arnamagnæanske Samling, AM 30 fol.

Short Description

AM 30 fol. is a paper manuscript in folio format comprising 56 leaves gathered into five quires of six conjoint leaves each, with the exception of the last quire, with consists of four conjoint leaves only. The whole codex is made of one type of relatively thick laid paper with a "bull's head" watermark. AM 30 fol. contains two main texts dealing with the Christianization of the Slavs.

Contents

- *Chronica Slavorum* by Helmold of Bosau
- *Chronica Slavorum* by Arnold of Lübeck

Copenhagen, Den Arnamagnæanske Samling, AM 30 fol.

Short Description

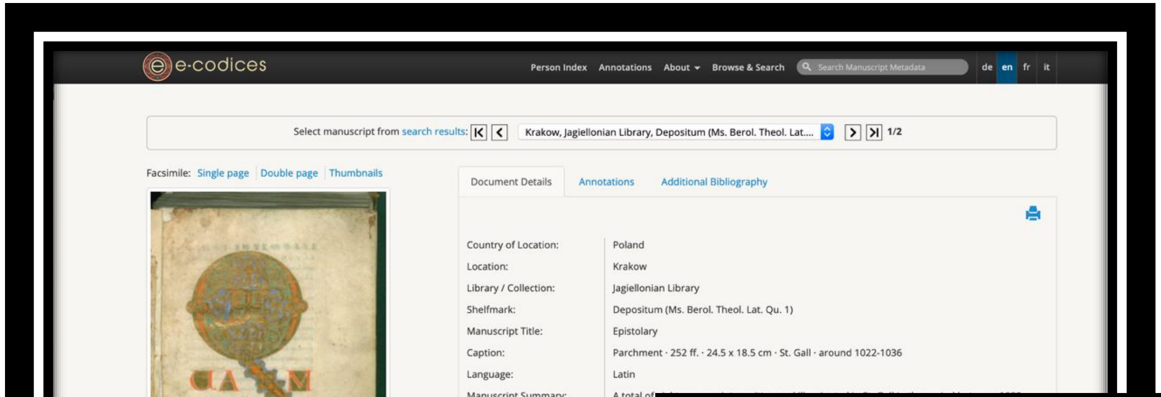
AM 30 fol. is a paper manuscript in folio format comprising 56 leaves gathered into five quires of six conjoint leaves each, with the exception of the last quire, with consists of four conjoint leaves only. The whole codex is made of one type of relatively thick laid paper with a "bull's head" watermark. AM 30 fol. contains two main texts dealing with the Christianization of the Slavs.

Contents

- Chronica Slavorum by Helmold of Bosau
- Chronica Slavorum by Arnold of Lübeck

Markup is everywhere

Behind the e-codices viewer: HTML & XML

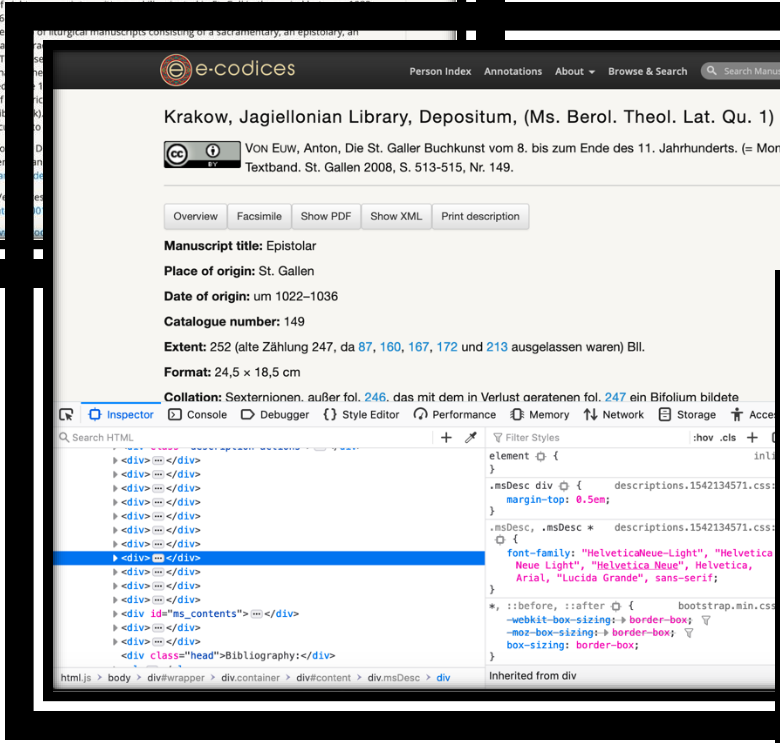


Select manuscript from search results: Krakow, Jagiellonian Library, Depositum (Ms. Berol. Theol. Lat.... 1/2

Facsimile: Single page | Double page | Thumbnails

Document Details | Annotations | Additional Bibliography

Country of Location:	Poland
Location:	Krakow
Library / Collection:	Jagiellonian Library
Shelfmark:	Depositum (Ms. Berol. Theol. Lat. Qu. 1)
Manuscript Title:	Epistolary
Caption:	Parchment · 252 ff. · 24.5 x 18.5 cm · St. Gall · around 1022-1036
Language:	Latin
Manuscript Summary:	A total of … and 1036 … complete … of liturgical manuscripts consisting of a sacramentary, an epistolary, an evangelia… *missae*. T… Unfortun… described… library of… (Staatsbib…k). was evac…
Standard description:	Euw Anto… Monaster… Show sta…
DOI (Digital Object Identifier):	10.5076/… Theol-Lat…
Permanent link:	https://w…



Krakow, Jagiellonian Library, Depositum, (Ms. Berol. Theol. Lat. Qu. 1)

VON EUW, Anton, Die St. Galler Buchkunst vom 8. bis zum Ende des 11. Jahrhunderts. (= Monasterium Sancti Galli, Bd. 3) Band I: Textband. St. Gallen 2008, S. 513-515, Nr. 149.

Overview | Facsimile | Show PDF | Show XML | Print description

Manuscript title: Epistolar

Place of origin: St. Gallen

Date of origin: um 1022–1036

Catalogue number: 149

Extent: 252 (alte Zählung 247, da 87, 160, 167, 172 und 213 ausgelassen waren) Bll.

Format: 24,5 × 18,5 cm

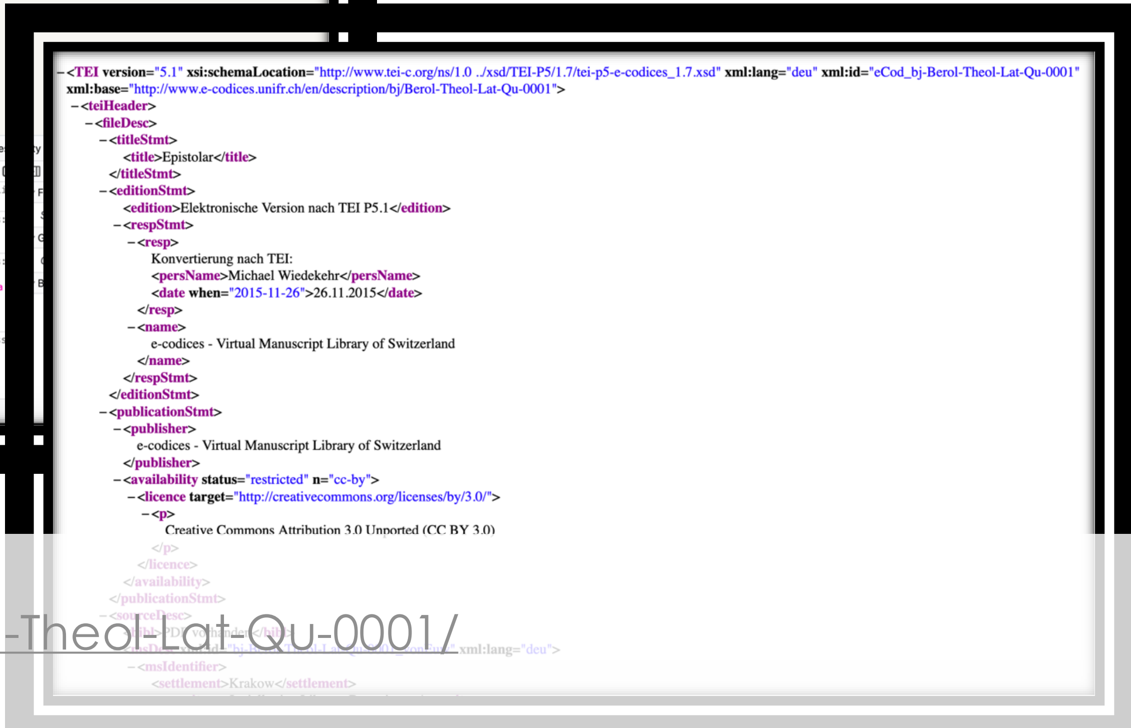
Collation: Sexternionen, außer fol. 246, das mit dem in Verlust geratenen fol. 247 ein Bifolium bildete

```
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div>⋯</div>
<div id="ms_contents">⋯</div>
<div>⋯</div>
<div>⋯</div>
<div class="head">Bibliography:</div>
```

html.js > body > div#wrapper > div.container > div#content > div.msDesc > div

```
element {
}
.msDesc div {          descriptions.1542134571.css:
  margin-top: 0.5em;
}
.msDesc, .msDesc *     descriptions.1542134571.css:
 {
  font-family: "HelveticaNeue-Light", "Helvetica
   Neue Light", "Helvetica Neue", Helvetica,
   Arial, "Lucida Grande", sans-serif;
}
*, ::before, ::after {         bootstrap.min.css
  ~~-webkit-box-sizing: border-box;~~
  ~~-moz-box-sizing: border-box;~~
  box-sizing: border-box;
}
```

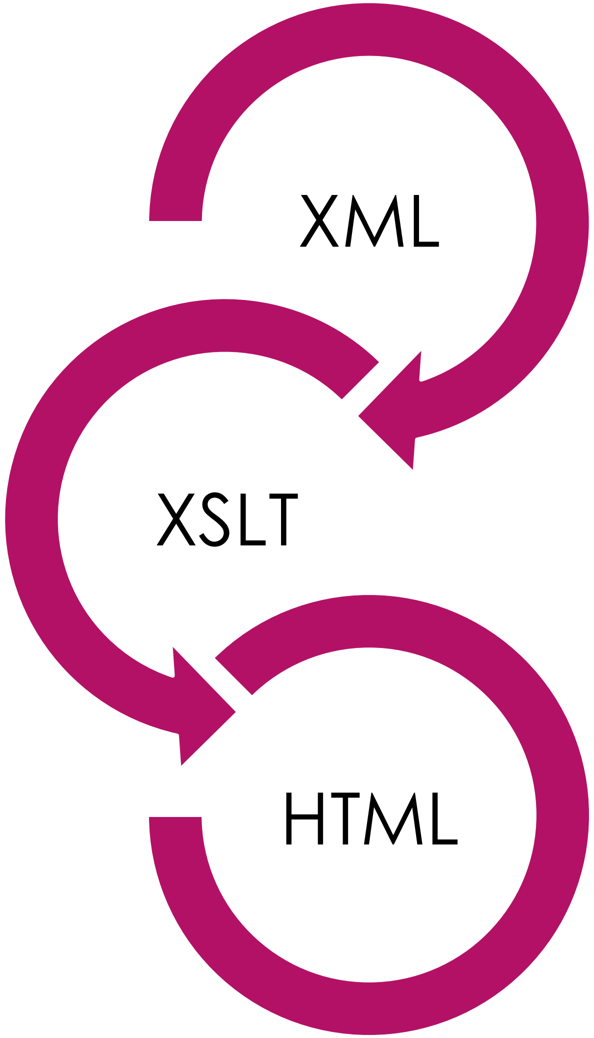
Inherited from div



```
-<TEI version="5.1" xsi:schemaLocation="http://www.tei-c.org/ns/1.0 ../xsd/TEI-P5/1.7/tei-p5-e-codices_1.7.xsd" xml:lang="deu" xml:id="eCod_bj-Berol-Theol-Lat-Qu-0001"
 xml:base="http://www.e-codices.unifr.ch/en/description/bj/Berol-Theol-Lat-Qu-0001">
  -<teiHeader>
    -<fileDesc>
      -<titleStmt>
          <title>Epistolar</title>
        </titleStmt>
      -<editionStmt>
          <edition>Elektronische Version nach TEI P5.1</edition>
        -<respStmt>
          -<resp>
              Konvertierung nach TEI:
              <persName>Michael Wiedekehr</persName>
              <date when="2015-11-26">26.11.2015</date>
            </resp>
          -<name>
              e-codices - Virtual Manuscript Library of Switzerland
            </name>
          </respStmt>
        </editionStmt>
      -<publicationStmt>
        -<publisher>
            e-codices - Virtual Manuscript Library of Switzerland
          </publisher>
        -<availability status="restricted" n="cc-by">
          -<licence target="http://creativecommons.org/licenses/by/3.0/">
            -<p>
                Creative Commons Attribution 3.0 Unported (CC BY 3.0)
              </p>
            </licence>
          </availability>
        </publicationStmt>
      -<sourceDesc>
          <bibl>PDF vorhanden</bibl>
          <msDesc xml:id="bj-Berol-Theol-Lat-Qu-0001_vonEuw" xml:lang="deu">
          -<msIdentifier>
              <settlement>Krakow</settlement>
```

Kraków, Jagiellonian Library, Ms. Berol. Theol. Lat. Qu. 1
Source: https://www.e-codices.unifr.ch/en/list/one/bj/Berol-Theol-Lat-Qu-0001/

From Data to Display



Structured Data

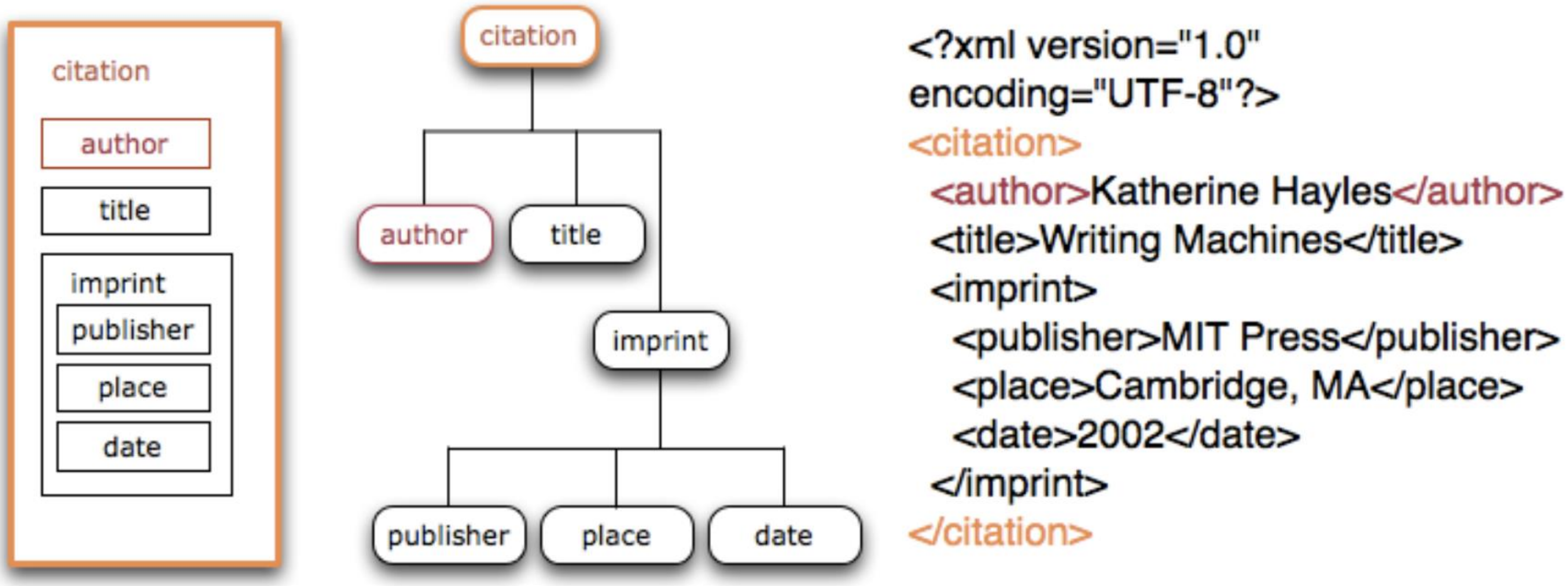
Transformation Scenarios

Basic Display

XML
(Extensible Markup Language)

XML (Extensible Markup Language)

- Storing structured data
- International standard, non-proprietary
- Standard text format (expressed in plain text)
- Easy to parse and read for computer programs.
- Widely used to export and share structured data.
- Hardware and software independent



```
<?xml version="1.0"
encoding="UTF-8"?>
<citation>
  <author>Katherine Hayles</author>
  <title>Writing Machines</title>
  <imprint>
    <publisher>MIT Press</publisher>
    <place>Cambridge, MA</place>
    <date>2002</date>
  </imprint>
</citation>
```

XML Elements & Tags

```
<!DOCTYPE html>
<html>
    <head>
        <title>This is a title</title>
    </head>
    <body>
        <p>This is a paragraph</p>
    </body>
</html>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<myRoot>
    <myContent>
        <content>
            Here is my content
        </content>
    </myContent>
</myRoot>
```

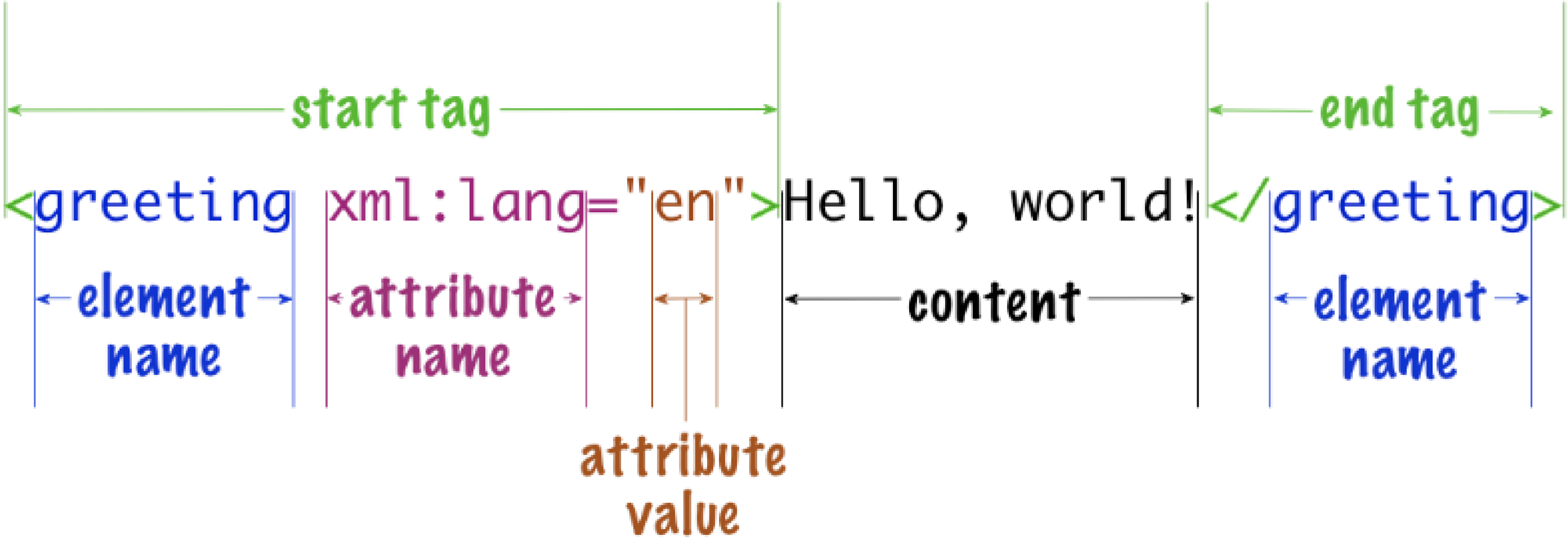
XML Elements & Tags

- Text is divided into elements (the nouns of the encoding — content objects).
- elements have start-tags and end-tags
  - <heading>My heading</heading>
- start-tags have < … >
  - <heading>
- end-tags have </ … >
  - </heading>

XML Attributes

- Attributes are adjectives of XML, they describe the properties of the elements
- any number of attributes can be specified on a given start-tag
- but only one with a given name

- **<person job="musician" age=”55">**Paul Simon**</person>**



Syd Bauman, Julia Flanders, and WWP, Creative Commons Attribution-ShareAlike 3.0 (Unported) license.

Quiz

I. Which of the following examples are well-formed XML?

1. <name type="person>Pearl S. Buck</name>
2. <name type="person">Toni Morrison<name>
3. <name="person">Carl Sagan</name>
4. <name type="person">Kurt Vonnegut</name>
5. <name type=person>John Cleese</name>
6. <name type="person"><forename>Frances</forename>
   <surname>Perkins</surname></name>

Course Materials:
https://github.com/KAKDH/TNAH_XML2025

Syllabus:
https://github.com/KAKDH/TNAH_XML2025/blob/main/Kapitan_TEI-XML_Syllabus.pdf

For Next Class:

1. Get your Git up and running
2. Get your Oxygen up and running