Manuscript Description and Facsimile

James Cummings

April 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.
Objectives

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

```xml
<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. Civitatus Grodnenfis (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 Sigismund 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 Rašciai (Roffeze) domain in Veliuona (Vyelona) district by the Kėdainiai prefect Georgij Andrichowich (Georgius Andrichowicz): the deed of purchase of December 13, 1518 from Kachna (Kachna) etc. in Kėdainiai (Kyeidanij) 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 (Sigismundus), King of Poland and Grand Duke of Lithuania, transumès and confirms two deeds of purchase of the Raščiai (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 (Kyeydanij) and the deed of purchase of April, 1518 from the Praczewicz (Praczewicz) family.</summary>
    <textLang xml:lang="lat" mainLang="lat" otherLanguages="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>
A structured `<msDesc>` (3)

```
<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>
    <p>Signed by Sigismundus rex – Sigismund the Old, King of Poland and Grand Duke of Lithuania.</p>
  </origin>
</history>
```
Identification (1)

The <msIdentifier>
Traditional three part specification:
  • place (<country>, <region>, <settlement>)
  • repository (<institution>, <repository>)
  • identifier (<collection>, <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:

```xml
<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>
```
Manuscript Description

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 <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>
<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>
</msContents>

<msContents>
<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 <objectDesc> contains just paragraphs, or <supportDesc> and <layoutDesc>

<objectDesc form="codex">
  <supportDesc material="mixed">
    <p>Early modern <material>parchment</material> and <material>paper</material>. </p>
  </supportDesc>
  <layoutDesc>
    <layout columns="1" ruledLines="25 32"/>
  </layoutDesc>
</objectDesc>
The carrier 2

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>
```
Working with facsimiles

<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, fols 1v-72v:4, is written in a practised Icelandic Gothic bookhand. This hand is not found elsewhere.

The second part of the manuscript, fols 72v:4-194, is written in a hand contemporary with the first; it can also be found in a fragment of Knýtlinga saga, AM 20b II fol.
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 þ 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 a 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</q> Roberti ordinis fratrum Pred<i>expan</i>icatorum</q>, 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 1pound 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>
And finally
A <msDesc> can contain a nested <msPart>, catering for composite MSS, formerly distinct.

<msDesc xml:id="ex5" xml:lang="en">
  <msIdentifier>
    <msName xml:lang="la">Codex Suprasiliensis</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>BO 3.201</idno>
    </altIdentifier>
  </msPart>
  <!-- More <msPart> elements -->
</msDesc>
Increasingly people want to do not just 'text' editions but text editions with facing page (or otherwise linked) facsimile images. Indeed, some people want to just have images and create and electronic facsimile (perhaps with a view to later eventual transcription). The <facsimile> element (a sibling of <teiHeader> and <text>) is provided to accommodate this desire.
Digital Facsimiles

- `<facsimile>` contains a representation of some written source in the form of a set of images rather than as transcribed or encoded text.
- `<surface>` defines a written surface in terms of a rectangular coordinate space.
  - `@start` points to an element which encodes the starting position of the text.
- `<zone>` defines a rectangular area contained within a `<surface>` element.
- Global `@facs (facsimile)` points directly to an image, or to a part of a facsimile element which corresponds with this element.
Simplest case: 1:1 mapping with @facs

If a digital text contains one image per page or column (or similar unit), and no more complex mapping between text and image is envisaged, then the @facs attribute may be used to point directly to a graphic resource.

```xml
<text>
  <body>
    <pb facs="147_13.jpg" n="13"/>
    <div>
      <head>ESSENCE OF PARLIAMENT. </head>
      <p> <!-- Page 13 continues --> </p>
    </div>
    <pb facs="147_14.jpg" n="14"/>
    <p>Business done. -- The Commons still... </p>
    <!-- Page 14 continues -->
  </div>
</body>
</text>
```
Using @facs in conjunction with <facsimile>, <surface>, and <zone>

Using these attributes and elements together enables an editor to
• associate multiple images with each page
• record arbitrary planar coordinates of textual elements on any kind of surface and link such elements to digital facsimile images of them
<facsimile>

The facsimile element is used to represent a digital facsimile. It appears within a TEI document along with, or instead of, the text element introduced in section 5 Default Text Structure. When this module is selected therefore, a legal TEI document may thus comprise any of the following:

• a TEI Header and a text element
• a TEI Header and a facsimile element
• a TEI Header, a facsimile element, and a text element
The facsimile element is used to represent a digital facsimile. It appears within a TEI document along with, or instead of, the text element introduced in section 5 Default Text Structure. When this module is selected therefore, a legal TEI document may thus comprise any of the following:

- a TEI Header and a text element
- a TEI Header and a facsimile element
- a TEI Header, a facsimile element, and a text element
The facsimile element is used to represent a digital facsimile. It appears within a TEI document along with, or instead of, the text element introduced in section 5 Default Text Structure. When this module is selected therefore, a legal TEI document may thus comprise any of the following:

- a TEI Header and a text element
- a TEI Header and a facsimile element
- a TEI Header, a facsimile element, and a text element
<facsimile> Example in Context

<TEI>
  <teiHeader>
  <!-- teiHeader here -->
  </teiHeader>
  <facsimile>
    <graphic url="147_13.jpg" xml:id="p1"/>
    <graphic url="147_14.jpg" xml:id="p2"/>
    <graphic url="147_15.jpg" xml:id="p3"/>
  </facsimile>
  <text>
    <body>
    <!-- body here -->
    </body>
  </text>
</TEI>
The `<surface>` element may be used to indicate that there are two image files corresponding with the same area of the work:

```
<facsimile>
  <surface>
    <graphic url="147_13.jpg"/>
    <graphic url="147_13-huge.jpg"/>
  </surface>
  <graphic url="147_14.jpg"/>
  <graphic url="147_15.jpg"/>
  <graphic url="147_16.jpg"/>
</facsimile>
```
The actual dimensions of the object represented are not documented by the surface element; instead, the surface is located within an abstract coordinate space, which is defined by the following attributes, supplied by the att.coordinated class:

- \(\texttt{@ulx}\) gives the x coordinate value for the upper left corner of a rectangular space.
- \(\texttt{@uly}\) gives the y coordinate value for the upper left corner of a rectangular space.
- \(\texttt{@lrx}\) gives the x coordinate value for the lower right corner of a rectangular space.
- \(\texttt{@lry}\) gives the y coordinate value for the lower right corner of a rectangular space.
Example drawing rectangles
<surf<ace> Example

<facsimile>
  <surface ulx="0" uly="0" lrx="700" lry="1000">
<!-- ... -->
</surface>
</facsimile>
<zone> in <surface>

To describe the whole image, we will also need to define a zone of interest which represents an area inside this surface. This zone of interest can be defined by a <zone> element, within which we can place the <graphic>:

```xml
<facsimile>
  <surface ulx="0" uly="0" lrx="993" lry="1639">
    <zone ulx="93" uly="681" lrx="967" lry="1568">
      <graphic url="147_13.jpg"/>
    </zone>
  </surface>
</facsimile>
```
The `<desc>` element may also be used within either `<surface>` or `<zone>` to provide some further information about the area being defined.
<desc> Example

<facsimile>
  <surface ulx="0" uly="0" lrx="993" lry="1639">
    <desc>Printed page</desc>
    <zone ulx="96" uly="89" lrx="750" lry="657">
      <desc>Cartoon</desc>
      <graphic url="147_13.jpg"/>
    </zone>
    <zone ulx="95" uly="681" lrx="990" lry="1568">
      <desc>Text section</desc>
      <graphic url="147_13.jpg"/>
    </zone>
  </surface>
</facsimile>
More uses for `<zone>`

In addition to acting as a container for `<graphic>` elements, `<zone>` elements may also be used to select parts of each surface for analytical purposes.

```
<facsimile>
  <surface ulx="0" uly="0" lrx="993" lry="1639">
    <desc>main text</desc>
    <zone ulx="393" uly="681" lrx="967" lry="890">
      <desc>Cartoon</desc>
      <graphic url="147_13.jpg"/>
    </zone>
    <zone ulx="507" uly="596" lrx="704" lry="768">
      <desc>Artist's signature</desc>
    </zone>
  </surface>
</facsimile>
```
Linking transcription and facsimile

1. give each relevant part of the facsimile an identifier
2. using the @facs attribute, point from the transcription into the <facsimile>
Linking transcription and facsimile: `<facsimile>`

```
<facsimile>
  <surface ulx="0" uly="0" lrx="993" lry="1639">
    <graphic url="147_10.jpg"/>
    <zone xml:id="cartoonfacs" ulx="96" uly="89" lrx="950" lry="657">
      <desc>Cartoon</desc>
    </zone>
    <zone xml:id="textfacs" ulx="93" uly="681" lrx="967" lry="1568">
      <desc>Text section</desc>
    </zone>
  </surface>
</facsimile>
```
There were eight pretty walkers...
Pointing from <facsimile> to transcription with @start attribute

It is also possible to point in the other direction, from a <surface> or <zone> to the corresponding text. This is the function of the @start attribute, which supplies the identifier of the element containing the transcribed text found within the <surface> or <zone> concerned.
@start attribute example: <facsimile>

<facsimile>
  <surface start="#p10" ulx="0" uly="0" lrx="993" lry="1639">
    <zone xml:id="cartoonfacs2" ulx="96" uly="89" lrx="950" lry="657">
      <desc>Cartoon</desc>
      <graphic url="147_10.jpg"/>
    </zone>
  </surface>
</facsimile>
@start attribute example: text

<body>
  <pb xml:id="p10" n="10"/>
  <div>
    <figure>
    <!-- cartoon info here -->
    </figure>
  </div>
  <div>
    <head>THE WALKERS</head>
    <lg>
      <l>There were eight pretty walkers...</l>
    </lg>
  </div>
</body>
Gravestone Example: <facsimile>

<facsimile>
  <surface xml:id="grave" ulx="0" uly="0" lrx="355" lry="678">
    <graphic url="gravestone-cropped.jpg"/>
    <zone ulx="83" uly="223" lrx="272" lry="256" xml:id="line1"/>
    <zone ulx="92" uly="251" lrx="256" lry="282" xml:id="line2"/>
    <zone ulx="21" uly="281" lrx="330" lry="308" xml:id="line3"/>
    <zone ulx="36" uly="306" lrx="320" lry="332" xml:id="line4"/>
    <zone ulx="85" uly="535" lrx="249" lry="556" xml:id="line5"/>
    <zone ulx="97" uly="556" lrx="241" lry="576" xml:id="line6"/>
    <zone ulx="58" uly="577" lrx="281" lry="595" xml:id="line7"/>
    <zone ulx="68" uly="595" lrx="271" lry="613" xml:id="line8"/>
  </surface>
</facsimile>
Gravestone Example: text

<div facs="#grave">
  <p>Private Moulds' gravestone</p>
  <div>
    <ab>
      <s facs="#line1">12851 PRIVATE</s>
      <lb/>
      <s facs="#line2">H. MOULDS</s>
      <lb/>
      <s facs="#line3">NORTHAMPTONSHIRE REGT.</s>
      <lb/>
      <s facs="#line4">23RD JULY 1916 AGED 21</s>
    </ab>
    <ab>
      <s facs="#line5">LOVING SON OF</s>
      <lb/>
      <s facs="#line6">MRS MOULDS</s>
      <lb/>
      <s facs="#line7">PETERBORO, ENGLAND</s>
      <lb/>
      <s facs="#line8">FOR EVER WITH US</s>
    </ab>
  </div>
</div>
Gravestone Example: rendered

facsimile-demo.html