Talk 6: Manuscript Description

James Cummings

July 2009
This section introduces the TEI's module on manuscript description, briefly surveying the nature of manuscript description and the structural components of a TEI `<msDesc>` element.
Why are manuscript descriptions special?

- Manuscripts are *unique objects*, sometimes (though not always) of great cultural or political value.
- Books, by contrast, exist in multiple copies, and can be described adequately by well-established and formalized bibliographic conventions.
- For manuscripts, there are several traditions, often descriptive or *belle lettriste*, and little consensus.

Similar concerns apply to other text-bearing objects.
The TEI `<msDesc>` element is intended for several different kinds of applications:

- standalone database of library records (*finding aid*)
- discursive text collecting many records (*catalogue raisonné*)
- metadata component within a digital surrogate (*electronic edition*)
- tool for ‘quantitative codicology’
Catalogue Raisonné

An `<msDesc>` can appear anywhere a `<p>` paragraph can

```xml
<div>
  <head>The Lithuanian National Martynas Mazvydas Library and its records</head>
  <p>Probably the finest collection of ……</p>
  <p>For example:
    <msDesc xml:id="F101-19" xml:lang="en">
      <!-- …-->
    </msDesc>
  </p>
  <p>In the following manuscript…
    <msDesc xml:id="F101-21" xml:lang="en">
      <!-- …-->
    </msDesc>
  </p>
</div>
```
Quantitative Codicology: is it possible?

Two conflicting desires:

- preserve (or perpetuate) existing descriptive prose
- reliable search, retrieval, and analysis of data

The `<msDesc>` tries, wherever possible, to have its cake and eat it.
Components of a manuscript description

We separate, and tag differently, aspects concerned with...

- identification (<msIdentifier>)
- intellectual content (<msItem>)
- physical description (<physDesc>)
- history (<history>)
- additional curation information (<additional>)
- and descriptions of parts of a composite manuscript (<msPart>)
msDesc structure

<msDesc xml:id="ex2" xml:lang="en">
  <msIdentifier/>
  <!-- Repository location, shelfmarks, etc. -->
  </msIdentifier>
  <msContents>
  <!-- Structured description of MS contents -->
  </msContents>
  <physDesc>
  <!-- Physical and codicological description -->
  </physDesc>
  <history>
  <!-- Origin, provenance, acquisition, etc. -->
  </history>
  <additional>
  <!-- Additional bibliographic and curatorial information, and associated materials etc. -->
  </additional>
  <msPart>
  <!-- Composite manuscript details -->
  </msPart>
</msDesc>

<msIdentifier> is the only one that is required.
A basic `<msDesc>`

```xml
<msDesc xml:id="F101-21" xml:lang="en">
  <msIdentifier/>
  <repository>Lithuanian National M. Mazvydas Library, Rare Book and Manuscript Department</repository>
</msIdentifier>
<p>F101–21: [August 24], 1522. Ciuitatus Grodneñfıs (Grodno). Stored in Lithuania, Vilnius, Lithuanian National M. Mazvydas Library, Rare Book and Manuscript Department. "In nomine Domini amen. Ad perpetuam rei memoriam...". Latin. There are postscripts in Polish. Parchment. Single sheet format: 515mm × 354mm with 115mm fold-up. Text: single column, 415mm × 221mm, 33 ruled lines. The ornamented initial "I" takes up 30 lines, 15 cm. The attached seal has not remained (parchment fold-up is 11.5 cm, two slits for the string are spaced 12 cm apart). Signed by Sigifmundus rex – Sigismund the Old, King of Poland and Grand Duke of Lithuania. Summary: An original copy of a charter where Sigismund the Old (Sigifmundus), King of Poland and Grand Duke of Lithuania, transumes and confirms two deeds of purchase of the Raściai (Roffcze) domain in Veliuona (Vyelona) district by the Kédainiai prefect Georgij Andrichovich (Georgius Andrichowicz): the deed of purchase of December 13, 1518 from Kachna (Kachna) etc. in Kédainiai (Kveydanij) and the deed of purchase of April, 1518 from the Praczewicz (Praczewicz) family.
</p>
</msDesc>
```
A structured <msDesc> (1)

<msDesc xml:id="F101-21" xml:lang="en">
  <msIdentifier>
    <country>Lithuania</country>
    <settlement>Vilnius</settlement>
    <institution>Lithuanian National M. Mazvydas Library</institution>
    <repository>Rare Book and Manuscript Department</repository>
    <idno>F101-21</idno>
  </msIdentifier>
  <msContents>
    <summary>Sigismund the Old (Sigifmundus), King of Poland and Grand Duke of Lithuania, transumes and confirms two deeds of purchase of the Raściai (Roffcze) domain in Veliuona (Vyelona) district by the Kėdainiai prefect Georgij Andrichovich (Georgius Andrichowicz): the deed of purchase of December 13, 1518 from Kachna (Kachna) etc. in Kėdainiai (Kyeypadnij) and the deed of purchase of April, 1518 from the Praczewicz (Praczewicz) family. </summary>
    <textLang xml:lang="lat" mainLang="lat" otherLangs="pol">Latin. Postscripts on the reverse are in Polish.</textLang>
    <msItem>
      <incipit>In nomine Domini amen. Ad perpetuam rei memoriam…</incipit>
    </msItem>
  </msContents>
A structured <msDesc> (2)

<physDesc>
  <objectDesc form="leaf">
    <supportDesc material="perg">
      <support>Parchment. Single sheet format: 515 x 354 + 115 mm.</support>
    </supportDesc>
  </objectDesc>
  <layoutDesc>
    <layout columns="1" ruledLines="33">Text: 415 x 221 mm.</layout>
  </layoutDesc>
</physDesc>
<decoDesc>
  <decoNote>The ornamented initial "I" takes up 30 lines, 15 cm.</decoNote>
</decoDesc>
<sealDesc>
  <seal type="attached">
    <p>The seal has not remained (parchment fold-up is 11.5 cm, two slits for the string are spaced 12 cm apart). </p>
  </seal>
</sealDesc>
</physDesc>
A structured <msDesc> (3)

```xml
<history>
  <origin>
    <origDate when="1502-06-24">[August 24], 1522. </origDate>
    <origPlace xml:lang="lat">ciuitatus Grodnenfis</origPlace>
    <origPlace xml:lang="en">Grodno</origPlace>
    <p>Authenticity – original. </p>
    Signed by Sigismundus rex – Sigismund the Old, King of Poland and Grand Duke of Lithuania. </p>
  </origin>
</history>
</msDesc>
```
Identification (1)

The `<msIdentifier>`
Traditional three part specification:

- place `<country>`<country>France</country>, `<region>`<region>, `<settlement>`<settlement>Troyes</settlement`
- repository `<institution>`<institution>, `<repository>`<repository>Bibliothèque Municipale</repository`
- identifier `<collection>`<collection>, `<idno>`<idno>50</idno>`

`<msIdentifier>`
`<country>France</country>`
`<settlement>Troyes</settlement>`
`<repository>Bibliothèque Municipale</repository>`
`<idno>50</idno>`
`</msIdentifier>`
Identification (2)

Alternative or additional names can also be included:

<msIdentifier>
  <country>Danmark</country>
  <settlement>København</settlement>
  <repository>Det ArnamagnæanskeInstitut</repository>
  <idno>AM 45 fol.</idno>
  <msName xml:lang="la">Codex Frisianus</msName>
  <msName xml:lang="is">Fríssbók</msName>
</msIdentifier>
Intellectual Content

• May simply use paragraphs of text...
• ... or a tree of <msItem> elements
• ... optionally preceded by a prose summary

We can describe the content in general terms:

<msContents>
  <p>An extraordinary charivari of heroic deeds and improving tales, including an early version of</p>
  <title>Guy of Warwick</title> and several hymns.</p>
</msContents>

or we can provide detail about each distinct item:

<msContents>
  <summary>An extraordinary charivari of heroic deeds, improving tales, and hymns</summary>
  <msItem>
    <!-- details of Guy of Warwick here -->
  </msItem>
  <msItem>
    <!-- other items here -->
  </msItem>
</msContents>
Manuscripts contain identifiable items, usually physically tied to a *locus*.

- `<locus>`, if present, must be given first
- then any of the following, in a specified order:
  - `<author>`, `<respStmt>`
  - `<title>`, `<rubric>`, `<incipit>`, `<explicit>`, `<colophon>`, `<finalRubric>`
  - `<quote>`, `<textLang>`, `<decoNote>`, `<bibl>`, `<listBibl>`, `<note>` ...
  - ... or nested `<msItem>`s
<msContents> with multiple <msItem>s</msContents>

<msContents>
  <msItem n="1">
    <locus>fols. 5r-7v</locus>
    <title>An ABC</title>
    <bibl>
      <title>IMEV</title>
      <biblScope type="pages">239</biblScope>
    </bibl>
  </msItem>
  <msItem n="2">
    <locus>fols. 7v-8v</locus>
    <title xml:lang="fr">Lenvoy de Chaucer a Scogan</title>
  </msItem>
  <!-- ... -->
  <msItem n="6">
    <locus>fols. 14r-126v</locus>
    <title>Troilus and Criseyde</title>
    <note>Bk. 1:71-Bk. 5:1701, with additional losses due to mutilation throughout</note>
  </msItem>
</msContents>
Physical Description

An artificial (but helpful) grouping of many distinct items. You can simply supply paragraphs of prose, covering such topics as:

- `<objectDesc>`: the physical carrier
- `<handDesc>`: what is carried on it
- `<musicNotation>`, `<decoDesc>`, `<additions>`
- `<bindingDesc>` and `<sealDesc>`
- `<accMat>`: accompanying material

Or, group your discussion within the specific elements mentioned above.
Similarly, within the specific elements, you can supply paragraphs of prose, or further specific elements.
The carrier 1

The \texttt{<objectDesc>} contains just paragraphs, or \texttt{<supportDesc>} and \texttt{<layoutDesc>}

\begin{verbatim}
<objectDesc form="codex">
  <supportDesc material="mixed">
    <p>Early modern \texttt{<material>parchment</material>} and \texttt{<material>paper</material>}.</p>
  </supportDesc>
  <layoutDesc>
    <layout columns="1" ruledLines="25 32"/>
  </layoutDesc>
</objectDesc>
\end{verbatim}
A more complex substructure with specific elements for 
<support>, <extent>, <foliation>, <collation>, <condition>. 
Multiple layouts may also be specified:

```xml
<layoutDesc>
  <layout ruledLines="25" columns="1">
    <p>
      <locus from="1r-202v"/>
      <locus from="210r-212v"/>
      Between 25 and 32 ruled lines.
    </p>
  </layout>
  <layout ruledLines="34 50" columns="1">
    <p>
      <locus from="203r-209v"/> Between 34 and 50 ruled lines.
    </p>
  </layout>
</layoutDesc>
```
＜handDesc＞ and ＜decoDesc＞

- ＜handNote＞ (note on hand) describes a particular style or hand distinguished within a manuscript.
- ＜decoNote＞ contains a note describing either a decorative component of a manuscript, or a fairly homogenous class of such components.
The manuscript is written in two contemporary hands, otherwise unknown, but clearly those of practised scribes. Hand I writes ff. 1r-22v and hand II ff. 23 and 24. Some scholars, notably Verner Dahlerup and Hreinn Benediktsson, have argued for a third hand on f. 24, but the evidence for this is insubstantial.
The first part of the manuscript, <locus from="1v" to="72v:4">fols 1v-72v:4</locus>, is written in a practised Icelandic Gothic bookhand. This hand is not found elsewhere.

The second part of the manuscript, <locus from="72v:4" to="194v">fols 72v:4-194</locus>, is written in a hand contemporary with the first; it can also be found in a fragment of <title>Knýtlinga saga</title>, <ref>AM 20b II fol.</ref>.
The `<additions>` element can be used to list or describe any additions to the manuscript, such as marginalia, scribblings, doodles, etc., which are considered to be of interest or importance.

```
<additions>
  <p>The text of this manuscript is not interpolated with sentences from Royal decrees promulgated in 1294, 1305 and 1314. In the margins, however, another somewhat later scribe has added the relevant paragraphs of these decrees, see pp. 8, 24, 44, 47 etc.</p>
  <p>As a humorous gesture the scribe in one opening of the manuscript, pp. 36 and 37, has prolonged the lower stems of one letter f and five letters p and has them drizzle down the margin.</p>
</additions>
```
<accMat> (accompanying material) contains details of any significant additional material which may be closely associated with the manuscript being described, such as non-contemporaneous documents or fragments bound in with the manuscript at some earlier historical period.

<accMat> A copy of a tax form from 1947 is included in the envelope with the letter. It is not catalogued separately. </accMat>
<history>

- <origin>: where it all began
- <provenance>: everything in between
- <acquisition>: how you acquired it

<origin> is *datable* element and thus has attributes @notBefore and @notAfter, @when etc.
Example

<history>
  <origin>
    <p>Written in <origPlace>England</origPlace> in the <origDate notAfter="1300" notBefore="1200">13th cent.</origDate></p>
  </origin>
  <provenance>
    <p>On fol. 54v very faint is <q>Iste liber est fratris guillelmi de buria de Roberti ordinis fratrum Pred<expan>icatorum</expan>, 14th cent. (?): <q>hanauilla</q> is written at the foot of the page (15th cent.).</p>
  </provenance>
  <acquisition>
    <p>Bought from the rev. <name type="person">W. D. Macray</name> on <date when="1863-03-17">March 17, 1863</date>, for 1 pound 10s.</p>
  </acquisition>
</history>
<additional> information

- `<adminInfo>`: administrative information
- `<surrogates>`: information about other surrogates eg pictures
- `<accMat>`: accompanying material
- `<listBibl>`: bibliography
Administrative information

- record history
- availability
- custodial history
- miscellaneous remarks

<adminInfo>
  <custodialHist>
    <custEvent type="conservation" notBefore="1961-03" notAfter="1963-02">
      <p>Conserved between March 1961 and February 1963 at Birgitte Dalls Konserveringsværksted.</p>
    </custEvent>
    <custEvent type="photography" notBefore="1988-05-01" notAfter="1988-05-30">
      <p>Photographed in May 1988 by AMI/FA.</p>
    </custEvent>
  </custodialHist>
</adminInfo>
A `<msDesc>` can contain a nested `<msPart>`, catering for composite MSS, formerly distinct.

```xml
<msDesc xml:id="ex5" xml:lang="en">
  <msIdentifier>
    <msName xml:lang="la">Codex Suprasliensis</msName>
  </msIdentifier>
  <msPart>
    <altIdentifier type="partial">
      <settlement>Ljubljana</settlement>
      <repository>Narodna in univerzitetna knjiznica</repository>
      <idno>MS Kopitar 2</idno>
      <note>Contains ff. 10 to 42 only</note>
    </altIdentifier>
  </msPart>
  <msPart>
    <altIdentifier type="partial">
      <settlement>Warszawa</settlement>
      <repository>Biblioteka Narodowa</repository>
      <idno>B0 3.201</idno>
    </altIdentifier>
  </msPart>
</msDesc>
```
EU-funded eContent+ program to:

• Create seamless access to distributed information about manuscripts and rare old printed books in Europe
• Connect existing digital libraries, bring aboard those who do not have them
• Provide access to these in Manuscriptorium in their own languages and their own virtual interface
• To standardise metadata on a subset of TEI P5 XML
Workpackage Leaders

- National Library of the Czech Republic, Prague
- AiP Beroun, s r.o., Beroun, Czech Republic
- Oxford University Computing Services, Oxford, United Kingdom
- Centro per la comunicazione e l’integrazione dei media, Florence, Italy
- SYSTRAN S.A., Paris, France
- Institute of mathematics and informatics, Vilnius, Lithuania
- Biblioteca Nacional de España, Madrid, Spain

But many many more participating institutions and content partners!
Workpackages

- WP1: Project Management
- WP2: Preparation for system implementation and content enhancement
- **WP3: Standardization of shared metadata**
- WP4: User personalization
- WP5: Personalization for contributors
- WP6: Multilingual and sophisticated access
- WP7: Evaluation, testing, and validation
- WP8: Dissemination and exploitation
ENRICH: Workpackage 3

- Development and documentation of the ENRICH TEI P5 Specification
- Conversion between MASTER (and other formats) and TEI P5 / ENRICH ODD
- Implementation of the OAI harvester into existing Manuscriptorium platform
- Creation of an transformation web service: The ENRICH Garage Framework
- Improvement and generalization of use of Unicode treatment throughout interface
WP3 Deliverables

• D.3.1: Revised TEI-conformant specification
• D.3.2: Documentation and training materials for use with specification
• D.3.3: Report on development and validation of migration tools
• D.3.4: Report and Documentation on the ENRICH Garage Framework, and best practice with respect to handling of Unicode and non-Unicode data in Manuscriptorium and P5 conversion techniques
Work done so far...

- A comparison between MASTER, Manuscriptorium and TEI P5, with resolutions of inconsistencies for ENRICH
- An ENRICH ODD specification recording the customisation from the TEI, but also providing additional constraints, and detailed internationalised project documentation
- A XSLT stylesheet for transforming both MASTER and EAD to TEI P5 (ENRICH) and a test corpus of documents before and after migration
- Report on development and validation of migration tools
- A prototype gBank web service for documenting non-Unicode characters
- Highly detailed documentation and reference information for the ENRICH specification
- Extensive training materials and slides for those teaching local encoders how to use the specification
- See http://tei.oucs.ox.ac.uk/ENRICH/
Towards a European Digital Library of Manuscripts

ENRICH is a project funded under eContent+ programme whose main aim is to create seamless access to distributed information about manuscripts and rare old printed books in Europe on the Manuscriptorium platform.

Read more

Successful dissemination event of ENRICH project organized in Madrid

Submitted by jakubh on Wed, 06/11/2008 - 22:00.

The National Library of Spain held an official presentation of the Enrich Project within the Workshop on Rare Books Librarianship last Thursday 5 June at the Main Hall of the Library in Madrid. Tomas Pichlauer of AIP introduced the project among an audience of 180 library specialists and representatives from CERL and Fundacion Biblioteca Cervantes Virtual.

Read more

ENRICH Project Web Site evaluation form

Submitted by jakubh on Fri, 05/09/2008 - 09:04.

The evaluation of our project website can be done using the form prepared by IMI. All partners are encouraged to disseminate this evaluation form amongst all potentially involved. We will really appreciate your cooperation

Digital library readers demands/needs online survey