The Menota handbook

Guidelines for the electronic encoding of Medieval Nordic primary sources

Appendices

Version 2.0
TEI P5 conformant

Odd Einar Haugen
(editor)

Tone Merete Bruvik, Matthew Driscoll, Odd Einar Haugen,
Karl G. Johansson, Rune Kyrkjebø and Tarrin Wills
(contributors)

The Medieval Nordic Text Archive
www.menota.org

Bergen 16 May 2008

A. Characters

Synchronisation with MUFI

As of v. 2.0 of the Menota handbook, the selection and encoding of characters have been synchronised with v. 2.0 of the MUFI character recommendation. This recommendation lists more than 1,300 characters in the Latin alphabet of potential use for the encoding of Medieval Nordic primary sources:

The MUFI character recommendation

The MUFI character recommendation lists Unicode code points and entity names for all characters in the file. Exactly the same entity names have been used in the Menota list of entities, cf. Appendix D.1.1.

The MUFI character recommendation and the Menota entity list are compatible with the Unicode Standard v. 5.0, published in July 2006. V. 5.1 of the Unicode Standard was published in April 2008. As soon as the MUFI character recommendations is updated with this revsion, a similar update will be made to the Menota entity list, probably before the end of 2008. In the meantime, the present entity list and MUFI compatible fonts can safely be used.

B. Fonts

MUFI compatible fonts

As of v. 2.0 of the Menota handbook, MUFI compatible fonts are recommended for the display of any texts encoded according to the Menota guidelines. A list of MUFI compatible fonts can be found on the MUFI web site:

MUFI compatible fonts

Note that several of the MUFI fonts can be downloaded free of charge. At present, all fonts are in the format Windows TrueType. Fonts of this type can be used on all major computer platforms, such as Linux, Mac (OS X), and Windows.

C. XML editors

C.1 Introduction

In order to write and edit XML one needs some sort of editing tool. Due to the nature of XML almost any text editing tool can be used, but there are also some specialised editors on the market. The decision on what kind of editor to use depends more or less entirely on personal preference and what kind of software the individual users are familiar with. There is no reason why a person cannot use Microsoft Word to edit or encode XML if he or she wishes – if that is the preferred editor for document editing.

The suggestions made below are only meant to be a guide to what we feel are convenient tools. For each of the editors we describe, there are probably a large number of almost identical alternatives, but these are the XML editing tools we have made use of and feel comfortable recommending to others.

C.2 TextPad

http://www.textpad.com

Operating system: Windows (all versions)
Version: 5.0.
Price: USD 33.00 (single user)

Full size image (1025 x 768 pixels)

C.2.1 Description

TextPad is basically a general text editor similar to the built-in Notepad in Windows. The main difference between these two editors is that TextPad has a number of additional functions that make text editing somewhat easier than in Notepad. Among the many extra functions are:

Unlimited undo/redo function. As opposed to many other editors that only remember the last command given, there is no limit to how far back you can go with the undo function i TextPad.

Expanded search and replace. TextPad does not only have the ability to do search and replace in text, but it can also search for regular expressions like tabulator, linebreak and space.

A built-in 'file manager' that makes it easier to edit several documents simultaneously.

In addition to the numerous built-in functions there are a large number of programs that are made by TextPad users to extend the functionality of the program. Among the many add-ons we find a program that 'cleans up' superfluous tabulators and blank lines and an option to run the Expat XML parser within TextPad. When it comes to XML support, TextPad has the ability to run DOS-commands from a menu. This basically means that validation of XML files with an external program like Xerces can be done without having to go to the command prompt. Command results are brought up in a file of its own in the file manager window.

Another useful function in TextPad is the ability to use so-called 'clip libraries'. In the context of HTML/XML this means that the user can be given a list of all tags in a particular code language in a menu. Instead of typing the tags, they can just be selected from the menu and inserted into the text. Examples of existing 'clip libraries' are: Dublin Core, HTML 4.0 and LaTeX. For Menota the biggest advantage of such clip libraries would probably be to create menus for the extensive entity list.

C.3

http://www.oxygenxml.com/

Operating system: Windows / Mac OS X / Linux / Unix / Eclipse
Version: 9.1 (as of 18 December 2007)
Price: from USD 48 (single user – academic version), 30 days free trial version
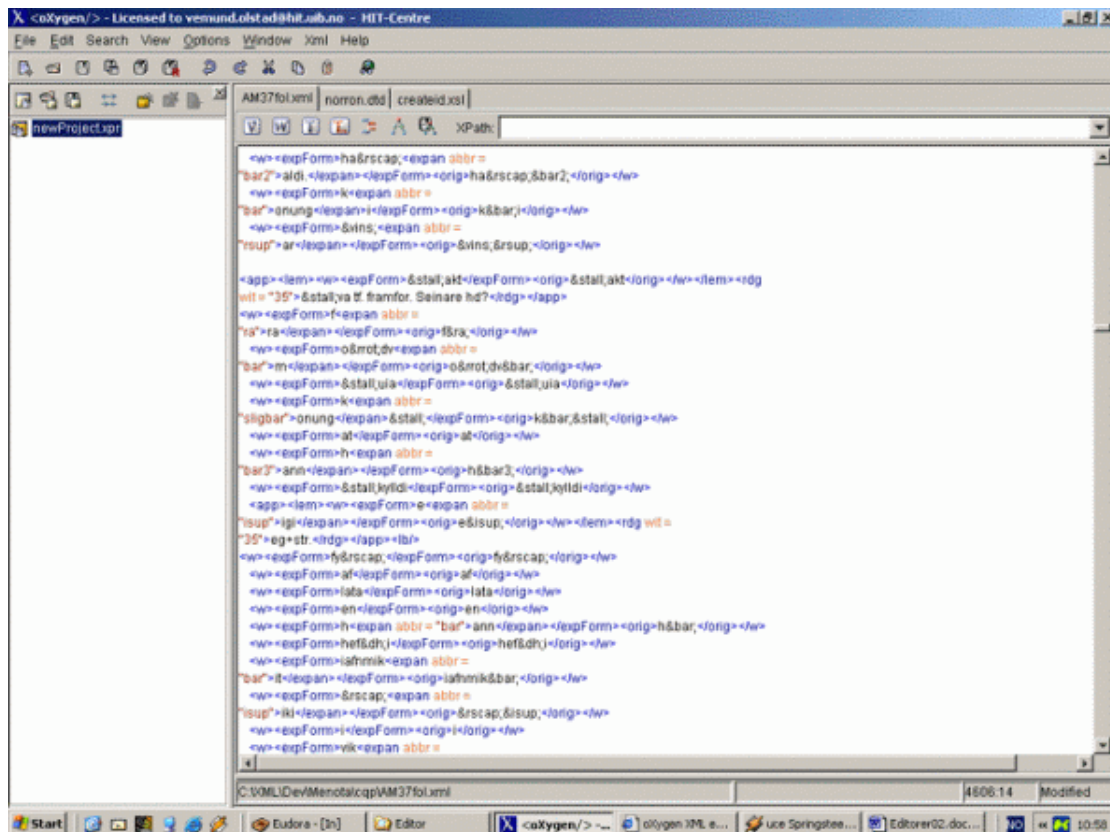
Full size image (1025 x 768 pixels)

C.3.1 Description

is a pure XML editor and does therefore have quite a few specialised functions that are not normally found in a general text editor of the kind TextPad represents. is, in many ways, similar to editors like XML Spy og Xmetal, but there are some points that, in our view, makes it a first choice compared to other editors.

This program is, like many other XML-related programs, written in Java and therefore available on most platforms, provided the computer has a Java environment installed. Like most specialised XML editors, has built-in functions to check for well-formedness and validity of the file you are working with. It is also possible to run XSLT and FOP transformations from within the editor, and it has full Unicode support.

One of the big advantages of , especially if working with premade schemas or dtds, is the so-called 'Code-insight' system which it uses. If the document being edited is connected to a DTD (or XML Schema), the editor will be able to read this and 'understand' the structure of the document. During text editing, context sensitive menus will be available – allowing for auto-completion of XML tags and attributes. This eliminates a lot of manual coding and makes editing easier for the author (see illustrations below).
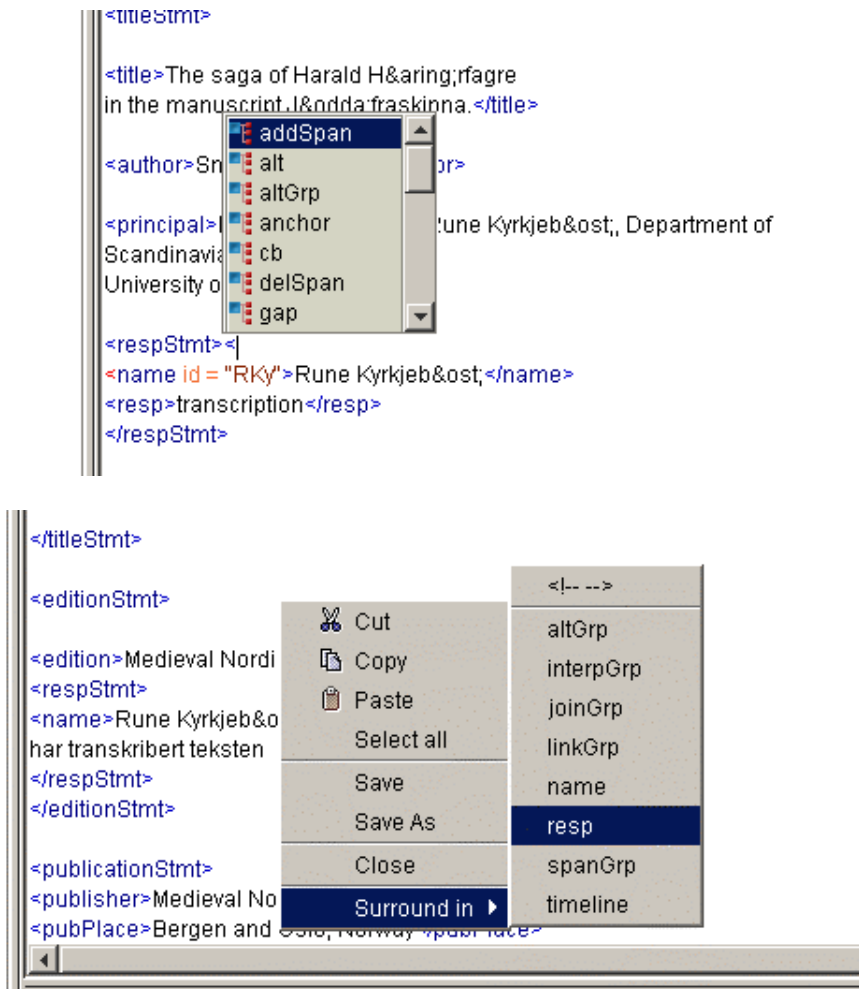
In addition to being context sensitive to existing DTD's, the program also has the ability to 'learn' the document structure of an XML file and build its own DTD if there is not one present already. <oXygen/> also has colour coding of the file formats it supports, and this can be changed to suit the users' preferences.

This editor is, in many ways, similar to emacs, which is a commonly used editor within the SGML/XML community – but it is somewhat easier to install and use.

C.4 Alternatives

The recommendation of <oXygen/> above is based on personal preference. There are many other XML editors out there that will work just as well - or in some cases even better. <oXygen/> was chosen for three reasons:
It is available for all platforms/operating systems
It has partial support for TEI built-in
It is relatively inexpensive.

If you need alternatives, some of the most frequently used editors are listed below:

C.4.1 XML Spy

http://www.altova.com/products/xmlspy/xml_editor.html

Operating system: Windows (2000, XP, VISTA, Server 2003)
Version: 2008

Price: from EUR 399
30 days free trial version

XML Spy is probably, as of today, the most widely used XML editor. It has won numerous industry awards, and is generally considered to be at the forefront of technological development - in terms of including new XML-related tecnologies into the software.

XML Spy is made specifically for Windows, which means it is less memory-intensive than Oxygen. The one negative point if you compare it to Oxygen is it's relatively high price.

C.4.2 XMetal

http://www.xmetal.com/

Operating system: Windows (NT 4.0, 2000, XP, Server 2003)
Version: 5.0 (as of 18 December 2006)
Price:from EUR 540
30 days free trial version

XMetal is a suite of programs, described as "XML-based content lifecycle solutions to help global organizations automate and streamline processes for creating, managing, and publishing high volumes of content.". The editor part of it is closer, both in terms of looks and function, to what we are used to seeing in WYSIWYG editors for HTML than XML Spy. But it is still one of the most widely used XML editors out on the market, and a good alternative to XML Spy. It is, however, very expensive, and probably something you should consider if you want to get involved in making/maintaining larger knowledge bases and sites using more advanced XML technologies.

D. Menota schemas (DTD, RELAX NG, and entity list)

D.1 Menota schemas

In version 1 of the Menota handbook, a number of changes were made to the TEI Guidelines version P4. These changes were implemented in two extension files, 'menota.extention.ent' and 'menota.extension.dtd', both written in the DTD syntax for SGML. The Document Type Definition (DTD) was thereafter generated using TEI Pizza Chef. For details on this procedure, see the page Document Type Definition in v. 1.1 of the handbook. Those who wish to encode texts according to TEI P4 should refer to this version of the handbook.

In version 2 of the Menota handbook, additions and changes have been made with reference to TEI Guidelines version P5. In TEI P5, the extensions are in XML in a schema called ODD (One Document Does it all). A new tool ROMA has been used to generate TEI schemas. ROMA allows users to choose between several outputs: a DTD, a RELAX NG schema (which, in spite of its name, can be a stricter schema than a DTD), or a W3C schema.

On the basis of the ODD file for Menota and the entity list, a Document Type Definition (DTD) and a RELAX NG schema for Menota have been generated with the help of ROMA. Those who prefer a W3C schema or a RELAX NG schema with compact syntax, may produce that themselves, using the Menota ODD file and the tool ROMA.

D.1.1 Current versions of Menota schemas

This is a list of all current Menota schemas. Earlier versions of these files have been moved down to the archive below and given a version number. Remember that many browsers will try to open XML files in a new window; in order to download these files, try using right-click (Windows) or alt-click (Mac).

(1) One Document Does it all (ODD)

The ODD file is an XML file which specifies all changes to the TEI Guidelines P5.
Menota 2.0 ODD file for TEI P5 (8 January 2008) – XML file

(2) Entity list

The entity list specifies recommended entities. This is not an obligatory part of a schema, but it is an important part of Menota encoding since a large number of characters have to be encoded by entities. It is essential that characters in the Private Use Area of the Unicode Standard are encoded with entities. The entity list is an ordinary text file.
Menota Entities (10 January 2007) – text file

(3) Document Type Definition (DTD)

The DTD is one of several schemas that are allowed in TEI P5.
Menota DTD (8 January 2008) – text file

Some browsers will attach the extension .xml to this file when downloaded. Just change the extension back to .dtd

(4) RELAX NG schema
The version given here has been generated on the basis of the same ODD file and entity file as the DTD above. Using a RELAX NG schema, data values can be validated to a higher extent than what is possible with a DTD. This applies both to attribute and content values. For example, the attributes 'notBefore' and 'notAfter' in the element **<origDate>** have to contain a valid date value if the RELAX NG schema is used. The RELAX NG schema also supports name spaces, which is useful if one wants to mix schemas as we do in Menota.
Menota RELAX NG (8 January 2008) – XML file

D.1.2 Archive of older versions of Menota schemas

(1) One Document Does it all (ODD)
Menota ODD #1 (20 December 2004) – XML file
Menota ODD #2 (7 June 2005) – XML file
Menota ODD #3 (21 June 2005) – XML file
Menota ODD #4 (6 December 2005) – XML file
Menota ODD #5 (6 February 2006) – XML file
Menota ODD #6 (5 December 2006) – XML file
Menota ODD #7 (13 November 2007) – XML file
Menota ODD #8 (6 Februar 2008) – XML file

(2) Entity list
Menota Entities #1 (19 November 2004) – text file
Menota Entities #2 (14 December 2005) – text file
Menota Entities #3 (6 December 2006) – text file
Menota Entities #4 (18 December 2006) – text file

(3) Document Type Definition (DTD)
Menota DTD #1 (16 December 2004) – text file
Menota DTD #2 (7 June 2005) – text file
Menota DTD #3 (21 June 2005) – text file
Menota DTD #4 (6 December 2005) – text file
Menota DTD #5 (6 February 2006) – text file
Menota DTD #6 (5 December 2006) – text file
Menota DTD #7 (27 December 2007) – text file
Menota DTD #8 (25 January 2008) – text file

(4) RELAX NG schema
Menota RELAX NG #1 (17 December 2004) – XML file
Menota RELAX NG #3 (21 June 2005) – XML file

Menota RELAX NG #8 (12 February 2008) – XML file

D.2 Referring to the Menota schema

Please note that as of TEI P5, a Menotic XML file should refer to two external files, one schema and one list of entities. The beginning of the file using the Menota P5 DTD schema should look like this:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE TEI SYSTEM "http://www.menota.org/guidelines-2/schemes/menotaP5.dtd"
[
<!ENTITY % Menota_entities SYSTEM
'http://www.menota.org/guidelines-2/schemes/menota-entities.txt'  >
%Menota_entities;]
>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
....
</TEI>
```

The beginning of the file using the Menota P5 RELAX NG schema might look like this if your are using Oxygen:

```
<?xml version="1.0" encoding="UTF-8"?>
<?oxygen RNGSchema="http://www.menota.org/guidelines-2/schemes/menotaP5.rng"
                type="xml" ?>
<!DOCTYPE TEI
[
<!ENTITY % Menota_entities SYSTEM
'http://www.menota.org/guidelines-2/schemes/menota-entities.txt'  >
%Menota_entities;]>

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

If you are using an other editor, the reference to the schema will be different.

Please note the namespace references in the TEI root element when the RELAX NG schema is used. The default namespace is the TEI namespace, while the additional Menota namespace is listed next as the "xmlns:me" (where 'me' is an abbreviation for 'Menota'). When the DTD schema is used, the Menota namespace can not the declared. When using the RELAX NG schema, all elements and attributes belonging to the Menota namespace have to use the prefix 'me:' in order to be valid, e.g.:

```
<interpretation me:lemmatized="completely" me:morphAnalyzed="completely">
  <p>The complete text has been lemmatised and morphologically analysed
  according to the rules specified in ch. 8 of the Menota Handbook, v. 2.0.</p>
</interpretation >
                ...
<w>
  <me:facs>karin<am>&us;</am></me:facs>
  <me:dipl>karin<ex>us</ex></me:dipl>
  <me:norm>Karinus</me:norm>
</w>
```

See the chapter on Menota header for downloadable samples of valid Menotic XML headers.

D.3 Changes from TEI P4 to TEI P5 which concerns Menota

TEI P5 1.0 was published on 2 November 2007.

D.3.1 Changes in the XML Document Prologue

The Document Type Declaration had to be changed from the P4 to the P5 version of the Menota schema because an entity list is now a separate document, and not included in the DTD as it was previously. The reason for making the entities into a sperate file is primarily that only DTDs can include entities (RELAX NG as well most other schema languages can not), and secondarily because it is easier to maintain. Section D2 above shows how documents using the Menota encoding should refer to the entities.

D.3.2 Encoding changes from P4 to P5 which concerns Menota

The root element has been changed from **<TEI.2>** to **<TEI>**

The attribute **@resp** in the element **<sic>** has been removed in P5. It was probably an error to have it declared in P4. Example 22 in chapter 7 of the Menota Handbook used it. This example has been removed from version 2.0 of the handbook.

The content model of **<change>** in **<revisionDesc>** is changed, the element **<respStmt>** is removed, and the content is more narrative. An example may look like this:

```
<change>On <date value="2006-02-02">2 February 2006</date>,
<name>Tone Merete Bruvik</name> edited the file according to the
TEI P5 0.3.1 published on 30 January 2006. </change>
```

The attribute **@langKey** in **<textLang>** has been changed into **@mainLang**:

```
<textLang langKey="oic">Old Icelandic</textLang>
```

->

```
<textLang mainLang="oic">Old Icelandic</textLang></change>
```

Encoding of hands has been changed, **<handList>** has been replaced with **<handNotes>** and **<hand>** has been replaced with **<handNote>**

The element **<date>** has an attribute **@when**, which in P4 was called **@value**

D.3.3 Changes from MASTER to P5

Menota v. 1.1 contained the recommendations from the MASTER project. TEI P5 includes the recommendations from the TEI Manuscript Description Taskforce. Although the recommendations from TEI Manuscript Description Task are based on the MASTER recommendations, there are some differences which Menota has to handle.

The element **<msHeading>** is not in P5, and the content of this element has to be encoded otherwise.

The elementet **<msWriting>** has been removed, **<handDesc>**should be used in its place.

Encoding in MASTER:

```
<msWriting hands="1">
  <handDesc script="Hybrida" scope="major">
    <p>The main hand (Hand 1) writes <locus>ff. 1r-9r and
    16r-118v</locus> in a practised Gothic hybrid.</p>
  </handDesc>
</msWriting>
```

Encoding in P5:

```
<handDesc hands="1">
  <handNote script="Hybrida" scope="major">
    <p>The main hand (Hand 1) writes ff. 1r-9r and
    16r-118v in a practised Gothic hybrid.</p>
  </handNote>
```

```
    </handDesc>
```

In MASTER, the attribute **@role** in the **<name>** element was used to encode the role of the person whose name was stated in the element. This attribute is not available in P5, so the Menota project had to decide how this should be encoded. This example is given in '10.6 The history of the manuscript':

```
<provenance>
  <p>According to AM 435 a 4to, ff. 54v-56v, the manuscript had been owned by
  <name type="person" role="owner">Sr. Þórður Jónsson á
  <name type="place">Staðastað</name> (1672-1720)</name>, who had got it
  from <name type="person" role="owner">Jón Hákonarson að
  <name type="place">Vatnshorni</name> ...
```

We suggest that **@subtype** is used to replace **@role** in this case.

It is not longer valid to have both **<p>** and **<msDesc>** as child elements of **<sourceDesc>**. A **<p>** element should be moved into a **<summary>** element like this (sample from *AM 233 a fol*):

```
<sourceDesc>
  <msDesc>
   <msIdentifier>
      <country key="DK">Denmark</country>
      <settlement>Copenhagen</settlement>
      <repository>The Arnamagn&aelig;an Institute</repository>
      <idno>AM 233 a fol</idno>
   </msIdentifier>
   <msContents>
     <summary>The source is an anonymous translation from Latin
     into Old Norse made in the second half of the 12th century,
     possibly in Norway. The original is lost and the text is only
     preserved in younger Icelandic manuscripts, including the two
     fragments in AM 233 a fol, fol. 28rA (the first column on the
     recto page) and 28vB (the second column on the verso page).
     This fragment is dated to 1350-1360 in "Ordbog over det
     norrøne prosasprog" (Index volume).</summary>
...
```

<msDescription> has been changed into **<msDesc>**

The value of the attribute **@defective** in **<msItem>** should have the values 'true' or 'false', not 'yes' or 'no'. This applies to other Boolean attributes as well.

The attributes **@technique**, **@figurative** and **@illustrative** in the **<decoNote>** element have been removed in P5.

The attribute **@reg** in the **<country>** has been removed. Use for instance **@key**.

D.4 Examples from the Menota Handbook

The samples in the *Menota Handbook 1.1* are available as a sample set at handbook_ex_1-1.xml

The samples from the *Menota Handbook 2.0* are found here:

Samples using DTD schema
Samples using RELAX NG XML schema

NB! If you open these files directly in your browser, you will most likely get an error message. These files are meant to be downloaded, not to be opened in a browser.

E. Menota header

E.1 Introduction

We recommend that all texts in the Medieval Nordic Text Archive contain a minimal amount of information about the text, its background and its encoding. This should be entered in the **<teiHeader>**, which is an integral and essential part of any XML/TEI document.

Below are two examples of headers for Medieval Nordic texts. The first example is of a header for a primary source with only a single text; the other for a source with more than one text.

A header may contain much more information than exemplified here. What we have specified in these examples is the recommended minimum amount of information.

Cf. the TEI P5 Guidelines ch. 2 'The TEI Header' for a discussion of headers in general and also the notion of minimal and recommended headers (ch. 2.6).

Important! Note that both headers will only be valid when using a Menota DTD schema. If you would like to use a RELAX NG schema you must add the Menota namespace to the very first line, as explained in ch. 1.9 and Appendix D.2:

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

The list in ch. 1.9 is a complete list of elements and attributes belonging to the Menota namespace. They must all be prefixed by 'me:' when using a RELAX NG schema.

E.2 Minimal header for a single-text source

Although many medieval manuscripts contain more than one text, there are some manuscripts which only contain a single piece of text. That is also generally true of charters (diplomas). We have chosen Holm perg 6 fol (also referred to as Sth. perg. fol. nr 6) as an example of a single-text source. This manuscript carries the text of *Barlaams ok Josaphats saga*, although with some parts missing.

Please note that the distinction between single-text and multi-text sources is not straightforward and ultimately rests on the definition of key concepts such as "text" and "work". If in doubt, we recommend using the established categories in catalogues and indices, e.g. the index volume of Ordbog over det norrøne prosasprog (Copenhagen 1989).

We have chosen Holm perg 6 fol also in order to demonstrate how a cumulative process of text encoding can be documented. This text was initially transcribed and edited in an electronic version by Magnus Rindal in the late 1970ies (the printed edition in 1981 was generated from this file). Subsequently, Jon Erik Hagen and Odd Einar Haugen lemmatised the text, i.e. added information about lemma and grammatical form for each running word. Finally, Christian-Emil Ore converted the file from its original customised format into Menotic XML. Thus, four people have contributed to the making of the present electronic version of the text, and should be credited accordingly.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader xml:lang="eng">
    <fileDesc>
```

```
<titleStmt>
  <title>Holm perg 6 fol (Barlaams ok Josaphats saga):
    an electronic edition</title>
  <respStmt>
    <resp>Edition by </resp>
    <name>Magnus Rindal
      <orgName type="affiliation">University of Bergen</orgName>
    </name>
  </respStmt>
  <respStmt>
    <resp>Lemmatisation and morphological encoding by </resp>
    <name>Jon Erik Hagen
      <orgName type="affiliation">University of Bergen</orgName>
    </name>
    <name>Odd Einar Haugen
      <orgName type="affiliation">University of Bergen</orgName>
    </name>
  </respStmt>
  <respStmt>
    <resp>Conversion to Menotic XML by </resp>
    <name>Christian-Emil Ore
      <orgName type="affiliation">University of Oslo</orgName></name>
  </respStmt>
</titleStmt>
<editionStmt>
  <p>First draft, <date when="2004-02-01">1 February 2004</date>
  </p>
</editionStmt>
<publicationStmt>
  <distributor>Medieval Nordic Text Archive</distributor>
  <idno type="Menota">Ms. 1</idno>
  <date when="2004-03-01">1 March 2004</date>
  <availability status="restricted">
    <p>This text is available for purposes of academic research
      and teaching only. Re-distribution in any form without prior
      permission is prohibited. Short extracts may be cited with
      full acknowledgment of the source.</p>
  </availability>
</publicationStmt>
<sourceDesc>
  <msDesc>
    <msIdentifier>
      <country  key="SE">Sweden</country>
      <settlement>Stockholm</settlement>
      <repository>Kungliga Biblioteket</repository>
      <idno>Holm perg 6 fol</idno>
      <msName>Barlaams ok Josaphats saga</msName>
    </msIdentifier>
    <msContents>
      <msItem n="1">
        <locus from="1r" to="102v">ff. 1r-102v (pp. 1-204)</locus>
        <title type="main">Barlaams ok Josaphats saga</title>
        <title type="abbreviated">Barl</title>
        <textLang mainLang="onw">Old Norwegian</textLang>
      </msItem>
    </msContents>
    <physDesc>
      <objectDesc form="codex">
        <supportDesc>
          <support>
            <p><material>Parchment</material>, but with
              five younger <material>paper</material> leaves
              added; two at the front (the last being blank) and
              three at the back.</p>
          </support>
          <extent>102 parchment leaves and 5 paper leaves;
            210 mm (height) by 155 mm (width).</extent>
          <foliation>
```

```
                    <p>The manuscript is paginated on the recto pages
                       of the parchment leaves, 1-204.</p>
                  </foliation>
                </supportDesc>
              </objectDesc>
            </physDesc>
            <history>
              <origin>
                <p>Written in <origPlace>Eastern Norway</origPlace>
                  <origDate notBefore="1250" notAfter="1300">c. 1275</origDate>.
                </p>
              </origin>
            </history>
          </msDesc>
        </sourceDesc>
      </fileDesc>
      <encodingDesc>
        <projectDesc>
          <p>This manuscript text has been encoded according to the standard
          set out in <title>The Menota handbook</title> (version 2.0), at
            <ref target="http://www.aksis.uib.no/menota/guidelines">
              http://www.aksis.uib.no/menota/guidelines</ref>
            as of <date>2006-03-30</date>.</p>
        </projectDesc>
        <editorialDecl>
          <correction status="high">
            <p>This text was proofread by Magnus Rindal and colleagues
            before the publication of the printed version in 1981. It is
            unlikely that it contains any significant number of errors.
            However, it can not be ruled out that the subsequent
            conversion of the file may have introduced some
            systemic errors.</p>
          </correction>
        </editorialDecl>
      </encodingDesc>
      <profileDesc>
        <langUsage>
          <language ident="oic">Old Icelandic</language>
          <language ident="onw">Old Norwegian</language>
          <language ident="osw">Old Swedish</language>
          <language ident="oic-onw">Old Icelandic with Old Norwegian traits
          </language>
          <language ident="onw-oic">Old Norwegian with Old Icelandic traits
          </language>
        </langUsage>
        <handNotes>
          <handNote xml:id="h1"/>
          <handNote xml:id="h2"/>
        </handNotes>
      </profileDesc>
      <revisionDesc>
        <change>
          <date>2008-01-09</date>
          <name>Tone Merete Bruvik</name>,
          <orgName type="affiliation">Aksis</orgName>: Updated the
          header to make it valid according to the Menota P5 DTD.
        </change>
        <change>
          <date>2006-04-18</date>
          <name>Tone Merete Bruvik</name>,
          <orgName type="affiliation">Aksis</orgName>: Changed the attribute use
          in the textLang element from xml:id into mainLang.
        </change>
        <change>
          <date>2006-03-30</date>
          <name>Tone Merete Bruvik</name>,
          <orgName type="affiliation">Aksis</orgName>: Revised the transcription
          in accordance with v. 2.0 of the Menota handbook
```

```
          and the corresponding Document Type Definition.
        </change>
        <change>
          <date>2003</date>
          <name>Christian-Emil Ore</name>,
          <orgName type="affiliation">University of Oslo</orgName>: Converted the
          electronic files (the transcription and the lemmatisation) into XML
          (Extensible Markup Language) according to the guidelines in
          the Menota handbook.
        </change>
        <change>
          <date>1988</date>
          <name>Jon Erik Hagen</name>,
          <orgName type="affiliation">University of Bergen</orgName> and
          <name>Odd Einar Haugen</name>,
          <orgName type="affiliation">University of Bergen</orgName>: Supplied
          lemmata (dictionary entry words) and grammatical forms for all words
          in the file. Published separately on microfiche in Norske
          språkdata vol. 22, Bergen 1989.
        </change>
        <change>
          <date>1980</date>
          <name>Magnus Rindal</name>
          <orgName type="affiliation">University of Bergen</orgName>:
          Finalised the electronic text and prepared it for publication
          in book form (published by Norsk Historisk
          Kjeldeskrift-Institutt in 1981).
        </change>
      </revisionDesc>
    </teiHeader>
    <text xml:lang="onw">
      <body>
        <div>
          <p>The text of the first chapter goes here.</p>
        </div>
        <div>
          <p><handShift new="h2"/>The text of the second
            chapter goes here.</p>
        </div>
        <div>
          <p>The text of the third chapter goes here.</p>
        </div>
      </body>
    </text>
 </TEI>
```

The body follows immediately after the header. This is where the actual transcription is located. Note that the **<div>** element is used to organise the various sections of the manuscript, in this case the individual chapters.

```
<text xml:lang="onw">
   <body>
     <div>
       <p>The text of the first chapter goes here.</p>
     </div>
     <div>
       <p><handShift new="h2"/>The text of the second chapter goes here.</p>
     </div>
     <div>
       <p>The text of the third chapter goes here.</p>
     </div>
   </body>
 </text>
```

The source text is placed within the **<text>** element. The attribute **@xml:lang** specifies the language of the text (cp. **<profileDesc>** above). If the text contains sporadic words or

sentences in another language, e.g. Latin, the attribute **@xml:lang** can be used to specify these exception from the rule, whether on the level of the word, **<w>**, paragraph, **<p>** or chapter, **<div>**.

This is also an example of how change of scribal hands in the source can be recorded. Note that rather than putting each hand in an element of its own (which would cause problems of overlapping) the actual change is recorded using a milestone element, **<handShift>**. Note that the number of hands must be recorded in the **<profileDesc>** element of the header.

E.3 Minimal header for a multi-text source

AM 242 fol – often referred to as Codex Wormianus or Órmsbók – is a good example of a multi-text source. It contains the prose version of *Edda* (by Snorri Sturluson), the four grammatical treatises, and a few other poetic texts (e.g. *Maríukvæði* and *Rígsþula*). *Snorra Edda* is usually divided into four sections; the *Prologue*, *Gylfaginning*, *Skaldskaparmál* and *Háttatal*. However, in the present manuscript, *Háttatal* is not placed in conjunction with the other sections of *Snorra Edda*, but towards the end of the manuscript, and *Maríukvæði* has been divided into two parts with a few other texts intervening. The structure is shown in fig. 1.
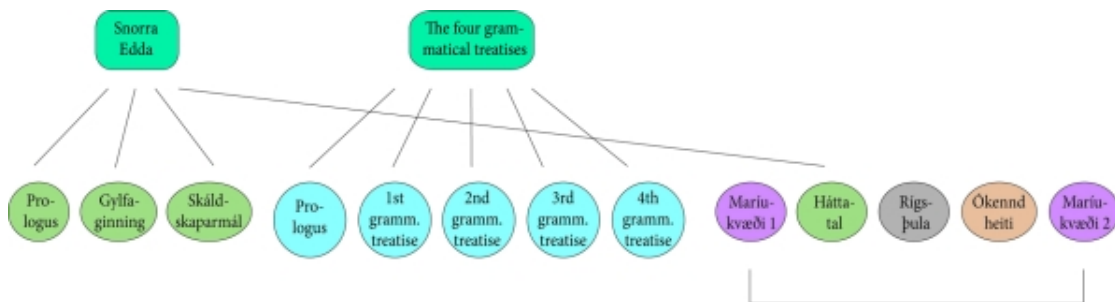
Fig. 1 The sequence of texts in Codex Wormianus (AM 242 fol). *Háttatal* is commonly regarded as the fourth part of *Snorra Edda*, but is located towards the end of the manuscript, and *Maríukvæði* has been divided into two parts.

Based on an analysis of the contents of the texts, *Háttatal* should be placed as the fourth part of *Snorra Edda*, but the two parts of *Mariukvæði* have been analysed as separate items and thus not joined. This interpretation is shown in fig. 2.

Fig. 2 The contents structure of texts in Codex Wormianus (AM 242 fol). *Háttatal* has been placed as the fourth part of *Snorra Edda*.

Although fig. 2 gives a better representation of the text structure of this particular manuscript, we recommend that the individual parts of the manuscript are encoded in sequential order, as shown in fig. 1. In general, we believe that the encoding of a source should keep closely to the source itself and not impose unnecessary interpretations on it. Even if there are good reasons for regarding e.g. *Háttatal* as the fourth part of *Snorra Edda*, the transcription should show that it is not located as the fourth part in the manuscript – and that this in fact may be intended by the scribe (or redactor).

The header of a multi-text source should specify the various parts of the source within the **<msContents>** element. Here, each part is described as an item, **<msItem>**, and linked with an **@id** number to the relevant section in the text itself. In the text, **<div>** elements identify the corresponding parts of the source.

If the source is encoded as recommended here, the text can be displayed in the sequence of the source itself, i.e. as in fig. 1, or by following the **@id** numbers, in the order of its contents (as analysed by the transcriber/editor), i.e. as in fig. 2.

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader xml:lang="eng">
    <fileDesc>
      <titleStmt>
        <title>AM 242 fol. (Codex Wormianus): an electronic edition</title>
        <respStmt>
          <resp>Transcription, lemmatisation and morphological encoding by
          </resp>
          <name><persName>Karl G. Johansson</persName>,
            <orgName type="affiliation">University of Oslo</orgName></name>
        </respStmt>
      </titleStmt>
      <editionStmt>
        <p>First draft, <date when="2002-11-01">1 November 2002</date>.</p>
      </editionStmt>
      <publicationStmt>
        <distributor>Medieval Nordic Text Archive</distributor>
        <idno type="Menota">Ms. 2</idno>
        <date when="2003-03-23">23 March 2003</date>
        <availability status="restricted">
          <p>This text is available for purposes of academic
            research and teaching only. Re-distribution in any form without
            prior permission is prohibited. Short extracts may be cited with
            full acknowledgment of the source.</p>
        </availability>
      </publicationStmt>
      <sourceDesc>
        <msDesc>
          <msIdentifier>
            <country key="DK">Denmark</country>
            <settlement>Copenhagen</settlement>
            <repository>The Arnamagn&aelig;an Institute</repository>
            <idno>AM 242 fol</idno>
            <msName type="nickname" xml:lang="lat">Codex Wormianus
            </msName>
          </msIdentifier>
          <msContents>
            <msItem n="1">
              <locus from="1v" to="35v">pp. 2-82 (including paper
                leaves)</locus>
              <title>Edda Snorra Sturlusonar</title>
              <title type="abbreviated">SnE</title>
              <textLang mainLang="oic">Old Icelandic</textLang>
              <msItem n="1.1">
                <locus from="1v" to="4v">pp. 2-8</locus>
                <title>Prologus</title>
              </msItem>
              <msItem n="1.2">
                <locus from="4v" to="20r">pp. 8-44</locus>
                <title>Gylfaginning</title>
              </msItem>
              <msItem n="1.3">
                <locus from="20v" to="35v">pp. 44-82</locus>
                <title>Sk&aacute;ldskaparm&aacute;l</title>
              </msItem>
            </msItem>
            <msItem n="2">
```

```
        <locus from="36r" to="54r">pp. 83-119</locus>
        <title>The four grammatical treatises</title>
        <title type="abbreviated">Gramm</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
        <msItem n="2.1">
          <locus from="36r" to="36r">p. 83</locus>
          <title>Prologus</title>
          <title type="abbreviated">GrammProl</title>
        </msItem>
        <msItem n="2.2">
          <locus from="36v" to="39v">pp. 84-90</locus>
          <title>First grammatical treatise</title>
          <title type="abbreviated">Gramm1</title>
        </msItem>
        <msItem n="2.3">
          <locus from="40r" to="41v">pp. 91-93</locus>
          <title>Second grammatical treatise</title>
          <title type="abbreviated">Gramm2</title>
        </msItem>
        <msItem n="2.4">
          <locus from="41v" to="50r">pp. 94-111</locus>
          <title>Third grammatical treatise</title>
          <title type="abbreviated">Gramm3</title>
        </msItem>
        <msItem n="2.5">
          <locus from="50r" to="54r">pp. 111-119</locus>
          <title>Fourth grammatical treatise</title>
          <title type="abbreviated">Gramm4</title>
        </msItem>
      </msItem>
      <msItem n="3">
        <locus from="54v" to="54v">p. 120</locus>
        <title>Mar&iacute;ukv&aelig;&eth;i</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
      </msItem>
      <msItem n="1.4" defective="true">
        <locus from="55r" to="60v">pp. 121-154 (including
          paper leaves)</locus>
        <title>H&aacute;ttatal</title>
        <title type="abbreviated">SnE</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
      </msItem>
      <msItem n="4" defective="true">
        <locus from="61r" to="61v">pp. 155-156</locus>
        <title>R&iacute;gs&thorn;ula</title>
        <title type="abbreviated">Edda</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
      </msItem>
      <msItem n="5" defective="true">
        <locus from="62r" to="63r">pp. 167-169</locus>
        <title>&Oacute;kennd heiti (Orms-eddu-brot)</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
      </msItem>
      <msItem n="6">
        <locus from="63v" to="63v">p. 170</locus>
        <title>Mar&iacute;ukv&aelig;&eth;i</title>
        <textLang mainLang="oic">Old Icelandic</textLang>
      </msItem>
    </msContents>
    <physDesc>
      <objectDesc form="codex">
        <supportDesc>
          <support>
            <p><material>Parchment</material>, but
              with some <material>paper</material> leaves
              added in the 17th century for the purpose of
              completing the text (pp. 61-72, 121-138,
              151-154, 157-166).</p>
```

```
                        </support>
                        <extent>85 leaves, including the younger paper
                          leaves; 280 mm (height) by 202 mm (width).</extent>
                        <foliation>
                          <p>The manuscript, including the younger
                            paper leaves, is paginated on the recto pages, 1-169.</p>
                        </foliation>
                      </supportDesc>
                    </objectDesc>
                  </physDesc>
                  <history>
                    <origin>
                      <p>Written in <origPlace>Iceland</origPlace> in
                        <origDate notBefore="1350" notAfter="1400">the second half of
                          the 14th century</origDate>; there are many
                        later additions, principally from the 17th century.</p>
                    </origin>
                  </history>
                </msDesc>
              </sourceDesc>
          </fileDesc>
          <encodingDesc>
            <projectDesc>
              <p>This manuscript text has been encoded according to the standard
                set out in <title>The Menota handbook</title> (version 2.0), at
                <ref target="http://www.aksis.uib.no/menota/guidelines">
                  http://www.aksis.uib.no/menota/guidelines</ref>
                as of <date>2006-03-30</date>.</p>
            </projectDesc>
            <editorialDecl>
              <correction status="high">
                <p>This text has been proofread by <name>Karl G. Johansson</name>
                  <name type="affiliation"> University of Oslo</name>.</p>
              </correction>
            </editorialDecl>
          </encodingDesc>
          <profileDesc>
            <langUsage>
              <language ident="oic">Old Icelandic</language>
              <language ident="onw">Old Norwegian</language>
              <language ident="osw">Old Swedish</language>
              <language ident="oic-onw">Old Icelandic with Old Norwegian traits</language>
              <language ident="onw-oic">Old Norwegian with Old Icelandic traits</language>
            </langUsage>
            <handNotes>
              <handNote xml:id="h1"/>
              <handNote xml:id="h2"/>
            </handNotes>
          </profileDesc>
          <revisionDesc>
            <change>
              <date>2008-01-09</date>
              <name>Tone Merete Bruvik</name>,
              <orgName type="affiliation">Aksis</orgName>:
              Updated the header to make it valid according to the Menota P5 DTD.
            </change>
            <change>
              <date>2006-04-18</date>
              <name>Tone Merete Bruvik</name>, <orgName type="affiliation">Aksis</orgName>:
              Changed the attribute use in the textLang element from xml:lang into mainLang.
            </change>
            <change>
              <date>2006-03-30</date>
              <name>Tone Merete Bruvik</name>, <orgName type="affiliation">Aksis</orgName>:
              Revised the transcription in accordance with v. 2.0
              of the Menota handbook and the corresponding Document Type Definition.
            </change>
            <change>
```

```
          <date>2004-05-10</date>
          <name>Karl G. Johansson</name>, <orgName type="affiliation">University
            of Oslo</orgName>
          Revised the transcription in accordance with v. 1.1 of the Menota handbook
          and the corresponding Document Type Definition.
        </change>
      </revisionDesc>
    </teiHeader>
```

The body follows immediately after the header. This is where the actual transcription is located. Note that **<div>** elements are used to organise the various parts of the manuscript, referring to the divisons set out in the **<msContents>** element of the header.

```
 <text xml:lang="oic">
   <body>
     <div n="1.1">
       <head>
         <supplied>Prologus</supplied>
       </head>
       <p>The text of part 1.1 goes here.</p>
     </div>
     <div n="1.2">
       <head>
         <supplied>Gylfaginning</supplied>
       </head>
       <p>The text of part 1.2 goes here.</p>
     </div>
     <div n="1.3">
       <head>
         <supplied>Sk&aacute;ldskaparm&aacute;l</supplied>
       </head>
       <p>The text of part 1.3 goes here.</p>
     </div>
     <div n="2.1">
       <head>
         <supplied>Prologus</supplied>
       </head>
       <p>The text of part 2.1 goes here.</p>
     </div>
     <div n="2.2">
       <head>
         <supplied>The first grammatical treatise</supplied>
       </head>
       <p>The text of part 2.2 goes here.</p>
     </div>
     <div n="2.3">
       <head>
         <supplied>The second grammatical treatise</supplied>
       </head>
       <p>The text of part 2.3 goes here.</p>
     </div>
     <div n="2.4">
       <head>
         <supplied>The third grammatical treatise</supplied>
       </head>
       <p>The text of part 2.4 goes here.</p>
     </div>
     <div n="2.5">
       <head>
         <supplied>The fourth grammatical treatise</supplied>
       </head>
       <p>The text of part 2.5 goes here.</p>
     </div>
     <div n="3">
       <head>
         <supplied>Mar&iacute;ukv&eth;i (1)</supplied>
       </head>
```

```
        <p>The text of part 3 goes here.</p>
      </div>
      <div n="1.4">
        <head>
          <supplied>H&aacute;ttatal</supplied>
        </head>
        <p>The text of part 1.4 goes here.</p>
      </div>
      <div n="4">
        <head>
          <supplied>R&iacute;gs&thorn;ula</supplied>
        </head>
        <p>The text of part 4 goes here.</p>
      </div>
      <div n="5">
        <head>
          <supplied>&Oacute;kennd heiti</supplied>
        </head>
        <p>The text of part 5 goes here.</p>
      </div>
      <div n="6">
        <head>
          <supplied>Mar&iacute;ukvæ&eth;i (2)</supplied>
        </head>
        <p>The text of part 6 goes here.</p>
      </div>
    </body>
  </text>
```

The source text is placed within the **<text>** element. The attribute **@xml:lang** specifies the language of the text (cp. **<msContents>** above). If the text contains sporadic words or sentences in another language, e.g. Latin, the attribute **@xml:lang** can be used to specify these exception from the rule, whether on the level of the word, **<w>**, paragraph, **<p>** or chapter, **<div>**.

The divisions of the text correspond to the list of **<msItem>** elements in **<msContents>** above, using the same numbers (1.1, 1.2, 1.3, 2.1, 2.2, etc.)

The **<head>** element specifies the title of the work (or a part of it). If the title is not stated in the source, the attribute type with the value supplied should be used. If the title is stated as a rubric in the manuscript, the type attribute should still be used, but with the value rubric.

E.4 Downloading sample XML headers

We offer the two headers in this chapter for free downloading:

Download the header of Holm perg 6 fol as a sample XML file

Download the header of AM 242 fol as a sample XML file

Please note that some browsers may try and interpret and open this sample file. In order to download the file to your disk, use alt-click (Mac) or right-click (Windows) on your browser, unless your browser has other preferences.

F. Extensible Stylesheet Language Transformation (XSLT)

F.1 Introduction

XSLT is a language for transforming XML documents into other XML documents, or other formats, e.g. HTML.

This page contains XSL Transformations (which we shall refer to as XSLT stylesheets) to be used with Menotic XML files in order to generate user-friendly HTML. To help users modify these stylesheets, they contain a number of comments explaining the intended effects of the stylesheet.

The XSLT stylesheets on this page have been developed by Vemund Olstad and Tone Merete Bruvik, Aksis, University of Bergen.

F.2 XSLT for Menotic texts

There are two XSLT stylesheets available:

menota-web-comment-P5.xsl, 9 January 2008, for texts encoded according to *Menota 2.0* (TEI P5 conformant)
menota-web-comment.xsl, 25 May 2005, for texts encoded according to *Menota 1.1* (TEI P4 conformant)

To download these stylesheets rather than opening them, right-click (on Windows) or ctrl-click (on Mac), depending on the preferences in your browser. Otherwise, you will most likely get an error report when your browser tries to open the XSL file.

These XSLT stylesheets will convert a text encoded with **<facs>**, **<dipl>** and/or **<norm>** tags into easily readable HTML display. Note that your computer must have a MUFI conformant font in order to display all characters correctly, especially on the **<facs>** level.
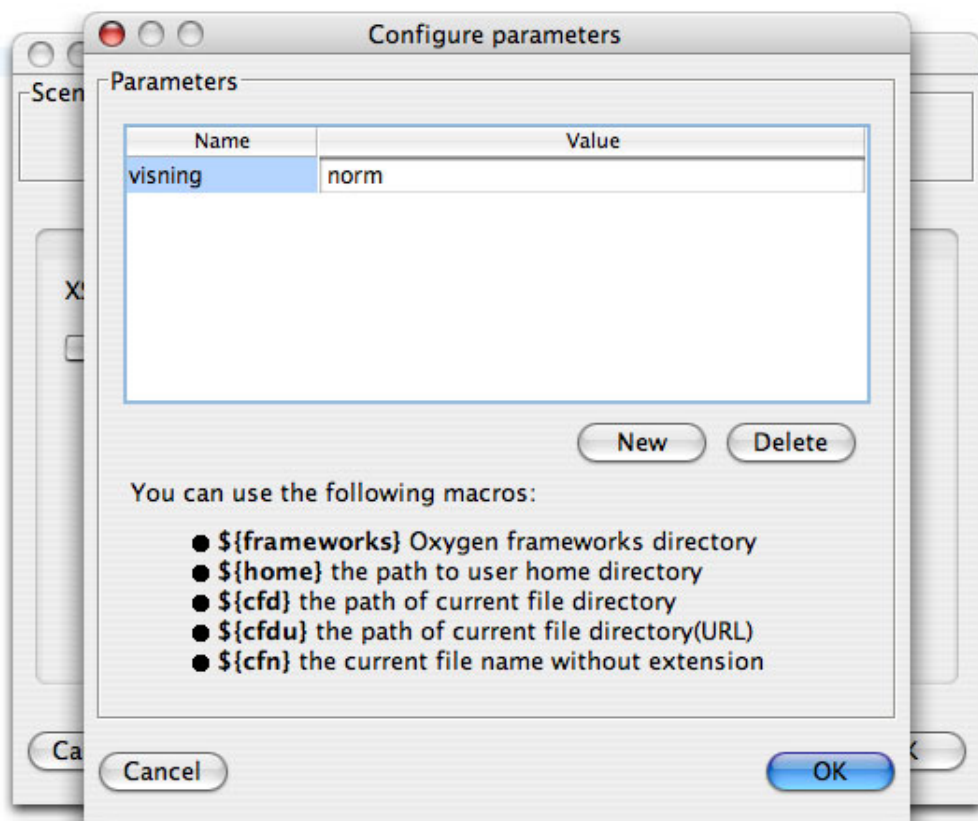
There are two ways of choosing the display level:

F.2.1 Using Oxygen (or other XML applications)

You may use the parameter option in e.g. Oxygen, or in a command line. In this case, specify 'facs', 'dipl' or 'norm'.
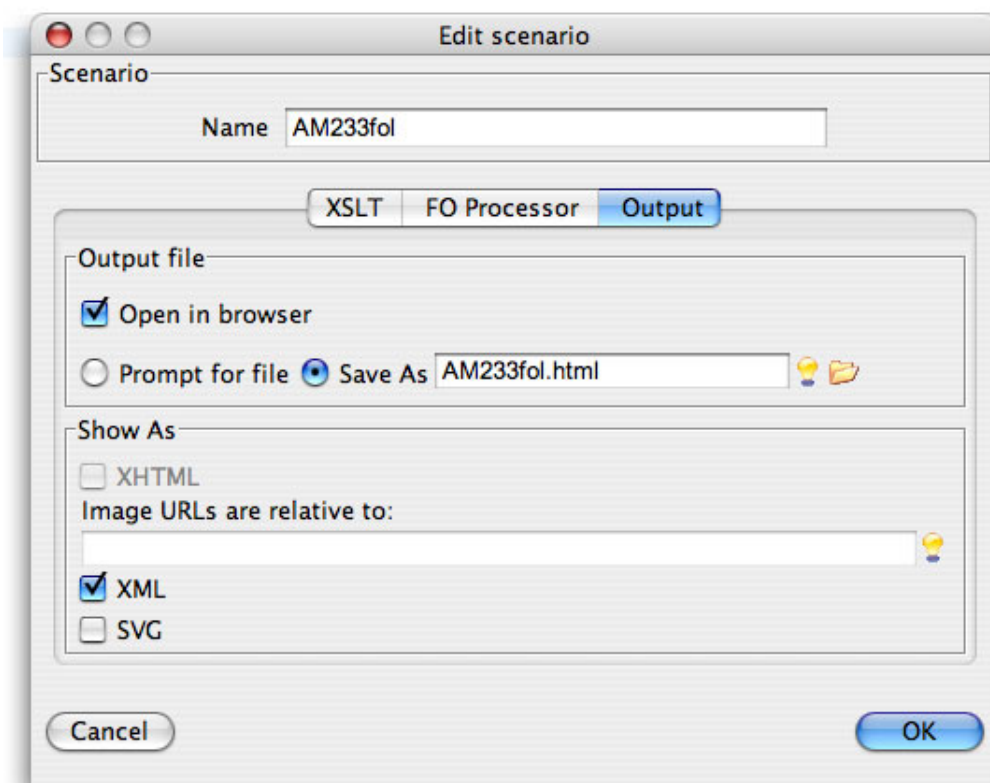
In Oxygen, name your scenario and choose your stylesheet (e.g menota-web-comment.xsl):
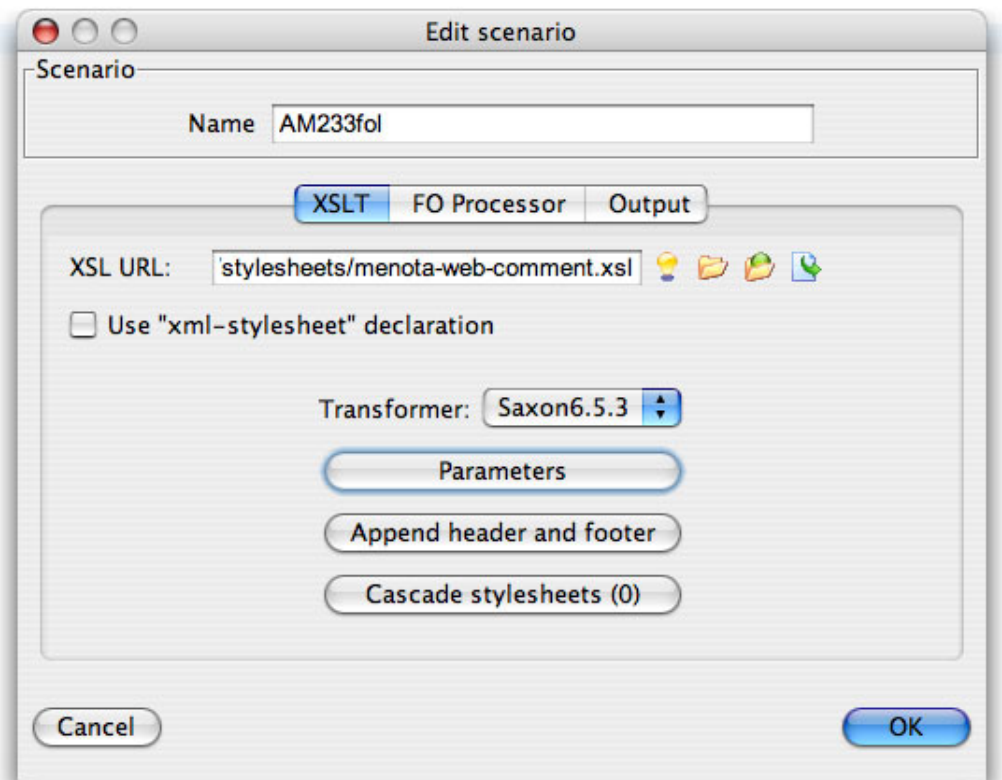
Then, set the desired parameter, e.g. for display ('visning') on the normalised level:

Finally, define a suitable output format. If giving the extension .html, the output file can be viewed immediately in a browser:

2. By defining one XSLT stylesheet for each level. In that case, you should switch comments in this pair of lines found at the beginning of the XSL file:

```
<!--<xsl:param name="visning" select="'norm'"/>-->
<xsl:param name="visning"/>
```

2.1. For **<facs>** display, make a stylesheet with this pair of lines:

```
<xsl:param name="visning" select="'facs'"/>
<!-- <xsl:param name="visning"/> -->
```

2.2. For **<dipl>** display, make a stylesheet with this pair of lines:

```
<xsl:param name="visning" select="'dipl'"/>
<!-- <xsl:param name="visning"/> -->
```

2.3. For **<norm>** display, make a stylesheet with this pair of lines:

```
<xsl:param name="visning" select="'norm'"/>
<!-- <xsl:param name="visning"/> -->
```

After having defined (and named accordingly) the XSLT stylesheets, you may generate HTML output in e.g. Oxygen as specified in (1) above. This time, you should not specify any parameter, since this has been done in the stylesheet itself (and for this reason you should define as many stylesheets as there are parameters).

F.2.2 Using an external transformator

If you do not have a version of Oxygen, there is an online solution available on the Henrik Ibsen project pages. Open this link:
http://gandalf.aksis.uib.no/cgi-bin/menota/xml/XSLT_konvert.cgi

Upload your XSLT stylesheet and XML file, and press 'Submit', as shown here:

Medieval Nordic Text Archive
www.menota.org ■ ■ ■ ■

Arkiv for nordiske middelaldertekster
Arkiv för nordiska medeltidstexter
Safn norrænna miðaldatexta

On-line conversion of XML files to HTML using XSLT stylesheets

Converting local XML files, using standard Menota stylesheets:

Select a menotic stylesheet (v 2.0 beta):   `menota_dipl.xsl` ▼

Choose a local XML file: [   ]   Bla gjennom ...

Send inn | Reset Form

Last changed: January 25, 2007 TMB

Note that with this solution you must have one XSLT stylesheet for each parameter, as explained above. Also note that the XML file must be valid and refer to an external DTD. Thus, your XML file should look like this:

```
<?xml version="1.0" encoding="UTF-8"?>

<!DOCTYPE TEI SYSTEM "http://www.menota.org/guidelines-2/schemes/menotaP5.dtd">

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

. . .
</teiHeader>
</TEI>
```