The <msDesc> module: describing primary sources in TEI

Matthew Holford, Bodleian Library

6 July 2017

History and development of the module

- See M. J. Driscoll, 'P5-MS: A general purpose tagset for manuscript description', *Digital Medievalist* 1 (2006)
  - http://www.digitalmedievalist.org/journal/2.1/driscoll/#mjd-b003

- New to TEI P5 (2007)
- Earlier versions assumed that the encoded text could be described using standard bibliographic tags (e.g. <bibl>)

- Developed from MASTER (1999-2001) and TEI Medieval Manuscript Description Work Group (1998-2000)
- Developed principally by specialists in Western medieval manuscript description
  - Not archivists
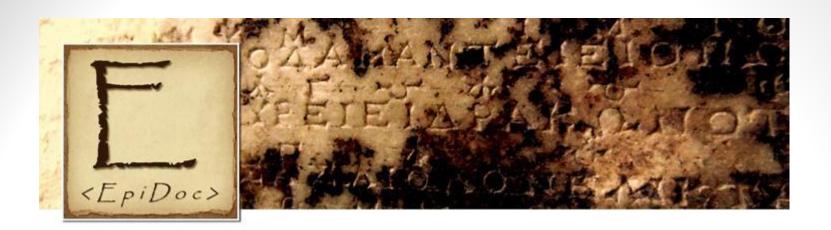 - Not 'Oriental' manuscripts
- Tags reflect this legacy

- Developed mainly for the purpose of manuscript description and cataloguing, rather than for text encoding
- Typically used for description and cataloguing, often linked to digital facsimiles

- Other standards are also important, e.g. MARC, EAD, but:
- TEI increasingly widely used:
  - Digital Walters
  - e-codices
  - Parker Library
  - Manuscriptorium
  - Fihrist
  - Genizah (Bodleian)

- Also used outside of 'manuscript' description:
  - Papyri (e.g. papryi.info) [sample xml]
  - EpiDoc
  - CEI (Charter Encoding Initative): Monasterium
  - early printed books
  - modern collections, e.g. Letters of Mary Russell Mitford [sample xml], Intellectual Berlin around 1800 [sample xml]

- Module was deliberately flexible, to deal with a wide range of data (legacy and modern descriptions)
  - Preserving descriptive prose
  - Allowing search and analysis
- Flexibility may conflict with interoperability

- More constrained versions of the module:
  - ENRICH (2008)
  - Europeana Regia (2010-12)
- Still several areas where further definition is possible / desirable (e.g. where should <secFol> go? What are the appropriate attributes of <material>?)

Browse by:

- **Manuscript Number**
  manuscripts by catalogue number

- **Text Title**
  manuscripts by title of texts contained within

- **Dimensions**
  manuscripts by physical dimensions

- **Date**
  manuscripts by date of origin

- **Language**
  manuscripts by language

- **Manuscript Title**
  manuscripts by manuscript title

- **Rubric**
  alphabetized list of rubrics

- **Incipit**
  alphabetized list of words that begin texts

- **Explicit**
  alphabetized list of words that end texts

- **Second Folio**
  alphabetized list of beginning words in second folio

Refine your search

Library / Collection	▸
Country of Location	▸
Text Language	▸
Century	▸
Dated	▸
Liturgica	▸
Material	▸
Illuminated	▸
Music Manuscript	▸
Restorated	▸
Descriptions	▸
With Annotations	▸
With Additional Bibliography	▸
With Additional Images	▸

e-codices, faceted search

\<msDesc>

Components of the module

```
<sourceDesc>
    <msDesc>
        <msIdentifier></msIdentifier>
        <msContents></msContents>
        <physDesc></physDesc>
        <history></history>
        <additional></additional>
    </msDesc>
</sourceDesc>
```

<msIdentifier>

```xml
<msIdentifier xml:id="n543.019027">
    <settlement xml:id="n543.019028">Baltimore</settlement>
    <repository xml:id="n543.019029">The Walters Art Museum</repository>
    <idno xml:id="n543.019030">W.20</idno>
</msIdentifier>


<msIdentifier>
    <settlement>Cologny</settlement>
    <repository>Fondation Martin Bodmer</repository>
    <idno>Cod. Bodmer 4</idno>
</msIdentifier>
```

```xml
<msIdentifier>
    <country>United Kingdom</country>
    <settlement>Oxford</settlement>
    <institution>University of Oxford</institution>
    <repository>Bodleian Library</repository>
    <idno type="shelfmark">Bodleian Library, Arch. G c.7</idno>
    <altIdentifier type="previous">
        <idno type="shelfmark">S 2.17 Art. [first Bodleian shelfmark,
            1624-1664?]</idno>
    </altIdentifier>
    <altIdentifier type="previous">
        <idno type="shelfmark">Arch. F c.13 [superscript z?] [second Bodleian
            shelfmark, 1906-?]</idno>
    </altIdentifier>
</msIdentifier>
```

\<head\>

```xml
<head>
    <title><persName role="author" key="pnd_100943764">Galterivs Anglicus</persName>, Fabulae</title>
    <origDate notBefore="1300" notAfter="1399">XIVe s.</origDate>
    <secFol><locus from="2r">f. 2</locus>: // Ecce natant<gap reason="sampling"/></secFol>
</head>
```

```xml
-<head>
    Letter from Mary Russell Mitford to B. R. Haydon,
    <date>31 October 1821</date>
    .
  </head>
```

<msContents>

```xml
<msContents>
    <p><!-- this is OK --></p>
</msContents>


<msContents>
    <summary><!-- optional --></summary>
    <textLang><!-- can go here as a top-level item --></textLang>
    <msItem></msItem>
    <msItem></msItem>
    <msItem></msItem><!-- etc -->
</msContents>
```

```xml
<msContents>
    <textLang mainLang="lat">The primary language in this
        manuscript is Latin.</textLang>


<msContents>
    <textLang mainLang="lat" otherLangs="fra"/>
```

<textLang>

- "In all cases languages should be identified by means of a standardized 'language tag' generated according to BCP 47"

  http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-textLang.html

- ISO 639-1 defined 2-letter codes, e.g. "la" = Latin

- ISO 639-2 defined 3-letter codes, e.g. "lat" = Latin

- 2-letter codes are preferred, but 3-letter codes are often found

<msItem>

```xml
<msItem>
    <title><!-- ok --></title>
</msItem>
<msItem>
    <note><!-- ok --></note>
</msItem>
<msItem>
    <p><!-- ok --></p>
</msItem>
```

```
<msItem>
        <locus></locus>
        <author></author>
        <title></title>
        <note></note>
        <rubric></rubric>
        <incipit></incipit>
        <explicit></explicit>
        <finalRubric></finalRubric>
        <colophon></colophon>
        <note></note>
        <bibl></bibl>
</msItem>
```

```xml
<msItem xml:lang="lat">
    <locus from="1r" to="23r">f. 1-23</locus>
    <author key="pnd_100943764">Galterivs Anglicus</author>
    <title>Fabulae</title>
    <note>L'ordre des fables non numérotées, dont quelques-unes manquent,
        est différent de celui de l'éd. de J. Bastin; chacune d'elle est
        suivie de la Moralitas.</note>
    <rubric><sic>Incpit</sic> liber Esopi.</rubric>
    <msItem>
        <note>(prol.)</note>
        <incipit rend="dots-after">Vt iuuet et prosit conatur pagina
            presens</incipit>
        <note>(12 v.).</note>
    </msItem>
```

Other types of text / document

```
<msItem>
    <locus from="1a" to="1b">Fol. 1</locus>
    <author/>
    <title>Deed of sale of a slave</title>
    <note>Deed of sale of a slave by <quote>לשמואל ור̇ור̇מר
יפת הלוי[</quote> to <quote>[ילהלוי [ב̇מרנ̇ ורבנ̇ יפת הלוי ... נ̇ע̇ הידוע בן אלבתנוני
ע̇נ̇ משה הלוי ור̇מר̇בר</quote>, dated the <date calendar="Hebrew-Seleucid"
when="1519">5th of Kislew, 1519 Sel.</date> = <date calendar="Gregorian"
when="1218">1218 A.D.</date>, <quote>בעיר אלקאהרה</quote>.
                Witnesses, <quote>א̇ב̇א צדוק ביר̇ הלל ביר̇ אברהם</quote>
ל̇צ̇ז</quote> and <quote>משה ביר̇ אלעזר</quote>.
                Verso blank.
    </note>
    <textLang mainLang="heb"/>
</msItem>
```

Printed books: a special case

```
<msContents>
    <titlePage>
        <docTitle>
            <titlePart>M<hi rend="superscript">r</hi> VVILLIAM <lb/>
                <hi rend="large">SHAKESPEARES</hi>
                <lb/>COMEDIES, <lb/>HISTORIES, &amp; <lb/>TRAGEDIES. </titlePart>
            <titlePart>Publiſhed according to the True Originall Copies.</titlePart>

        </docTitle>

        <docImprint>London : Printed by Isaac Iaggard, and Ed. Blount [at the charges
            of W. Iaggard, Ed. Blount, I. Smithweeke, and W. Aspley].,
            <docDate>1623</docDate>.</docImprint>
    </titlePage>
</msContents>
```

In this instance <author>, <title> etc. are given earlier in the TEI header

```xml
<titleStmt>
    <title type="statement">The Tempest from Mr. William Shakespeares comedies, histories, &amp; tragedies.
        Published according to the true originall copies.</title>
    <title type="variant">Mr. VVilliam Shakespeares comedies, histories, &amp;
        tragedies</title>
    <title type="distinctive">Bodleian First Folio, Arch. G c.7</title>
    <author key="LCCNn78095332">Shakespeare, William, 1564-1616.</author>
    <editor>Heminge, John, approximately 1556-1630</editor>
    <editor>Condell, Henry, -1627</editor>
    <respStmt>
        <persName>Droeshout, Martin, 1601-</persName>
        <resp>engraver</resp>
    </respStmt>
    <respStmt>
        <persName>Jaggard, Isaac, -1627</persName>
        <resp>printer</resp>
    </respStmt>
    <respStmt>
        <persName>Blount, Edward, fl. 1594-1632</persName>
        <resp>printer</resp>
    </respStmt>
    <respStmt>
        <persName>Jaggard, William, 1569-1623</persName>
        <resp>publisher</resp>
    </respStmt>
```

<physDesc>: physical description

```
<physDesc>
    <objectDesc>
        <supportDesc></supportDesc>
        <layoutDesc></layoutDesc>
    </objectDesc>
    <handDesc></handDesc>
    <typeDesc></typeDesc>
    <decoDesc></decoDesc>
    <bindingDesc></bindingDesc>
    <sealDesc></sealDesc>
    <accMat></accMat>
</physDesc>
```

```
<objectDesc form="codex">
    <!-- ... -->
</objectDesc>
```

- Suggested values (Enrich / Europeana Regia):
  - codex
  - fascicle [ER only]
  - leaf ['a loose leaf']
  - scroll
  - other

```xml
<objectDesc form="codex">
    <supportDesc>
        <support></support>
        <extent>
            <dimensions></dimensions>
        </extent>
        <foliation></foliation>
        <collation></collation>
        <condition></condition>
    </supportDesc>
```

```xml
<supportDesc xml:id="n543.019052" material="parchment">
    <support xml:id="n543.019053">
        <p xml:id="n543.019054">Medium to thick goatskin showing
            significant handling, including smudges and soiling at the
            lower edges of folios, ink-blot stains, and the last lines
            of text on many folios appear faded from rubbing; all folios
            are of a faded yellow to gray color; fore-edges and lower
            margins have been trimmed on many folios, sometimes
            extensively (see fols. 7, 9, 11, 37, and 54); trimming
            appears to have been started on fol. 55 but halted and later
            mended by the Walters Art Museum's conservation laboratory;
            prickings visible on fore-edges of folios; occasional holes
            in parchment or evidence of flaws in the goatskin; edges of
            folios show traces of red pigment</p>
    </support>
```

Suggested values: "perg" (parchment), "chart" (paper), "mixed" (paper and parchment), "papyrus", "other"

```xml
<extent>Foliation: iii+55+iii
    <dimensions type="leaves"><height
            unit="cm">23.0</height>
        <width unit="cm">14.5</width>
    </dimensions>
    <dimensions type="written" n="item 1040"
            ><height unit="cm">22.0</height>
        <width unit="cm">13.5</width>
    </dimensions>
</extent>


<extent>
    <measure type="leavesCount" n="II + 23 + I">II + 23 ff. + f. [24]</measure>
    <measure type="pageDimensions" n="21.7 x 15.4 cm">217 x 154 mm</measure>
</extent>
```

There is currently no agreed/satisfactory markup
for representing the extent (e.g. iii + 55 + iii) of a manuscript

```xml
<collation>Cahiers: <locus from="V2r">f. II</locus>,
    <locus from="1r" to="4v">1-4</locus>:
    <term>quinion</term> avec <catchwords>réclame</catchwords>,
    <locus from="10r" to="17v">f. 10-17</locus>: <term>quaternion</term> et
    <locus from="18r" to="23v">f. 18-23</locus>: <term>ternion</term>,
    sans <catchwords>réclames</catchwords>. Réglure à la mine de plomb.</collation>
```

More common to enclose the collation formula in <formula>;
but no standard tags for more detailed markup of formulae

```xml
<handDesc>
    <handNote>Transitional Carolingian
        miniscule; rustic capitals and uncials used at the beginning of
        lines of poetry</handNote>
</handDesc>
<decoDesc>
    <decoNote>Incipit "S" on fol. 1r
        rendered in blue ink and red outline with floral motif; first
        letters of major text divisions marked in red ink with decorative,
        calligraphic forms; text in brown ink</decoNote>
    <!-- ... -->
</decoDesc>
```

```xml
<handDesc n="1">
    <handNote script="gotica"><locus from="1r" to="23r">f. 1-23</locus>,
        écriture gothique arrondie, une main, </handNote>
</handDesc>
<decoDesc>
    <decoNote type="initial">Titres rubriqués;
        au <locus from="1r">f. 1</locus> grande initiale bleue ajourée de blanc ornée
        de dessins filigranés rouges et bleus; autres initiales bleues, à filigranes
        rouges seulement jusqu'au <locus from="11r">f. 11</locus>, et initiales rouges,
        à filigranes bleus seulement au <locus from="1r">f. 1</locus>;
        initiales des vers rehaussées de jaune.</decoNote>
</decoDesc>
```

<physDesc> in other contexts

```
<support>
 <p>The greater part of a moulded <material>marble</material>
  <objectType>base</objectType>, recomposed from four pieces (together,
 <dimensions unit="metre">
   <width>0.62</width>
   <height>0.33</height>
   <depth precision="low">0.34</depth>
  </dimensions>).</p>
</support>
```

http://www.stoa.org/epidoc/gl/latest/supp-descobject.html

<history>

```xml
<history>
    <origin>
        <origDate></origDate>
        <origPlace></origPlace>
    </origin>
    <provenance></provenance>
    <acquisition></acquisition>
</history>
```

```xml
<history>
    <origin><p>Italienne, au <locus from="2r">f.
2</locus>  les premières lignes du texte sont accompagnées en interligne
d'une traduction italienne: <quote>ecote si nodano<gap
reason="sampling"/>, quel tira, quela tira, quela strapoza</quote>, ligne
5: <quote>quela medesima paura</quote>, etc.</p></origin>
    <acquisition>Acquis par <persName
role="former_possessor" key="pnd_116216387">Martin Bodmer</persName> en
<date>1963</date> chez le libraire <persName role="seller"
key="pnd_120146894">B. Rosenthal, <placeName>New
York</placeName></persName>.</acquisition>
</history>
```

Refine your search

Library / Collection	▸
Country of Location	▸
Text Language	▸
Century	▸
Dated	▸
Liturgica	▸
Material	▸
Illuminated	▸
Music Manuscript	▸
Restorated	▸
Descriptions	▸
With Annotations	▸
With Additional Bibliography	▸
With Additional Images	▸

e-codices, faceted search

```xml
<history>
    <origin notAfter="1199" notBefore="1175">
        <origDate>Last quarter of the 12th century
            CE</origDate>
        <origPlace>Northern France</origPlace>
    </origin>
```

```xml
<provenance xml:id="n543.019084">Owned by Antoine Loisel (1539-1596), based
    on the signature on fol. 55v, written vertically at the center of the
    page</provenance>
<provenance xml:id="n543.019085">Owned by the Le Peletier de Rosanbo (or
    Rosambo) family in the eighteenth or early-nineteenth century, based on
    the first bookplate on the front pastedown</provenance>
<provenance xml:id="n543.019086">Owned by the Duchess of Berry in the
    early-nineteenth century, based on the second bookplate on the front
    pastedown, which reads "Bibliothèque de Rosny," referring to the Château
    de Rosny at Seine-et-Oise</provenance>
```

```xml
<additional>
    <adminInfo>
        <recordHist>
            <source></source>
        </recordHist>
        <availability></availability>
        <custodialHist></custodialHist>
    </adminInfo>
    <surrogates></surrogates>
    <listBibl></listBibl>
</additional>
```

```xml
<adminInfo>
    <recordHist>
        <source>
            <bibl><persName>Pellegrin Elisabeth</persName>,
                Manuscrits Latins de la Bodmeriana, Cologny-Genève <date>1982</date>, pp. 13-14.</bibl>
        </source>
    </recordHist>
</adminInfo>


                <additional>
                    <surrogates>
                        <listBibl>
                            <bibl type="digitalFacsimile">Digital facsimile images available at: <ref
target="http://firstfolio.bodleian.ox.ac.uk/">http://firstfolio.bodleian.ox.ac.uk/</ref>.</bibl>
                        </listBibl>
                    </surrogates>
                </additional>
```

<msPart>

```xml
<msPart>
    <altIdentifier><idno>Cod. Bodmer 4 (flyleaves)</idno></altIdentifier>
    <msContents>
        <msItem>
            <locus>(fols. I-Iv, 24-24v)</locus>
            <title>Genesis</title>
            <note>(French translation with commentary)</note>
            <textLang mainLang="fr">French</textLang>
        </msItem>
    </msContents>
    <physDesc>
        <objectDesc form="leaf">
            <supportDesc material="perg">
                <support>Parchment</support>
            </supportDesc>
        </objectDesc>
    </physDesc>
    <history>
        <origin>
            <origDate notAfter="1299" notBefore="1275">13th century, end</origDate>
        </origin>
    </history>
</msPart>
<msPart><altIdentifier><idno>Cod. Bodmer 4 (main text)</idno></altIdentifier>
    <!-- etc -->
</msPart>
```