

## **Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions**

*Andrej Bojadžiev*

### *Preface*

#### *Introduction*

These Guidelines are the first full introduction to the Repertorium Initiative XML model for the electronic description of Medieval manuscripts. The model is part of the activities of the *Special Commission for Computer Processing of Medieval Slavonic Manuscripts and Early Printed Books* to the Executive Board of the International Committee of Slavists. This version of the Guidelines differs significantly from all previous versions in both form and content. The most important new feature is the synchronization with the “P5: Guidelines for Electronic Text Encoding and Interchange” of the Text Encoding Initiative Consortium (TEI) XML model. There do remain differences between the TEI and the Repertorium manuscript description models, but we have made every effort to bring two as close as possible, taking into consideration primarily the three principal Medieval manuscript traditions in Europe – Slavic, Greek, and Latin. We hope that the result will contribute to the general discussion on how to encode a manuscript description in electronic form. A special Appendix at the end of this text explains and discusses remaining differences between the TEI and Repertorium schemas and guidelines.

This edition has been rewritten completely, and is not based directly on earlier attempts to document the Repertorium XML model. It contains many editorial changes, new examples, revised lists of elements and attributes, and discussions about a wide variety of topics.

Previous versions of these Guidelines circulated only in electronic form. The first was compiled for the Annual TEI Members’ Meeting in 2005 in Sofia. Subsequently, within the framework of the project „Slovo: Toward a Digital Library of South Slavic Manuscripts” (2007-2008, cf. Miklas & Miltenova 2008), funded by the *Austrian Science and Liaison Offices*, a new electronic version was prepared. Both are now

archived and available on request. The present, new *Guidelines* are based on the ideas and proposals contained in various project papers, proceedings and special volumes (Birnbaum & Miltenova 2000, esp. Miltenova & Bojadžiev 2000, Bojadžiev 2000, Dobрева 2000, Stojkova 2000, Radoslavova 2000; Miltenova et al. 2003, esp. Cleminson 2003, Bojadžiev et al. 2003, Mussakova 2003, Krushelnitskaya 2003; Miklas & Miltenova 2008, cf. Birnbaum et al. 1995; Cleminson 1998; Gagova 1998, 2000; Radoslavova 1998; Stojkova 1998), as well as in the Commission's journal *Scripta & e-Scripta* (e.g., Miltenova 2008; Boyadzhiev 2009). These ideas have also been presented and discussed at several conferences and workshops. It goes without saying that this published document can reflect only the current level of our work and the model may undergo further change, but we consider the present version representative of a stable stage in the development of the Repertorium schema that has been used effectively in production.

### *About this document*

This document is provided as an introduction to the XML model for the description of Medieval Manuscripts known as the Repertorium Initiative or Template for Slavic Manuscripts (*TSM*). It is intended for those who have some previous knowledge of XML technologies who would like to begin to describe Medieval manuscripts. It should not be considered a general introduction to markup technologies, nor a tutorial on the TEI approach to linguistic and literary markup. For the latter, see the fifth chapter of the current P5 edition of the TEI Guidelines „A Gentle Introduction to XML” (Burnard & Bauman 2007).

### *Acknowledgements*

Although I can not possibly do justice to everyone who has provided useful comments on or changes to the XML model described in these Guidelines, or by editing and proofreading various passages of the text, I owe special thanks to the following for their patience and courage: Anisava Miltenova, Margaret Dimitrova, Dilyana Radoslavova, Antoaneta Granberg, David J. Birnbaum, Ralph Cleminson, Sebastian Kempgen, and Predrag Matejić. All errors that remain are, of course, my own.

### *Abbreviations and special signs used*

#### *Special signs and reserved words in the code*

- > – end of a comment
- , – sequence, followed by
- ;- – end of entity
- ? – must appear 0 or 1 time (optional)

?> – end of a processing instruction  
\, " – (paired) enclose an attribute value  
@ – attribute, e.g., read @type as *attribute type*  
\* – 0 or more times (optional and repeatable)  
% – introduces a parameter *entity*  
+ – 1 or more times (required and repeatable, i.e., must appear at least once)  
< – less than (recognised as reserved entity &lt;);, start of a declaration of any sort  
<! – start of an XML declaration  
<!-- – start of a comment  
<!DOCTYPE – start of a document type declaration  
<!ELEMENT – start of an element declaration  
<!ENTITY – start of an entity declaration  
<? – start of a processing instruction  
< . . . /> – empty element (self-closing; combining start and end tag in one)  
< . . . > – start tag  
</ . . . > – end tag  
> – greater than (recognised as reserved entity &gt;);, end of a declaration of any sort  
| – *alternation, or*  
a . – abbreviation for *attribute class*  
CDATA – in attributes, any valid *character data*, including spaces or punctuation marks  
DOCTYPE – *document type*, encloses the document type declaration

### *Abbreviations and Acronyms in the text*<sup>1</sup>

*BCP*: Best current practice (<http://www.rfc-editor.org/categories/rfc-best.html/>)

*CSS*: Cascading Stylesheet (<http://www.w3.org/TR/CSS2>, Bos et al.2008)

*ISO*: International Organization of Standardization <<http://www.iso.org/>>

*MASTER*: Manuscript Access through Standards for Electronic Records <http://www.ucl.ac.uk/~ucgadkw/mscat.html/>

*MS* – manuscript

*MSS* – manuscripts

*OASIS*: The Organization for the Advancement of Structured Information Standards

<http://www.oasis-open.org/>

*PDF*: Portable Document Format (<http://www.adobe.com/products/acrobat/adobe.pdf.html>)

*PI*: Public Identifier (Bray et al. 2008)

*RFC*: Request for comments <http://www.ietf.org/rfc.html/>

*SGML* – Standard Generalized Markup Language (ISO 8859:1987)

*SI*: System Identifier (cf. Bray et al. 2008)

*TEI*: Text Encoding Initiative Consortium (<http://www.tei-c.org>)

*W3C*: World Wide Web Consortium (<http://www.w3.org>)

---

<sup>1</sup> More information on standards, recommendations and specifications mentioned is available in the *References section* at the end of the text.

*XHTML*: Extensible HyperText Markup Language (Altheim, McCarron 2001)  
*XLink*: XML Linking Language (DeRose et al. 2001)  
*XML*: Extensible Markup Language (Bray et al. 2008)  
*XPath*: XML Path Language (Clark, DeRose 1999)  
*XPointer*: XML Pointer Language (Grosso et al. 2003a, 2003b, DeRose 2003)  
*XSL*: Extensible Stylesheet Language (Berglund 2006)  
*XSL:FO*: Extensible Stylesheet Language Formatting Objects (cf. XSL)  
*XSLT*: Extensible Stylesheet Language Transformations (Clark 1999)

## ***Introduction to the model***

### *Short History*

The Repertorium Initiative started during 1994-1995 as „Computer-Supported Processing of Medieval Slavic Manuscripts”, an international project with coordinators Prof. David J. Birnbaum, University of Pittsburgh, USA and Dr. Anisava Miltenova (Institute of Literature, Bulgarian Academy of Sciences). The project was supported by the International Research & Exchanges Board (IREX), Washington, DC. During this period an initial SGML model was developed by David J. Birnbaum, Anisava Miltenova, Milena Dobрева (Institute of Mathematics and Informatics, BAS, Bulgaria), Berend Dijk, and Harry Gaylord (both from the University of Groningen, the Netherlands). Further work on defining content models, especially for textological, linguistic and palaeographical features, was implemented in 1995 by David J. Birnbaum, Anisava Miltenova and Andrej Bojadžiev.

During the next three years (1996–1998) the project continued as a joint initiative under the guidance of Prof. Ralph Cleminson (Central European University – Budapest, Hungary and the University of Portsmouth, Great Britain) and Dr. Anisava Miltenova, with the name „Computer Supported Processing of Medieval Slavic Manuscripts and Early Printed Books” (a joint project supported by Central European University). At this time more than one hundred manuscript descriptions were encoded electronically and the project gained in popularity among scholars.

At the Twelfth International Congress of Slavists in 1998 in Kraków the need for markup technologies and standards and the lack of unified and well-established terminology in the field of Slavic Medieval Studies became obvious, and a new „Commission for the Computer Processing of Medieval Slavonic Manuscripts and Early Printed Books” was chartered by the Executive Board of the International Committee of Slavists (Miltenova et al. 1998). Members of the Commission are scholars from more than twelve countries and the main function of the Commission is to serve as a coordinator among various projects in the field.

In 1999-2000 the SGML model was transferred to XML syntax by David J. Birnbaum. The team decided then to change not only the markup language, but also the

declarations of some elements and attributes, both to improve the overall structure and to enhance the functionality of the schema. This stage of development was completed in 2001-2002. Since then the minor changes and improvements that have been made to the model are, in most cases, backwards compatible. An experimental *Repertorium Workstation* database was build upon the Exist Native XML database (<http://exist.sourceforge.net/>) by David J. Birnbaum (Birnbaum 2009). It uses SVG technology to represent textological differences and similarities graphically.

This model was presented at conferences in Bulgaria, Poland, Greece, the Netherlands, France, and the USA. Additionally, participants in the Repertorium project conducted courses, lectures, and workshops about the Repertorium in Bulgaria, Sweden, Hungary, Germany, the USA, Serbia, and Russia.

### *Sources for the XML model*

The Repertorium model relies fully on structured markup technologies. The first variant was compiled in 1994-1995 as an SGML Document Type Definition (ISO 8859:1987), based on TEI P3. The model presented here is based on Extensible Markup Language (XML), a simplified and web-oriented variant of SGML, and specifically on the Text Encoding Initiative Consortium (TEI) P5 schemas. The Repertorium model reflects fifteen years of electronic philological descriptions of more than three hundred Slavic Medieval and Early Modern Manuscripts.

From the beginning, the Manuscript Description Model, also called the Template for Slavic Manuscripts (*TSM*), has been related closely to the ideas of the Text Encoding Initiative and the schema fragments used there (TEI). The current Repertorium Initiative XML model for manuscript description is build on top of TEI P5. For the publication of Repertorium descriptions in electronic form, the project relies on XSLT scripts that transform the input XML to XML, XHTML, and SVG.

### *Text Encoding Initiative Consortium and Repertorium*

The Repertorium Initiative uses the TEI models as a principal framework for encoding and interchange. However, the Repertorium model extends some of the TEI element and attribute definitions, thus providing opportunities for the encoder to describe fully specific features of Medieval Slavonic texts. Simultaneously with our work, the TEI Consortium published its own model for manuscript description, which is based primarily on the MASTER project, a European Initiative for the description of Western European Manuscripts. The main differences between the Repertorium and TEI models are described in full in Appendices to these Guidelines, along with formal definitions of Repertorium-specific changes and additions to the TEI P5 models.

### **Standards and recommendations implied in TEI P5**

Like every project that relies on TEI P5, the Repertorium 3.0 model is repre-

sented in XML 1.0 DTD (Bray et al. 2008) and Relax NG formats (Clark, Murata 2001, Clark 2002), and provides the possibility of transformation into other available schema languages, which currently means W3C Schema language (Shudi, Sperberg, Thomson 2009, Peterson et al. 2009).

From the TEI P5 the Repertorium inherits the possibility of using all modern standards technologies from the XML family, including, for example, XSL<sup>2</sup>, XML Namespaces (Bray et al. 2006), XQuery language (Boag et al. 2011), XML Base (Marsh, Tobin 2009), and others.

Repertorium 3.0 inherits all constructions in the syntax of attribute values from the TEI P5 infrastructure, where they are, in turn, expressed according to recommendations of the W3 Consortium and ISO standards, and various RFCs and BCPS. The attribute values for language codes follow BCP 47 (Philips, Davis 2006c), which comprises RFC 4646 (Philips, Davis 2006a) and RFC 4647 (Philips, Davis 2006c), which, in turn, follow ISO standards (cf. ISO 639-1, ISO 639-2). Four-letter codes for the names of scripts are from ISO 15924 (ISO 15924). Data in attributes that provide pointers to other locations in the same or in other documents (e.g., target or lemmaRef) should be expressed in terms of RFC 3986 (Berners-Lee et al 2005). Temporal expressions in attributes follow the W3C XML Schema recommendation or ISO 8601 (cf. Shudi, Sperberg, Thomson 2009, Peterson et al 2009, ISO 8601).

#### *Visualization and publication*

Previous versions of the Repertorium use the XLink recommendation (DeRose et al. 2000) and CSS techniques (Bos et al. 2008) to visualize XML data directly without transformation. In the current version this approach has been abandoned in favour of transformation to XHTML (Althem, McCarron 2001) and possibly to PDF through XSLT and XSL-FO scripts.<sup>3</sup> TEI P5 versions are managed with an XSL framework (made available by Sebastian Rahtz)<sup>4</sup> which can be further customized to meet the needs of a particular project.

#### *Public Identifier and XML Namespaces*

Until now the Repertorium Initiative has had no Public Identifier (PI), and instead used system identifiers (SI) and XML catalog files (cf. Bray et al. 2008, Walsh 2001). Beginning with the current version of the Repertorium guidelines, the following PI will be used: "-//Repertorium//DTD Repertorium XML V3.0/EN".

For use with modern schema languages an XML Namespace designation is needed. A possible candidate is the `xmlns re` with the value "<http://www.ilit.bas.bg/repertorium/ns/3.0>". All elements from the TEI namespace `tei` preserve their value despite the fact that their content could be changed in some way, and only elements that are not present in the TEI models are implemented in the Repertorium namespace.

---

<sup>2</sup> Which in turn consists of the XPath (Clark, DeRose 1999), XSLT (Clark 1999) and XSL-FO (Berglund 2006) languages.

<sup>3</sup> See, in the TEI P5 Guidelines, the chapter about linking and segmentation, which represents XPointer schemes (Burnard, Bauman 2007: 16.2.4.TEI XPointer schemes, cf. Grosso et al. 2003a, 2003b, DeRose 2003)

<sup>4</sup> XSL Stylesheets for TEI XML <<http://tei-c.org/Tools/Stylesheets/>>

### *The modules*

The Repertorium Initiative was intended from the ground up to be constructed as a model for the description of medieval Slavic manuscripts. However in its XML implementation it also has the ability to serve as a framework for the description of other medieval traditions, such as Greek or Latin, because the philological approaches for such investigations are much more similar than different. The structures of catalogues of medieval European manuscripts share most of the same features and should give answers to the same questions related to manuscript contents, palaeography and codicology.

Using the TEI infrastructure makes it possible to implement several parts of an electronic text project within the same model. For example, it becomes possible to express all imaginable aspects of text edition, source description and graphical representation of a text object, providing ready-to-use models at least in the following areas:

- description of the source (in our case, a manuscript)
- edition of the text (with the appropriate segmentation into separate text units, apparatus criticus, annotation of linguistic features, pointers to the graphical representation of the original source, etc.)
- encoding bibliographic records related to description and edition
- encoding metadata, including terms of use
- construction of a model of a print dictionary
- technical documentation of the XML infrastructure (model classes, elements, entities, attributes, etc.)

The TEI model that has served as the foundation for the Repertorium schema is designed specifically to support the description of medieval manuscripts.

The required modules include:

1. `tei` – the main module of the TEI Infrastructure (Burnard, Bauman 2007: *1 The TEI Infrastructure*);
2. `header` – common metadata information (Burnard, Bauman 2007: *2 The TEI Header*);
3. `textstructure` – the default text structure (Burnard, Bauman 2007: *4 Default Text Structure*);
4. `core` – common core (Burnard, Bauman 2007: *3 Elements Available in All TEI Documents*);
5. `linking` – linking, segmentation, and alignment (Burnard, Bauman 2007: *16 Linking, Segmentation, and Alignment*);

In addition to these required modules several optional ones have been included:

6. `analysis` – associates simple analyses and interpretations with text elements (Burnard, Bauman 2007: *17 Simple Analytic Mechanisms*);
7. `certainty` – elements and attributes to indicate the degree of certainty attached to the encoded text (Burnard, Bauman 2007: *21 Certainty and Responsibility*);
8. `figures` – representation of tables, formulae, figures and graphics (Burnard, Bauman 2007: *14 Tables, Formulae, and Graphics*);
9. `namesdates` – representation of names (of people, places, organizations), times and dates (Burnard, Bauman 2007: *13 Names, Dates, People, and Places*);



10. `transcr` – transcription of primary sources (Burnard, Bauman 2007: *II Representation of Primary Sources*).

In the most recent Repertorium model, based on TEI P4, the modules for drama and performance, verse and feature structures were represented. In practice they have never been used in our work, and they have therefore been removed from the current model.

In comparison with the customized TEI P5 manuscript description schema, three more modules have been included: analysis, certainty and figures.

#### *Some principles*

It was necessary to make a very precise comparison between the TEI P5 module for manuscript description and the Repertorium model in order to reach the maximum level of compatibility and interchangeability. To fulfill this task several principles were formulated to govern that transformation of the prior Repertorium schema into the present one:

1. All attributes and elements from the core and additional tag sets that have been changed from TEI P4 to TEI P5 are changed in Repertorium 3.0, as well.

2. All attributes and elements that share the same functionality between TEI P5 and Repertorium are renamed in favour of TEI P5.

3. No customizations are made on the level of TEI macro structures or on the level of model classes with the exception of

- model classes related to specific manuscript description elements
- exclusion of all numbered *divLike* models and their corresponding elements. This includes removal of `<div1>`, `<div2>`, `<div3>`, `<div4>`, `<div5>`, `<div6>`, and `<div7>`.

4. When there is an element with the same name in TEI P5 and the Repertorium, it is declared in the TEI, not the Repertorium, namespace, even if the given element name has its roots in Repertorium Initiative.

5. When there is no functional possibility in the Repertorium model to express particular features, the elements from TEI P5 with their names are taken without further customization. This is the case, for example, with element `<additional>`.

6. When the same information with the same functional characteristics is expressed in different places in TEI P5 and the Repertorium, the choice is made in favour of the former. This is the case with the definition of manuscript origin. In early versions of the Repertorium schema this was expressed with a special element `<manuscript-creation>`, placed before the content description. However, in TEI P5 there is an element `<origin>` with a similar function, which is part of the manuscript history, placed almost at the end of the schema. In this case the latter element has been chosen.

7. When elements with different names share the same functional characteristics even if their declarations differ to some extent, the elements from the TEI namespace are chosen. This is the case, for example, with the element `<msItemStruct>`, which replaces the original Repertorium `<articleContentDesc>`, and the element



<filiation> with the special attribute @type which replaces <antigraph>, <apograph>, <litRedaction>, <protograph>, <source> and <translation>.

8. If one of the models provides a more general and simpler view and the other has a structured model of the same information, an attempt is made to integrate both approaches in the declaration of element. This is the case, for example, with the <bindingDesc>. The declaration of this element in TEI P5 is:

```
<!ELEMENT bindingDesc ((%model.pLike; | decoNote | condition)+ |binding+)>
```

Then the content of <binding> is declared as

```
<!ELEMENT binding (%model.pLike; | condition | decoNote)+>
```

The proposed model leaves the declaration of <bindingDesc> without changes but modifies the content of the element <binding> to include structured information:

```
<!ELEMENT binding (%model.pLike; | bindingNote |condition | decoNote)+>
```

9. If there is an element in the Repertorium that has the same name as in TEI P5 but a different meaning, the Repertorium element is removed or renamed. An example of such an element is <writing>, which was used in earlier Repertorium schemas to provide general palaeographic information. An element with the same name exists in P5 with a completely different meaning („a passage of written text revealed to participants in the course of a spoken text”, cf. Burnard, Bauman 2007), so the Repertorium element was changed to <summary> in the schema.

10. Specific Repertorium elements and structures from earlier versions are retained only if there are structural or semantic differences between the Repertorium and TEI models. One example is the element <scribeDesc>, which represent different points of view for manuscript description and is presented in a different way by the TEI P5 Guidelines.

11. If in the TEI P5 Guidelines an element can be used with more than one meaning, then an attempt is made to clarify the usage. This pertains to elements such as <msIdentifier>, <altIdentifier>, <msPart> and <accMat>.

### *The main Structure*

A manuscript description can occur in two different places in a Repertorium XML document. When the description (cataloguing information) serves as an introduction to the digital representation of transcribed text, it is encoded inside the <tei-Header> element that provides metadata for the edition itself. However, when the goal of the project is focused on manuscript descriptions, without transcribed full text, the descriptive elements and content are regarded as primary data and the description itself becomes a primary object. In this case it serves no longer as metadata and must be entered as part of the content of the element <body>.

The basic, wrapper element for a description is called <msDesc> (*manuscript description*) and it serves as a container for seven additional parts: general cataloguing

information (<msIdentifier>); headings (<head>); information on physical characteristics and layout of the manuscript (<physDesc>); palaeographical and linguistic data (<scribeDesc>); the description of the manuscript content as a whole, as separate texts, and its notes and colophon(s) (<msContents>); conclusions about the time and place of writing of the manuscript and a sketch of the manuscript's history (<history>); as well as a place for additional information the describer may need to encode (element <additional>).

Convolutives and palimpsest parts are described as part of the element <msPart><sup>5</sup>. Additional notes by the authors of earlier descriptions and by owners of the manuscript form part of the element <accMat>.

The specification of the source could be extended further as part of the <recordHist> (*recorded history*) element at the end of the description (cf. *Additional information*).

Only one part of this information is mandatory and must always appear in the description: <msIdentifier>. All other main elements are optional.

Information is entered step by step, in an ordered and purposeful way. The first step concerns general cataloguing information, then comes information related to physical features of the manuscript, palaeography, language, content, sources of manuscript description and, finally, administrative information, such as custodial history, restoration, available photocopies and bibliographic information.

#### *Levels of representation*

Descriptions of medieval manuscripts may differ significantly both in structure and length. There exist many traditions in cataloguing medieval materials, which means that similar information can look different in different frameworks. The first thing to consider is the completeness and length of the description.

#### *Description with the <msIdentifier> and <head> elements*

The briefest possible description consists of the elements <msIdentifier> and <head>, as shown in the following example:

```
<msDesc>
  <head>Homilies from Macarius of Egypt, Russian,
17th c., paper, XI+229 ff., 185x145 mm. Slavic
translation from Greek from the year 1599. Notes from
the scribe Stefan Jovanović</head>
  <msIdentifier>
    <settlement>Belgrade</settlement>
    <repository>Patriarchal Library</repository>
    <idno>178</idno>
  </msIdentifier>
</msDesc>
```

(Богдановић 1982: 25, № 153)

---

<sup>5</sup> Cf. *Description of convolutives*, *Description of palimpsests*

Even this check-list description could be further segmented into more discrete units, so that the above information could be recorded differently from the encoder's point of view as

```
<msDesc>
  <head type="briefDesc">
    <term type="textClass">Homiles</term> from <name
type="author">Macarius of Egypt</name>, <term type="lang">Russian</
term>, <origDate>17th c.</origDate>,
    <material>paper,</material><seg type="extent"> XI+229 ff.,</
seg>
    <dimensions type="folia" unit="mm"><height>185</
height><width>145</width></dimensions>
    Slavic translation from Greek from the year <date
type="translation">1599</date>. Notes from the scribe <name
type="scribe">Stefan Jovanović</name></head>
  <msIdentifier>
    <settlement>Belgrade</settlement>
    <repository>Patriarchal Library</repository>
    <idno>178</idno>
  </msIdentifier>
</msDesc>
```

We attach @type to the element <head> to make a distinction between the proper headings in some editions and the heading that serves as a brief statement here. Next we encode each valuable piece of information by giving it a name. A possible rationale for this more complete approach would be the extraction of or search for specific pieces of data, or the attachment to it of specific formatting during rendering.

*Description with <msIdentifier> and paragraphs*

An alternative to the preceding involves the use of the element <p>, which can unite all data related to physical description, palaeography, language and the manuscript's provenance and history, e.g.:

```
<msDesc>
  <msIdentifier>
    <country>Bulgaria</country>
    <settlement>Sofia</settlement>
    <repository>Scientific Archives of the
BAS</repository>
    <idno>90</idno>
  </msIdentifier>
  <p>Damaskin of priest Joan from 1788, 272 ff., paper of two types, both of bad quality, cf.
Nikolaev 473b (1751)
```

608b (1779), 664 (1783). 1 column, 23-29 lines per page.

Literature: Иванов 1891: 83-112, 1892: 169-193; Цонев 1940: 180-181; Ангелов 1958;

Кодов 1969 : 227-232 </p>

</msDesc>

(From the electronic description of MS BAS 90 by Elena Tomova)

If we decide to use just paragraphs, we can nonetheless select whatever granularity we consider most appropriate to our purposes. For example, we can divide the preceding information into smaller units, as in this example:

<msDesc>

<msIdentifier>

<msName>Leitourgikon</msName>

<country>Greece</country>

<geogName>

<geogFeat>Mount</geogFeat>

<name>Athos</name>

</geogName>

<placeName>Zograph Monastery</placeName>

<repository>Monastery's library</repository>

<idno type="shelfmark">III.r.7</idno>

...

</msIdentifier>

<p xml:id="material">Paper - thin, rough and white. Watermark: scissors, similar to Briquet, Nr 3688 from the year 1496.</p>

<p xml:id="binding">Binding: new, made in the library.</p>

<p xml:id="writing">Writing: middle semi-uncial, 23 lines per page. Without decoration.</p>

<p xml:id="language">Serbian.</p>

<p xml:id="photo">Photograph from the f. 1r</p>

<p xml:id="biblio">Райков, Кожухаров, Миклас, Кодов 1994, tabl. 176.</p>

</msDesc>

(Райков и др. 1994: 105)

In this example the cataloguing model consists of the identification of the codex (<msIdentifier>) and a sequence of paragraphs, each of which provides information that is identified with the help of @xml:id. The choice of paragraph types can be determined by the goals of the individual project.

*Extraction of information from a full description*

Another way to obtain information in the form of a checklist or short catalogue is

to extract already existing data from an analytical description. For this purpose, the form of the full-length catalogue must follow encoding rules that will allow data extraction without loss of information. For example, small summaries or overviews might coexist with more elaborated, structured features of manuscript or text analysis, e.g.:

```
<msContents>
  <summary>The fragment contains 11 Lenten lessons; the last one includes only the first
  three verses from 2nd Lent
  Sunday to the 6th Sunday at Orthros (Palm day)</summary>
    <msItemStruct xml:id="ACD1"><locus n="1">lr</locus>
  <title>Mc 1:35-44;</title>
    <note>The beginning is missing. <date>Saturday II
  </date></note>
    </msItemStruct>
    <msItemStruct xml:id="ACD2"><locus n="2">lr-lv
  </locus> <title>Mc 2:1-12 </title>
  ...
    </msItemStruct>
</msContents>
```

(From an electronic description of NBKM 848  
by Margaret Dimitrova)

In this example the element `<summary>` is used as an introduction to the more extensive, formalized description of each text from the manuscript. This hybrid format permits the use of *XSLT* or *XQuery* to extract whatever information might be needed.

## ***Cataloguing Information***

### *General remarks*

Cataloguing information is an obligatory part of each manuscript description and it should provide an introduction for researchers consisting of:

- identification of manuscript name, as a type or genre (e.g., *Menaion* or *Triodion*) or as a specific name that has become standardized in the research tradition (e.g., *Euchologium Sinaiticum*);
- information about where the manuscript is kept, together with its shelfmark(s);
- optional additional notes

The description thus begins with answers to questions like: What is the name of the manuscript? Where it is now kept? and Are there earlier descriptions? The answers to all of these questions constitute part of the content of the element `<msIdentifier>` (i.e., manuscript identifier, Burnard & Bauman 2007).

The encoder can provide information about the name of the manuscript (`<msName>`) if it exists, and about the country, geographical region and/or city, and the institution, repository, or collection (Burnard & Bauman 2007: `model.placeNamePart`).

Between `<msName>` and following structured information it is possible to insert a brief summary description in the form of paragraph (cf. below). More information, including optional bibliographic records or lists of records, may be entered as part of the element `<note>`.

The element `<msIdentifier>` (manuscript identifier) has `@type` and `@subtype` attributes. These attributes may be used, for example, when the manuscript is now lost or its place is now unknown, so that the type of identifier is not current, and the electronic form represents an encoding of an old description, and not the *de visu* work of the current encoder.

### *The Manuscript Name*

The name of the manuscript is a word or a short phrase that can be used to classify the manuscript according to a known scheme and to provide a special or supplied name that has become established by the research tradition.

There are many characteristics involved in assigning a name to a manuscript. The most important is the so-called generic or unified name, based on typological characteristics, with values such as *Psalter* or *Miscellany*. In some cases manuscripts may also already have unique names, such as *Codex Zographensis* or *Kodov Triodion*. These names should be provided separately.

This element may contain only plain text. The `@type` attribute may be used to give further information whether the name is *general* (i.e., unified or generic) or *individual* (nickname).

The following example from the description of *Codex Marianus* demonstrates the use of this element:

```
<msName type="general"
xml:lang="la">Tetraeuangelion</msName>
<msName type="general"
xml:lang="bg">Четвероевангелие</msName>
<msName type="individual" current="yes"
xml:lang="bg">Мариинско евангелие </msName>
<msName type="individual" subtype="current"
xml:lang="la">Codex Marianus </msName>
<msName type="individual" subtype="old"
xml:lang="en">Grigorovič tetraeuangelion</msName>
<msName type="individual" subtype="old"
xml:lang="en">Athonite tetraeuangelion</msName>
```

Here we have the following information: the general information on the manuscript name is encoded with the values *general* or *individual* of the `@type` attribute in Latin and in Bulgarian. The manuscript name currently accepted among scholars is *Codex Marianus* and the older names are *Grigorovič tetraeuangelion*, and *Athonite tetraeuangelion*.

When the manuscript has different nicknames for separate parts that are kept in different repositories, this information goes into corresponding sections of description. As in case with *Codex Marianus*, there are two leaves now kept in Vienna that have been known by two names that are no longer generally used: Mihanović folia and Miklošič folia:

```
<altIdentifier xml:id="OENB146">
  <msName type="individual" subtype="old">Mihanović
folia</msName>
  <msName type="individual" subtype="old">Miklošič
folia</msName>
  ...
</altIdentifier>
```

*A paragraph between the name of the manuscript and its location*

The use of this paragraph is restricted to cases when special annotation is needed to summarize the locations of different parts of the same manuscript. The current location in different countries, archives, or collections could require a special introduction to this state with possible links to sections on manuscript history, acquisition, and to the various signatures or numbers in the catalogues. The following extract from the description of *Manuilev Apostolos* illustrates this approach:

```
<p>Six fragments belonging to an once larger codex have
been identified by Tixomirov ...</p>
```

(From the electronic description of Manuilev Apostle  
by Margaret Dimitrova and Andrej Bojadžiev)

This short introduction to the cataloguing information in the description of the Manuilev Apostle is needed to understand the rest of the description. The formal declaration of the element `<p>` (meaning *paragraph*) is used exactly as the same element in TEI models (Burnard & Bauman 2007: p).

## ***Location of the Manuscript***

### *General Remarks*

The next part of the electronic cataloguing information is devoted to the place where the manuscript is kept. There are three levels for this information: geographical pointers, institutional identification, and (optionally) data regarding the collection in which the manuscript is kept. The latter often appears as part of the shelfmark designation.

The specified order of elements is important, going from largest to smallest:

1. country
2. geographical region



3. city or place name
4. repository
5. collection
6. shelfmark, inventory number or some other designation

This set of elements is followed by optional information on identification based on *secundo folio* and bibliographic records.

Following example from the description of Ms *NBKM 848* illustrates the usage of the element:

```
<country>Bulgaria</country>  
<settlement>Plovdiv</settlement>  
<repository>Plovdiv National Library</repository>  
<idno type="shelfmark">25</idno>
```

(From the electronic description of MS Plovdiv 25 by Desislava Atanasova)

### *Country*

The most simple content of the element `<country>` is the name of country, e.g.:  
<country>Bulgaria</country>

In some cases it may be useful to provide a country code as well as a country name. This can be achieved by supplying a value to the `@key` attribute. It is recommended by the TEI that this value follow the country names and codes defined by ISO 3166 (Burnard & Bauman 2007: `country`). The encoder can choose among three representations defined in the standard: alpha-2 code (two-letter, ISO 3166-1-alpha-2-code), alpha-3 code (three-letter, ISO 3166-1-alpha-3-code), and numeric code (ISO 3166-numeric). The three variants are illustrated below:

```
<country key="BG">Bulgaria</country>  
<country key="BGR">Bulgaria</country>  
<country key="100">Bulgaria</country>
```

For the full list cf. Common Data Hub: ISO 3166 Country Codes.

### *Settlement (city or town, etc.)*

The next part of the information is related to the place where the manuscript is currently kept. Two different elements are provided for this purpose. In most cases the manuscript is kept in libraries or archives in cities, for which the element `settlement` is used:

```
<settlement>Sankt Peterburg</settlement>
```

The type of settlement may be further specified by adding a standard type, as in this example:

```
<settlement type="town">Kazanlāk</settlement>
```

### *Region*

When the name of place is not well-known or there could be two or more places with the same name, an additional element `<region>` may be added.

By <region> is meant an administrative unit smaller than a country but larger than a city, town or village (Burnard & Bauman 2007: region).

```
<region type="district">Sofia</region>
<settlement type="village">Alino</settlement>
```

### *The place and geographical names*

When the codex is kept in a specific place not normally considered a region or settlement, e.g., in a monastery, the alternative elements <placeName> and <geogName> may be used. These elements are part of the TEI element infrastructure and may contain plain text, sub-elements like <region> (for some places that are smaller than countries, but bigger than cities) or <district> (for names of non-administrative, geographical places). With the help of <geogFeat> we can encode such geographical features as valley, river, mount, etc. (Burnard & Bauman 2007: geogFeat). Then we can provide the name of the place using general purpose <name> element.

The choice in this case should be made on the level of the whole project, unifying the appearance of the information in order to format, transform or search the data. For example, a possible encoding for manuscripts kept at the Zograph Monastery on Mount Athos could look like the following:

```
<geogName>
  <geogFeat>Mount</geogFeat>
  <name>Athos</name>
</geogName>
<placeName type="monastery">Zograph monastery</placeName>
```

The same information could be encoded in two other ways. The first is using a general purpose element <name> with @type, e. g.:

```
<name type="monastery">Zograph monastery</name>
```

The other is to use a special referencing string as the value of the @key attribute, e. g.:

```
<placeName key="ZM" type="monastery">Zograph
Monastery</placeName>
```

In the last example the <placeName> element has two special attributes. As explained in the TEI Guidelines @key „provides an externally-defined means of identifying the entity (or entities) being named, using a coded value of some kind.” (Burnard & Bauman 2007: att.canonical/key) and @type is intended to be used as a general characteristic of the place name. That means that project should have a list of abbreviations of some sort, one of which must be *ZM* (= *Zograph Monastery*). For the manuscripts kept at Mount Sinai „St. Catherine” the same model is also valid, e.g.:

```
<country>Egypt</country>
<geogName>
  <geogFeat>Mount</geogFeat>
```

```
<name>Sinai</name>  
</geogName>
```

*Repository (library, archive, etc.)*

The next piece of information in the electronic description, <repository>, identifies a public or private library or archive. This element should contain a common or official designation, as in this example:

```
<country>Czech Republic</country>  
<settlement>Prague</settlement>  
<repository>National Museum</repository>
```

(From the electronic description of MS IX.H.16  
by Irina Kuzidova)

*Name of the institution*

Sometimes the repository part of another institution, and in such cases the name of the institution comes before the name of the archive, as in the following examples:

```
<institution>University of Sofia</institution>  
<repository>Central University Library</repository>
```

or

```
<institution>Bulgarian Academy of Sciences</institution>  
<repository>Scientific Archives</repository>
```

*Shelfmarks, inventory and catalogue designations*

This information forms the content of the element <idno>. This element has an @type value that specifies whether the identification is a shelfmark, inventory number, or catalogue number. This attribute value is especially important in archives that combine several collections that may be arranged and identified according to former owners or thematic principles. In practice, shelfmarks are often formed from the initials or special codes for separate collections. Typical examples could be:

1. shelfmark:

```
<idno type="shelfmark">Пор. 8</idno>
```

2. Designation in the catalogue:

```
<idno type="catalogue">1</idno>
```

3. Information on identification in inventory catalogue:

```
<idno type="inventory">5/1946</idno>
```

*Current and former shelfmarks, catalogue and inventory numbers*

The best way to differentiate between current and former designations is to wrap the element <idno> around <altIdentifier> elements. Consider the following examples:

1. Manuscript with shelfmark, inventory and catalogue numbers

```
<msIdentifier>
```

```
...
    <settlement>Sofia</settlement>
    <repository>National Library SS Cyril and Methodios<
/ repository>
    <idno type="shelfmark">1415</repositSignature>
    <altIdentifier>
        <idno type="inventory">2/977</idno>
    </altIdentifier>
    <altIdentifier>
        <idno type="catalogue">1415</idno>
    </altIdentifier>
...
</msIdentifier>
```

(From the electronic description of MS 1415 of NLCM, Sofia  
by Dilyana Radoslavova)

## 2. Manuscript with current and several former shelfmarks:

```
<msIdentifier>
...
    <country>Ukraine</country>
    <settlement>Odessa</settlement>
    <repository>State Library</repository>
    <collection>V. Grigirovich</collection>
    <idno type="shelfmark">1/102</idno>
    <altIdentifier type="former">
        <idno type="shelfmark">1/211</idno>
    </altIdentifier>
    <altIdentifier type="former">
        <idno type="shelfmark">20</idno>
    </altIdentifier>
    <altIdentifier type="former">
        <idno type="shelfmark">34</idno>
    </altIdentifier>
    <altIdentifier>
        <idno type="catalogue">20</idno>
    </altIdentifier>
...
</msIdentifier>
```

(From the electronic description of MS 1/102 of Odessa State Library  
by Anisava Miltenova)

In this example the manuscript has one current and three former shelfmarks and it is described in the catalogue as item number 20.

By setting the value of @type in the <altIdentifier> element the encoder

can choose between the values *current* and *former* and then further define values in its sub-elements, e.g.:

```
<altIdentifier type="current">
...
    <idno type="shelfmark">Ms 3045</idno>
...
</altIdentifier>6
```

### Collections

Some repositories use a complex archiving system with special divisions between the separate collections and archives. Very often the names of the sections bear the names of previous owners or commissioners, which can be represented with the `<collection>` element. The `@type` and `@subtype` attributes are available for cases where a more detailed characterization is necessary.

The `<collection>` element may contain unstructured (plain) text, e.g.:

```
<collection>П. Сырку</collection>
<idno>15</idno>
```

In this fragment it is stated that the collection belongs to the gatherings of P. Syrku under the number 15.

When the `<collection>` element appears, it is immediately followed by the shelfmark or inventory number as an identification inside the collection:

```
<msIdentifier>
...
    <collection>M. Pogodin</collection>
    <idno type="shelfmark">Pog. 1204</idno>
        <altIdentifier type="current">
            <idno type="catalogue">77</idno>
        </altIdentifier>
...
</msIdentifier>
```

(From the electronic description of a Vidin Miscellany  
by Antoaneta Granberg)

In some cases the current collection of some repository contains an entire earlier repository collection. This is often the case in Russian collections like the National Russian Library in Sankt-Petersburg, the Historical Museum, the State Russian Library and the Russian State Archive of Ancient Acts. In this case it is helpful to include this information, as well, e.g.:

```
<msIdentifier>
    <msName>Hitrovo Evangelion</msName>
```

---

<sup>6</sup> Other possible values could include *partial*, *internal* for a project's internal identifier, etc.

```
<country>Russia</country>
<settlement>Moscow</settlement>
<repository>Russian State Library</repository>
<collection>ф. 304. III.</collection>
<idno type="shelfmark">3</idno>
<altIdentifier type="former">
  <collection>M. </collection>
  <idno type="shelfmark">8657</idno>
</altIdentifier>
<altIdentifier type="former">
  <collection>Tr.-Serg. Lavra</collection>
  <idno type="shelfmark">3</idno>
</altIdentifier>
</msIdentifier>
```

*Various parts of the manuscript in one repository*

In this case the information about the country, city, geographical or place names and repository remain the same and only information about the shelfmarks, collections or catalogue numbers differ. This approach is illustrated with the description of the *Pirdop Apostle*:

```
<msIdentifier>
  <msName>Пирдопски апостол</msName>
  <msName>Pirdop Apostle</msName>
  <country>Bulgaria</country>
  <settlement>Sofia</settlement>
  <repository>National Library SS Cyril and Methodios</repository>
  <altIdentifier xml:id="NBKM497">
    <idno type="shelfmark">497</idno>
    <idno type="inventory">237</idno>
    <note place="inline"><ref type="bibl"
target=" ../bibliography/TsonevB1923">Цонев II, с. 70, Nr 497</ref>
</note>
  </altIdentifier>
  <altIdentifier xml:id="NBKM90">
    <idno type="shelfmark">90</idno>
    <idno type="inventory">588</idno>
    <note place="inline"><ref type="bibl"
target=" ../bibliography/TsonevB1910">Цонев I, с. 31-33, Nr 90</ref>
</note>
  </altIdentifier></msIdentifier>
```

(From the electronic description of the Pirdop Apostle  
by Dimitrinka Dimitrova)

*One and the same manuscript in different repositories*

When the manuscript is separated into different repositories the encoder should use the element `<altIdentifier>` to designate each part. There are two main cases:

1. If the manuscript is separated into one larger part and several smaller pieces, the cataloguing part of the description could be included immediately after `<msIdentifier>` and the other parts inside an `<altIdentifier>` element.

2. A better approach is to designate all parts inside `<altIdentifier>` elements. This approach could be used also when there is no unambiguous main part:

```
<msIdentifier>
  <msName xml:lang="bg">Слепченски апостол</msName>
  <msName xml:lang="en">Slepcha Apostle</msName>
  <altIdentifier type="partial" xml:id="SPLFPi101">
    <country>Russia</country>
    <settlement>St. Petersburg</settlement>
    <repository>State Public Library</repository>
    <idno type="shelfmark">F. p. I. 101</idno>
  </altIdentifier>
  <altIdentifier type="partial" xml:id="RAS2446">
    <country>Russia</country>
    <settlement>St. Petersburg</settlement>
    <repository>Russian Academy of Sciences
    </repository>
    <idno type="shelfmark">24.4.6.</idno>
    <collection>I. Sreznevskij</collection>
    <idno type="CS">Srez. 53</idno>
  </altIdentifier>
  <altIdentifier type="partial"
xml:id="RSL_Grig14">
    <country>Russia</country>
    <settlement>Moscow</settlement>
    <repository>Russian State Library</repository>
    <idno type="shelfmark">Grig. 14</idno>
    <collection>Grig.</collection>
    <altIdentifier type="former">
      <idno type="shelfmark">M. 1696</idno>
    </altIdentifier>
  </altIdentifier>
</altIdentifier>
```



```
    </altIdentifier>  
</msIdentifier>
```

*Once again about the usage of <altIdentifier> and <idno> elements*

In an analytical catalogue where there may be combinations of former and current collections and shelfmarks in a repository it is often useful to encode all these designations. The encoder can then further explain the history of the manuscript in sections about its origin, provenance and acquisition.

The usage of <altIdentifier> in the Repertorium approach as a subelement of <msIdentifier> is to define the current or old identification of a codex or to distinguish among parts of one and the same manuscript dispersed in one or in various repositories.

The element <idno>, on the other hand, is used for various kinds of designations, such as shelfmarks and inventory or catalogue numbers. In this approach, if we have a part of the manuscript in one repository with an old shelfmark or collection identification that is still in use, it is necessary to allow <altIdentifier> to appear elsewhere, in the description of other parts, as in:

```
<altIdentifier type="partial"  
    xml:id="RSL_Grig14">  
    <country>Russia</country>  
    <settlement>Moscow</settlement>  
    <repository>Russian State Library</repository>  
    <idno type="shelfmark">Grig. 14</idno>  
    <collection>V. Grigorovic</collection>  
    <altIdentifier type="former">  
        <idno type="shelfmark">M. 1696</idno>  
    </altIdentifier>  
</altIdentifier>
```

In this example both Grig. 14 and M. 1696<sup>7</sup> are used as identifications in the tradition in spite of the fact that the latter points to the previous owner. In this case the shelfmark signifies the intermediate change of owner. Further details should be given in the section about the history of the codex.

*Secundo folio*

This element is very rare in the Slavic medieval and early modern tradition, but for western medieval manuscripts it has been (mainly between the thirteenth and sixteenth centuries) a common component of manuscript description. The term refers to transcribing the first word or words from the second leaf of a codex. The element

---

<sup>7</sup> An abbreviation of Muzejnoe sobranie, or Museum collection. This points to the previous name of the library, which existed until 1925 as a private collection of Count H. Rumjancev.

<secFol> (Lat. *secundo folio*, or Eng. *second folio*) is used for this purpose (Burnard & Bauman 2007: secFol). The following example illustrates the usage:

```
<secFol>dixit ad me</secFol>
```

(Description of MS Crispin 16, Dutschke et al. 1987: 35–36)

### *Additional information in the catalogue description*

All other data could be entered at the end of <msIdentifier> or <altIdentifier> elements as part of the content of a <note> element. Typical information includes internal or external bibliographic or other pointers, notes about the attribution of a manuscript, etc.

## ***Codicological information***

### *General remarks*

The markup definition of the physical structure should be entered in the element <physDesc> (Burnard & Bauman 2007), which unites information about the physical condition of the manuscript with observations about the material, layout, binding and decoration. The encoder can choose between two options:

1. To make a very short sketch of different features of the manuscript in paragraphs, or
2. To use special elements for the description of material, layout, binding, decoration, etc.

For the former approach the encoder should use only paragraphs, as in this example:

```
<physDesc>
```

```
<p>4+516 ff. (Stojanov wrongly gives 273ff.). Paper with watermark three cylinders with letters FA. 230 x 180 mm. In 1 column, 30 lines per page</p>
```

```
<p>Greek cursive of the 19th century.</p>
```

```
<p>The original binding is preserved: leather on boards. Remains of clasps  
Blind-tooled frames of floral motifs and blind stamped central medallions on both covers.  
></p>
```

```
<p>The codex is in good condition. The folios containing the Table of contents  
are now unfolded. Originally they formed the last quire of the manuscript</p>
```

```
</physDesc>
```

(From the electronic description of MS CIAI 369 by Margaret Dimitrova, simplified)

The element <physDesc> has no special attributes. Its only use is to contain other elements, which encode the physical characteristics of the manuscript.

The second approach also permits the use of the element <p>, in this case for an introduction to the physical description.

### *Description of material and layout*

Information about manuscript materials and their characteristics and about the layout of the leaves is encoded within `<objectDesc>`. This element has one special attribute, `@form`. The proposed values (Burnard & Bauman 2007) of this attribute could be *codex*, *roll*, *fragment*, *leaf*, *partial leaf* or some typologically relevant value, e.g.

```
<objectDesc form="codex">
...
</objectDesc>
or
<objectDesc form="roll">
...
</objectDesc>
```

This information could be entered as an unstructured sequence of paragraphs, or as part of the content of the specialized elements `<supportDesc>` and `<layoutDesc>`.

### *Description of physical support*

The information in `<supportDesc>` could be also entered in two ways: 1) as an unstructured (paragraph) view, or 2) using specialized elements. Here we illustrate an overview of the use of structured data for various aspects of cataloguing the support.

`<supportDesc>` has one important attribute, `@material`. The value of this attribute, as suggested by TEI Guidelines (Burnard & Bauman 2007: `supportDesc`), is most commonly *parchment*, *paper*, or *mixed*, as in this example:

```
<supportDesc material="paper">
...
</supportDesc>
```

The meaning of the value of this attribute is to give a general type of the material of the manuscript, and it is not associated with the possible construction of manuscript's binding or the box in which the manuscript is kept. Especially in cases when the value is „mixed,“ a further explanation should be provided.

The order of information in this element is significant, and follows traditional paper-based descriptions. There is no required minimal content, and the range of information depends on the opportunities the researcher may have had to work with the manuscript *de visu*. The brief explanation of the proposed elements is as follows:

- `<support>` – describes in more detail the material of the leaves or roll
- `<extent>` – describes the size of manuscript in terms like number of folia or pages, or the length of the roll;
- `<foliation>` and/or `<pagination>` – information about whether the

manuscript is foliated (paginated), with further explanation;

<collation> – description of the folia collation, ruling, pricking and other techniques related to the quire structure;

<condition> – an overview of possible damage, absence or additions of folia, blank leaves, etc.

### *Detailed description of material*

This information could be entered as plain text in the form of paragraphs (<p>) in the element <support> or as one or more instances of the element <material>. It contains a brief statement that can serve as an overview of a more in-depth description that can follow. Here is an example of the first type:

```
<support>
  <p>Paper with <watermark>Harbalet in Ring</watermark>: similar to
<bibl>Harlfinger Nr. 2/a</bibl> from the years 1471-1472</p>
</support>
```

(From the electronic description of MS Rila 3/6  
by Dilyana Radoslavova, simplified)

However, in the case of an analytical description we may choose to use a paragraph as just an introduction to a more elaborate description of the material, or we may use only one or more <material> elements, often depending on the completeness of information we have. For example, if we are describing a manuscript that we have not been able to examine *de visu*, or are making an electronic edition of an old catalogue, we may lack information about the paper or quality of parchment. In such cases we may ignore the <support> section of description and provide the information we do have only as the value of the attribute @material in the element <supportDesc>, as in the example above.

In contrast with this, when we use one or more <material> elements not as part of <p>, that means that we have enough information to complete a parchment or paper description of the codex. Consider the following example:

```
<supportDesc material="mixed">
  <support>
    <material key="parchment">Here is the information on
parchment</material>
    <material key="paper">Here is the information on paper.
</material>
  </support>
...
</supportDesc>
```

Here the @key attribute is used to distinguish the different materials in the manuscript. The element <material> can include an attribute @material, with values

that might include (but are not limited to):

- *original*: original parchment used for the main body of the manuscript;
- *reused*: reused parchment (palimpsest) or paper;
- *supplement*: a supplement written for the original manuscript;
- *restoration*: added during the restoration process;
- *endleaf*: use for the end leaves;
- *pastedown*: leaf or leaves glued to the inside of a board;
- *binding*: part of the binding (but not the only part of it);
- *wrapper*: a wrapper for the manuscript;
- *distinct*: an originally distinct leaf or leaves (bound with the manuscript at a latter date);
- *other*: all other possible cases that could not be formalized clearly in advance.

### *Description of parchment*

The parchment description is restricted to the technological process and its outer appearance. However, the encoder might wish to add a statement regarding the origin, quality, colour and thickness of the parchment, as in this example:

```
<material key="parchment" usage="original">More probably sheep  
than calf...of middling thickness and quality, with a good mat (velvety) finish" (Doyle and  
Parkes 1979: xxi), trimmed and very stained. </material>  
<material key="parchment" usage="restoration">Portions eroded  
(gnawed by rats?) have been replaced with blank parchment (1956). </material>  
(http://www.canterburytalesproject.org/pubs/HGMSDesc.html)
```

### *Description of paper*

#### **General remarks**

The description of paper in recent catalogues is much more complicated than the description of parchment. On the one hand, there are various levels of detailed information in the catalogues, and, on the other, there are much more specialized research initiatives in the field of paper and watermark studies. Therefore, the description of the paper in a manuscript could consist of an overview and comments followed by a description of watermark(s) and of other aspects of the paper production:

```
<material usage="original">  
  <dim type="format">4°</dim>thick, rough, non-  
glazed  
  paper </material>
```

(From the electronic description of Add. Ms 37233  
by Dilyana Radoslavova, simplified)

The description of obviously different kind of paper, possibly from various manufactures, and of varying quality and age could be accomplished by repeating the ele-

ment `<material>` and its attributes exactly in the same manner as it is implemented for the description of parchment.

### Watermarks

The central and most important information in the description of the paper is the watermark. The information on its characteristics is part of the element `<watermark>`. The content of this element includes phrase level elements, as well as specialized elements like `<countermark>` and `<motif>`.

There are several special attributes for the content of this element. `@tracing` is used to describe very briefly (with one word) the method of watermark tracing. The most usual method is still *drawing* by hand, but there could be other possibilities like *rubbing* by hand, usage of special Dylux print paper, *beta-* or *x-radiography*, *scanning*, etc. Further more the `@place` from the class `att.placement` can refine where on the pages the watermark could be seen. This feature could be described in relation to wire lines and *chain (laid) lines*. The possible content might read something like: chain lines regularly spaced, between chain lines more widely spaced, on a chain line, on a supplemental chain line, across 2 chain lines, across 3 chain lines, or could be expressed in some other terms, e.g., in millimeters between various points of the leaf.

#### *Motif*

The element `<motif` is the most important part of the watermark description because it is intended to give a special terminological apparatus of watermark design.

There are several ways to add information on this subject:

- In the absence of reference books and watermark albums, simply as plain text;
- If a more complicated motif is present and the encoder would like to provide full compatibility with professional watermark databases, with the attribute `@type`, as described below;
- By *secondary* is meant designs above or below the basic motif. Frequently used motifs above the basic one include bird, cross, flower, monograms, monograms with a combination with figures (e.g., monogram BA and cross), ramrod, different star shapes, star and monogram, trefoil, etc. Frequently used motifs below the basic one include grapes, monograms, etc.;
- By *supplemental* is meant supplemental characteristics of the basic motif, e.g., in circle, in oval, in shield, etc.

In this approach the usual description of the motif “anchor in circle with 5-point star above” could be encoded as:

```
<motif type="basic">anchor<motif type="supplemental">in
```

```
circle</motif>
<motif type="secondary">with
star/5-point above</motif>
```

### Countermark

A countermark has the same content and attributes as the <watermark> element. Example:

```
<material>
  <watermark tracing="drawing"><motif
type="basic">Leo</motif>within <motif
type="supplemental">circle</motif> surmounted by <motif
type="secondary">quatrefoil</motif> and with
<countermark>"AP"</countermark> similar to <ref>Nikolaev
#125</ref> from the year <date cert="yes">1585</date>
  </watermark>
</material>
```

*A note inside* <material>

The meaning of this note is to enable the encoder to describe other features of the material, which is especially important for the description of paper. Although not structured in a systematic way, the information to be entered here might involve:

*Wire lines:* The relevant information of wire lines could contain:

- The density of the wire lines as the measure of 20 wirelines (in mm) — a standard procedure common in paper descriptions;
- Description of wire lines, usually in terms like: thin wire lines; thick wire lines; alternating thick and thin; very thick wire lines; thick wire lines, thin lines at edges; wire lines in corrugated pattern; wires wrapped around the chain lines; alternate wires wrapped around the chains; etc.

*Chain or laid lines:* The most important feature of the chain lines is the spacing between them, which could be described by entering the element <dim> and its sub element <height>. The following are recommendations from the Watermark Initiative: “If the watermark is on the chain line the spaces on either side of it should be given in parentheses: 28,25,(15,14),28,25. If the watermark is positioned between or across two chain lines, the space between them should be given in parentheses: 44,45,(50),47,43 “ ([http://abacus.bates.edu/Faculty/wmarchive/FORM\\_descriptions.html](http://abacus.bates.edu/Faculty/wmarchive/FORM_descriptions.html)). An example of such description is provided as well: “04. Wire/Chain line description: Density of Wirelines: 26 mm. for 20 wirelines. (style: thin wirelines) Chain line Interval(s): / 50 mm.”

*Paper thickness:* The data on paper thickness could be entered in the element <dim> with the @type attribute and a *thickness* value. This value could be relative, e.g.: very thin, thin, medium thin, medium, medium thick, thick, very thick; or, alternatively, it could be measured in another way.



*Paper colour*: This information could be entered as content of the element <colour>.

*Texture*: Terms in common use include: *smooth*, *matte*, *fuzzy*, *patina*, *glazed*, *bumpy*, etc. One might also record whether the dark fibres and wire lines are visible on the surface of the paper.

*Stiffness*: Information about the stiffness of paper is rare in manuscript descriptions, but possible values could include: *stiff*, *crisp*, *supple*, *limp*, etc.

*Date of the paper*: Usually the dating of the paper is part of the watermark description. This is the most familiar place because the watermark image is the most characteristic part of the paper description. In addition to providing the date or date range, the encoder could further specify the justification for the dating. Possible values could include:

1. published watermark(s) cited in references;
2. codex date given in scribal note;
3. codex dated on palaeographical grounds;
4. codex dated by comparison of papers;
5. codex dated from other evidence;
6. the paper associated with binding, etc.

The date range could be entered with the asterisk (\*) as a wildcard, so that, for example, the expression 15\*\* would mean the range from 1500 to 1599, 156\*, from 1560 to 1569, and 156\*-157\* - each year from the 1560 up to the 1579.

*References*: The references are given as external reference element inside the elements <bibl> or <listBibl>.

#### *Number of leaves or folia*

The number of folia in the manuscript is entered in the element <extent>, which may be accompanied by the element <measure> with the attribute @unit in order to say whether we mean leaves or pages. Consider the following example from the description of the manuscript 369 from the Church Historical and Archival Institute in Sofia:

```
<extent><measure unit="folia">4+516 (Stojanov wrongly gives  
73ff.)</measure><extent>
```

It is possible to include links to the information in the same file or outside it, e.g.:

```
<extent>2 <ref target="OENB146">(Vienna)
```

```
</ref>+171 (f.134 is a later insertion)
```

```
<ref target="RGB1689">(Moscow) </ref></extent>
```

This shows the location of the folia in both repositories. The *Codex Marianus* consists of 2 ff. kept in Vienna and 171 in Moscow. The element <ref> adds a link to the corresponding part of the description (<msIdentifier> or <altIdentifier>).

An even more complicated example can be illustrated from the description of the Manuilev Apostle, which is now separated into six pieces housed in three different libraries:

```
<extent unit="folia">lacuna+1 <ref target="NBKM499">(f. 3, NBKM
499)</ref> + 1 <ref target="BAN4.5.23">(BAN 4.5.23)</ref> +
    9 lacuna + f. 4 + lacuna + f. 5 + lacuna + ff. 6-10 + lacuna + f.11-12<ref
target="NBKM499">(NBKM 499)</ref> +
    lacuna + 1
        <ref target="RSLGrig15.III">(Grig. 15.III, 1r/v)</ref> + 4
<ref target="BAN24.4.7">(BAN 24.4.7)</ref> + 1
    <ref target="RSLGrig15.III">(Grig. 15.III, 2r/v)</ref>
+lacuna + 1 <ref target="NBKM500">(NBKM 500)</ref>+2
    <ref target="NBKM499">(ff. 1-2, NBKM 499)</ref> + 1 <ref
target="BAN4.5.24">(BAN 4.5.24)</ref></extent>
```

## Dimensions

### General remarks

All sorts of dimensions could be entered into the element `<dimensions>`, and the attribute `@type` gives a hint to the encoder of possible uses for the element. The attribute `@unit` provides three possible characteristics and `@scope` emphasizes the range of these dimensions.

The element `<dimensions>` could be used in many places where measurement is requirement, and it has three special attributes for that purpose. The attribute `@type` provides a common typology of the dimensions of a given manuscript source. Possible values comprise all important parts of the manuscript physical description and layout. Commonly used dimensions are those of *leaves*, *page written* area and the *height* of the characters or lines. The author of the description may, however, find it useful to mention also the dimensions of the *ruled* or *pricked* region of the page and quire, as well as to provide measurements for the *borders* as part of the binding. It is even possible to distinguish the dimension of a single letter in comparison to the whole line.

The most useful units of measurement in manuscript description are mm and cm, but sometimes, especially in the description of initials, a measurement in *lines* per page or characters per line (*chars*) could be more useful than exact measures.

To supply information on the possible range of dimensions the compiler of the catalogue information may find it convenient to use the attribute `@scope` with the exact value of a particular part of the manuscript.

### Height

The height of some object given in units as part of the element

`<dimensions>`.

### Width

The width of an object given in units as part of the element `<dimensions>`.

### Depth

The depth of an object given in units as part of the element `<dimensions>`.

A possible example could be:

```
<dimensions unit="mm" type="folia">  
  <height>205</height><width>140</width>  
</dimensions>
```

### *Other dimensions*

Any other dimensions can be supplied through the general-purpose element <dim>. For example, in measuring binding panels:

```
<dimensions type="panels" unit="mm">  
  <height scope="all">300</height><width  
scope="all">220</width><dim type="thickness">85</dim>  
</dimensions>
```

### *Foliation and pagination*

Both elements have the same content and structure, and include information about when the foliation or pagination was made, whether it is original, and how the folia or pages are marked or numbered, e.g.:

```
<foliation>original, from <date>1805</date></foliation>
```

From the electronic description of MS Pogodin 1204  
by Antoaneta Granberg

### *Collation*

The description of collation offers three alternatives: 1) to describe the general collation using the element <p> for an unstructured description; 2) to begin with a general overview using the element <p> but also to describe such characteristics as signatures, catchwords, pricking, ruling or dispositions of folia in the gatherings in separate elements; or 3) when some of the quires deviate from the general rules, to describe them in one or more separate <quire> elements.

A simple description could include the element <p> with a formula, e.g.:

```
<collation><p><formula>I (8 -i), II(8), III(8-i), IV-XIX(8), XX(4.4+2), XXI(8  
-8)</formula></p>  
<signatures>Gatherings signed in the <locus>mid-lower margin, first  
recto</locus>  
<ruling>Leaves have been ruled, though the ruling pattern is hardly visible.  
</ruling></signatures>  
</collation>
```

### *Signatures*

An important part of the description of quires is how they are marked at their beginning and/or end. The description should include a phrase defining the form and place at the beginning and end of the quire, as well as any changes, e.g.:

```
<signatures>The original markings of the quires through letters is preserved  
</signatures>
```

### *Description of each quire*

The element `<quire>` brings together all possible features for the characterization of the quire, including number, composition, how and where there are marked, and with the assistance of what tool.

The element `<formula>` provides information about how the quire is composed, of how many folia, etc. The content of this element is a formula. Consider the following example:

```
<collation>
<quire n="I"><extent>2</extent><formula>8 (-1-2-3-4-5-6)
</formula></quire>
<quire n="II-XXI"><extent>8</extent></quire>
<quire n="XXII"><extent>7</extent><formula>7 (-7)</formula>
</quire>
<quire n="XXIII"><extent>2</extent><formula>6 (-1-2-3-4)
</formula></quire>
...
</collation>
```

This means that first quire of the manuscript was formed of 8 ff., all of which, with the exception of the f. 7, are now missing. Quires II to XXI have 8 ff. each. From quire XXII one folio (7) is missing. The last, quire XXIII, consisted originally of 6 ff. but only 2 of them (5 and 6) are preserved intact.

### *Gregory rule or the rule of parchment arrangement in the quire*

The arrangement of the parchment folia as part of the quire structure is entered here. Only *de visu* observations can identify the sequence and order of parchment folia. There are two common abbreviations: HS – for the hair (dark) side, and FS – for the flesh (light) side of the parchment.

Therefore, the content of this element could be a sequence of HS and FS abbreviations, e.g.:

```
<gregoryRule>HS|FS-FS|HS-HS|FS-FS|HS =
HS|FS-FS|HS-HS|FS-
FS|HS</gregoryRule>
```

Here the sign `|` separates pages, `-` separates folia, and `=` signifies the middle of the quire.

In cases where there is no rule, or there are deviations from it, this should be stated explicitly. If the deviations are caused by the loss of leaves, it is possible to insert a link pointing to the place where this loss is explained, e.g., in the content of the element `<condition>`.

### *Ruling patterns and types*

Two kinds of information regarding medieval manuscripts could be entered in the element `<ruling>`. The first is related to the ruling type defined by researchers like Leroy (1976) or Gilissen (1977), who attempted to propose well-defined typologies of ruling types. In the other, information is related to the place of ruling in the

quire, or the so called „scheme.” The former could be expressed as a formula of ruled lines for writing, and the latter could be visualized as simple graphic showing the places in the quire where the ruling is accomplished, e.g.:

>> | > | |< |<< = >>| >| |< | <<

This scheme means that in this quire the ruling is made for each of the two folia: once for the outer double leaf and once for the inner one (double sign >> or <<). The symbols > or < mean that on the leaf there is just a trace of the ruling made on the previous leaf<sup>8</sup>.

The attribute @type should provide information about the classification scheme (Gillisen or Leroy for example) employed.

### *Prickings*

Such observations are relatively rare in the literature. However, as described by Jones (1946a, 1946b), they could bear information on scriptorial techniques. Thus, it is necessary to provide an element for such types of information.

<pricking>Pricking in outer margins only, probably through folded quires, from front to back.</pricking>

(Example from DUL MS Cosin V.II.17 Anseis de Carthage, Ogiez le Danois (in French) s. xiii2 description  
(<http://www.dur.ac.uk/library/asc/theme/medmss/apvii17/>)  
encoded in Repertorium model)

<pricking>Most prick marks have been trimmed off. On p 112, 6 prick marks are visible on the outer margin. On p 417, thirteen prick marks at 9mm intervals remain on the outer margin. There is no pricking on the inner margins.</pricking>

(Example from the web-based electronic description of St. Alban's Psalter  
<http://www.abdn.ac.uk/~lib399/english/essays/codicology.shtml>) encoded  
in Repertorium model)

### *The manuscript condition*

Information about the condition of the manuscript is entered in the element <condition>. It is used mainly to describe damage to folia, quires, the binding or the whole manuscript. Traces of restoration processes could be encoded in this element as well. A possible example could be:

<condition>The manuscript is in good condition except for the last gathering, in which there is damage along the lower edge of the leaves, which have been repaired and the text restored in a style very close to the original.</condition>

From the electronic description of British Library Add MSS 23999  
by Dilyana Radoslavova

---

<sup>8</sup> The symbols < and > are reserved characters in XML. In the encoded file they should be written as &lt; for < and &gt; for >, respectively.

### *Description of the page layout*

This description is entered as part of the element `<layoutDesc>` (Burnard & Bauman 2007). As in many other places in the model you can choose between unstructured description (with element `<p>`) or structured data in the element `<layout>`, probably preceded by an introduction (element `<summary>`).

The element `<layout>` has three special attributes: `@columns`, `@writtenLines` and `@ruledLines`. If there are no further explanations by the encoder, all the information can be supplied through these attribute values:

```
<layoutDesc>
<layout columns="1" writtenLines="18" ruledLines="20"/>
</layoutDesc>
```

### *Ink*

Information about the ink is presented usually as a simple phrase like “brown dark ink for the basic text and red ink for the headings and small initials”. However, such descriptions could then be further developed in order to supply information on ink composition, or its brightness or its loss of colour. A simple description could be:

```
<ink>Brown ink for the text and dark red for the initials</ink>
```

### *Decoration*

The element `<decoDesc>` is used to give information about manuscript illumination. There are two basic ways of using this element and the subelements of its structure. The first is to provide basic notions on the styles and art objects in the codex with the help of the element `<p>` (paragraph). Another other way is to provide a thorough description of each kind of manuscript decoration with the element `<decoNote>`, possibly preceded by some kind of introduction (in the element `<summary>`). In the second approach each `<decoNote>` is supplied with the attribute `@type` and an appropriate value, e.g., miniatures, headpieces, tail-pieces, initials, borders, line-fillers, marginal, wash, etc. Here we will give some examples:

1. For miniatures:

```
<decoDesc>
<decoNote type="miniatures">
  <p>Three miniatures of St. Mark, St. Luke, and
  St. John, crude drawings in <colour>ink</colour> and <colour>pink</colour>,
  <colour>green</colour> or <colour>ocher</colour>
  <term>washes</term>. </p>
  <list><item xml:id="StMark"><locus>F. 44r</locus>, St. Mark
  in full length, under arch, holding the Gospel.</item>
  <item xml:id="StLuke"><locus>F. 77v</locus>, St. Luke in full
  length, under ornamented arch.</item>
  <item xml:id="StJohn"><locus>F. 133v</locus>, St. John in full
  length, under arch. The image is stylistically different from the others.</item></list>
```

```
</decoNote>
```

```
</decoDesc>
```

## 2. For headpieces:

```
<decoDesc>
```

```
<decoNote type="headpieces">
```

```
<list>
```

```
<item><locus> f. 43, the Chapters of St. Mark's Gospel, </locus>  
an interlaced headband with floral endings at the corners and the short sides  
of the band.</item>
```

```
<item><locus>f. 44v, the beginning of St. Mark's Gospel</locus>,  
an interlaced headband with floral endings at the corners and a central motif on the upper  
rim.</item>
```

```
</list></decoNote></decoDesc>
```

## 3. For initials:

```
<decoDesc>
```

```
<decoNote type="initials"><label>Floral elements and  
interlaces</label>
```

```
<list><item><locus>f. 44v, beginning of St. Mark's Gospel  
</locus>initial <c xml:lang="cu" rend="glag">3 </c> with floral  
elements and interlaces. </item>
```

```
<item><locus>f.78r, beginning of St. Luke's Gospel </locus>initial  
<c xml:lang="cu" rend="glag">Π</c> with floral elements and interlaces.
```

```
<ref target=" ../images/CodexMarianus78r.jpg">  
Image</ref>
```

```
</item>
```

```
</list>
```

```
</decoNote>
```

```
</decoDesc>
```

## 4. For decorations in margins

```
<decoDesc>
```

```
<decoNote type="marginals">Marginal pointers of floral motifs on f. 1v, 5v,  
39r, 55v, 71v, 88r, 89r, 102r, 111v, 120r, 131r, 141r, 143r, 163r, 179v, 198v, 213, 229r;  
245v; 257v, 269v, 281r; 279r, 296r, 297r-v, 298v, 300r-v, 301r-v, 302v, 303r, 304r, 305v,  
311r-v, 312r-v, 313r, 314v; 274r - picture of the Holy Cross; </decoNote>
```

```
</decoDesc>
```

In the last example the various appearances of floral motifs could be further wrapped within the element `<locus>`. As can be seen in the examples above, multiple occurrences of one decoration type could be encoded as a list of items, each with a separate `@xml:id` attribute and a pointer to some image outside the file.



## *Binding*

### *General remarks*

Information about binding in most of the descriptions is very brief, and consists of statements like “Wooden panels covered with ornamented leathers. Not original. Done probably in 16th century”. In most cases this information is entered in brief descriptions using the sub-element `<p>` of the element `<bindingDesc>` (Burnard & Bauman 2007). The describer can supplement this with further notes about the binding decoration (`<decoDesc>`) and its recent condition (`<condition>`).

If the encoder prefers to use a structured description, a short statement should be entered in the element `<binding>` as value of the attribute `@contemporary`, such as *true*, *false*, *unknown*.

The encoder can include a short statement before continuing the description of the binding, or or she can provide only this information. In the latter case he or she could give information about all of the main parts of the book-binding structure and decoration. For example:

```
<bindingDesc>
<binding contemporary="true"><p>Brocade on wooden boards,
<dimensions unit="mm"
type="plate"><width>145</width><height>205</height></dimensions>.
Two clasps. The paste-downs and first and last flyleaves are backed with mauve cloth.
Illumination and marginalia have suffered a little from excessive trimming. The edges of the
pages are gauffed and gilded. Two metal bosses on the lower side of each cover.</p>
<condition>Good preserved.</condition>
</binding>
</bindingDesc>
```

(from the description of British library Ms Egerton 3045  
by Dilyana Radoslavova)

All other aspects of the binding description, not covered by elements like `<decoDesc>` or `<condition>`, could be entered in the element `<bindingNote>` with different values of the attributes `@type` or `@subtype`. For example, all aspects of the description of the cover should be entered in the content of the element `<bindingNote>` with attribute `type="cover"`. The material of the cover should be then described as part of the sub-element `<material>`. The same is true also for other parts of bookbinding, such as edges (e.g., technology, beveling of edges), fly-leaves, etc. The observations in `<fastening>` could include how the fastening is done, e.g., by peg and strap, by hook and clasp; the length or width of the elements could be measured; their place(s) could be described, etc. In `<furnishing>` the encoder should describe all the bosses, corner pieces, etc. that can be observed on the binding.

A description of the sewing could include the types of attachments to the quires

and leaves or to the main boards. A text block can be sewn through the side of the leaves, e.g., through as over sewing or side sewing, etc. In addition, the text block can also be sewn through the fold of the sections by hand. Three different patterns are observed: sewing structure, sewing stitches, and sewing pattern.<sup>9</sup>

*Sewing structures*: supported = parts other than thread involved in the connection of the sections. Materials for support usually are: cord, parchment, skin, linen tape, etc. *unsupported* = only sewing thread connecting the sections to each other. An unsupported structure is commonly found in Greek, Islamic, and Oriental books.

*Sewing stitches*, which is the motion of a thread from a sewing station (hole) exit to a sewing station entry. The classification could be as follows:

- *Slip*: thread passes under itself
- *Link*: thread passes under another thread
- *Loop*: circle around a support
- *Lap*: passes over a support
- *Forward*: continues in the direction of progression
- *Reverse*: continues to the previous direction of progression
- *Drop*: moves downward
- *Climb*: moves upward
- *Incline*: moves diagonally

*Sewing patterns*. The main patterns are:

• *Fold patterns*: sequence of thread visible inside the fold of the section. These are of two types: Continuous pattern: thread passing in and out each sewing station to form a pattern; and Periodic pattern: thread skipping sewing holes and leaves intervals between some of the stations.

• *Spine patterns*: the pattern is formed by the stitches across the uncovered spine of the book. Some spine patterns are: (1) Chain (formed by link stitches), (2) Bridge (by span), (3) Ribs (by loop), (4) Herringbone (by loop and link), (5) Steps (by lap), etc.

Interesting questions pertaining to the description of spines include:

- whether the form of the spine is rounded or flat (as an attribute @type);
- whether there are raised or smooth bands;
- the process of reinforcing the spine of a sewn book after gluing-up, rounding and backing, and before covering or casing-in, known as lining or spine lining;
- whether there are headbands, whether they are double or triple, etc.; how they are connected to the board; what their colours are; what material they're made from.

### *Date of binding and place where it is made*

In most of the cases the binding of medieval manuscripts is latter then the

---

<sup>9</sup> The following definitions and information are based on Pamela Spitzmueller's paper – "Trial Terminology for Describing Sewing Through the Fold," (the Paper Conservator, vol.7, 1982/83, and on the web-article "Paper Conservation Course" (<http://www.scmre.org/bookbinding.html>) and on Szirmai's book (1999).

codex itself. Therefore the date and place of producing the binding should be provided separately from the conclusions on date and place of origin of the described book.

```
<date certainty="likely">15 century</date>
```

Sometimes the exact date is known from a note by the copyist or bookbinder. In these cases a reference to the description and its edition should be made, e.g.:

```
<date certainty="yes">1667<ref target="note1">cf. bookbinder's  
note</ref></date>
```

(from Tachiaos 1981: 69–70, No 25)

In this case the `<date>` element bears the attribute `@certainty` with the value *yes* and this information is further referenced with the `@target` attribute value *note1*. As in all other cases with the elements `<ref>` or `<ptr>`, such an identifier must be present. Another possible solution is to use in this case the element `<ptr>` instead of `<ref>`. The element `<ptr>` is an empty element and all its information is encoded with attributes. This means that we would have to omit, e.g., a bookbinder's note, yielding something like:

```
<date certainty="yes">1667<ptr target="note1"/></date>
```

### *Description of seals*

The description of seals attached to the manuscript follow the same model as mentioned in the description of binding. The catalogue's compiler could make a short statement in form of paragraph or paragraphs (`<p>`) or provide a structured description in the form of decoration description (`<decoNote>`), the condition of the seal(s) (`<condition>`) and the seal itself (`<seal>`). The element `<seal>` has the same special attribute `@contemporary` with the same choice of values as `@binding`.

### *Description and edition of various notes*

#### *General remarks*

There are several types of additional notes, some original to the manuscript, that could be described in different ways. The first one, the colophon, containing the name of the scribe, date and possibly place should be described as part of the `<msItemStruct>` as stated below. Other notes go elsewhere. Some examples follow.

#### *Description and edition of later note as part of the <additions> element*

This is the most obvious place for the description of additions. The element `<additions>` is the last part of the structure of `<physDesc>` element. If there is just one addition, then the description and edition should be ordered in the following order: locus, name, date. The description could consist of one or more paragraphs or it could be encoded as a list containing two or more items. If there is just one note, the

structure could be as follows:

```
<additions type="inscription">
  <p><locus>the place in the manuscript, e.g., f. 104r, inner
margin</locus>
  <persName>if available</persName>
  <date>if available</date>
  <quote>an edition of the original text</quote>
  <note>other notes and/or bibliographic reference, e.g., previous
editions and interpretations.</note>
  </p>
</additions>
```

In cases with more than one note, it might be most convenient to represent them as a list of separate items:

```
<additions>
  <list><item><locus>front cover</locus>
  <persName type="binder">Krastio Gramatik</persName>
  <date evidence="conjecture">1785</date>
  <quote xml:lang="cu">сѣа дамаскинъ на сѣаго архангела повеза
го кр(с)то граматикъ и тѣгава бѣха два попа на сѣаго архангела попъ Бено и ѡзѡнъ попъ
кр(с)то догава отиде на сѣа гора мѣъ маѣа</quote></item>
  ...
  <item><locus>f. 61v</locus>
  <persName type="reader">Stefan Stojanovich
</persName>
  <date evidence="internal">1834</date>
  <quote xml:lang="cu">Извѣстную какво на време то на Дѡнисиѡ
епѣскопа ловчанскагѡ собираха ѿ патраршина та ѿ секи попъ по 43 гро(ш). заради да права(т)
сѣо мѣро. наже сѣпени(к) помажет новорожденныа мл(д)нцы съ сегѡ сѣаго мѣро бывате сыни
бжїа, и наслѣбники цр(с)твїя его. 1834 мѣрд. 23 гѡв дна</quote></item></list>
</additions>
```

The elements in both approaches should contain information about the place in the manuscript, the name of the scribe and the date if they are known, as well as the edition of note itself. The encoder could further provide a brief statement about the marginal formula and some typologization about colophons (cf. Gagova 1998; 2000).

#### *Additional notes and texts on separate manuscript folia*

When there is a later note or text written on a separate folio or folia it would be more convenient to represent it as a separate part of the manuscript. In this case we can make a simple reference in the element

```
<additions>:
<additions>
```

```
<p>A separate folio 72 inserted during the early book re-binding in 16 c. now forms
```

a separate `<ref target="MS_Part2">part</ref>` of the manuscript.

```
</p>
```

```
</additions>
```

Then this part is fully described elsewhere in the document as part of the element `<msPart>`.

## *The Scribe: palaeography and features of the language*

### *General remarks*

In the Repertorium Initiative approach all data about the palaeographic and linguistic aspects of the manuscripts are gathered together under the element `<scribeDesc>`. This element represents an information level between the pure physical description of the manuscript and a description of its content. As elsewhere in this electronic model, both unstructured and structured approaches are available. The unstructured description means entering the information only as part of paragraph (`<p>`) elements. The structured approach involves encoding information about each scribe (element `<scribe>`), possibly preceded by an overview (element `<summary>`). When the second approach is adopted, the first elements to be entered are the name and the place in the manuscript where the scribe has written it. If the name of the scribe is not available or unknown, the encoder can record that as:

```
<scribeDesc>
<scribe n="1">
<name>Unknown</name>
```

```
...
```

```
</scribe>
```

```
</scribeDesc>
```

This information is followed by a description of palaeographic and linguistic features, along with additional notes as needed.

### *Palaeographical features*

The analysis of the palaeographical features of the manuscript should be accomplished in the element `<palaeoCharact>`. The content of this element is a combination of paragraph elements with others specific for the description of palaeography. This means that the encoder may enter either unstructured or structured information, and under the latter approach such features as musical notation or encrypted writing can be described with specialized tags. The `<palaeoCharact>` element has one important attribute – `@script`. The value of this attribute should be short code that is compatible with international standards of script designations (ISO 15924), e.g., *Grk* for Greek, *Cyrl* for Modern Cyrillic, *Cyrs* for Old Cyrillic, *Latn* for Latin or Glag for Glagolitic (<http://unicode.org/iso15924/iso15924-codes.html>):

```
<palaeoCharact script="Glag"> ... </palaeoCharact>
```

A brief introduction to the palaeographical characteristics of the scribe can be included in the element <summary>, e.g.:

```
<summary>Uncial script very close to the writing in  
Enina Apostle and Vatikana Palimpsest</summary>
```

*The description of the shape of each sign (<palaeoNote>)*

This optional element may be useful when the exact shape of a letter or other sign is important, as in:

```
<palaeoNote>The letter <c xml:lang="cu">a</c> is written ...  
</palaeoNote>
```

```
<palaeoNote>The letter <c xml:lang="cu">б</c> is written ...  
</palaeoNote>
```

*Musical notation*

If musical signs are found in the manuscript, the system of their usage along with the places in the manuscript where they are found are described in the element <musicNotation>. The element has a special attribute @type, which could further systematize the notation:

```
<musicNotation type="theta">The theta-notation is used together with  
double okseia, double vareia (Piesma) and double apostrophos.</musicNotation>
```

*Cryptography*

The element for the description of a cryptographic system in the Repertorium model is called <crypto> and should be used in cases where encrypted letters or passages of text occur. The additional attribute @type might identify the method, e.g., simple substitution cipher:

```
<crypto type="simpleSubstitution">A = a, B = H, Γ = 3, Δ = S, K = Π, Λ = o, M  
= ž, ρ = u, c = w, T = ψ, γ = x. The other letters are used without  
substitution.</crypto>
```

*Description of the scribe's language*

*General remarks*

In the framework of Repertorium Initiative there is a clear distinction between the language of the *text* and the language of the *scribe*. The language of the text is understood in relation to the text history, cf. the section pertaining to manuscript content. The language of the scribe or scribes may or may not be related with the text history and text analysis. Therefore, a special section in the manuscript description is need to allow it to be treated separately.

The language of the scribe can be described as unstructured paragraphs or as a sequence of the elements <orthography>, <grammar> and <lexis>, optionally introduced by general observations in the element <summary>.

*Orthography*

Encoders of Slavonic manuscripts have two possible ways to describe orthographic features. The first is to provide only a brief characteristic in the element p, e.g.:

```
<orthography>
```

```
<p>The scribe writes with two jers and with nonetymological appearance of juses. After the letters ш, ж, in the word-initial position (after j), and after the vowel letters front nasal is turned into the back one. After ч this change is not observed. A relatively small number of ligatures and supralinear signs. A few examples of accent marks ...
```

```
</p></orthography>
```

The other way to describe the orthography of the scribe is with a structured description, where each piece of the description is entered as a separate element.

The element `<orthNote>`

The rules for the usage of different letters and signs are described as content of the element `<orthNote>`. It contains the attributes `@type` and `@subtype`, which are used to typologize the orthography of the scribe in appropriate ways, especially with respect to three groups of letters: jers, juses and letters for jotated vowels.

#### *The orthography of the jers*

Information regarding this feature is entered in the element `<orthNote>` with value of attribute `@type="jer"`. The encoder can describe the positions of jers and their relationship to the etymological condition, along with examples of their usage.

This element has one required attribute: `@subtype`. Possible values are:

- *front*: only front jer is used
- *back*: only back jer is used
- *etymReg*: etymologically regular use of jers
- *nonEtymReg*: not etymological but regular use of jers
- *irregular*: the use of jers cannot be formalized in general rules.

For example, manuscripts written with Tarnovo orthographic norms use both jers regularly, but in a non-etymological way. The same is true for the Zograph Folia from the tenth century.

```
<orthography>
```

```
<orthNote type="jer" subtype="nonEtymReg">Both jers are used but jery is written only with front jer. There is a tendency to substitute back with front jer before a syllable with front vowel.</orthNote>
```

```
</orthography>
```

#### *The orthography of the juses*

The description of the orthography of the juses is very similar to that of jers, addressing the positions where juses are or are not used, etymologically correct or incorrect usage, etc. The element `<orthNote>` is used here, too, but with `type="jus"`. Here, the possible values of `@subtype` attribute are:

- *etymReg*: etymologically regular use of juses (e.g., as in Old Bulgarian)
- *nonEtymReg*: non etymological, but otherwise regular use of the juses (Middle Bulgarian manuscripts, and in part East Slavic examples)
- *nonConsis*: nonconsistent use of the juses
- *nonJus*: orthography without juses
- *jusTrace*: orthography without or with very few examples of the juses, but with trac-

es, e.g., manuscripts written generally with non-jus orthography but with traces of the Middle Bulgarian alternation of juses

```
<orthography>
<orthNote type="jus" subtype="jusTrace">The text is written with non-jus
orthography but with some traces of Middle Bulgarian confusion of
nasals.</orthNote>
</orthography>
```

#### *The orthography of the letters for jotated vowels*

In many cases the distribution of jers and juses may provide enough data to determine the main features of a specific medieval Slavic orthographic system. To be more precise, however, the description of letters for jotated vowels may also be important. This information is also as part with the element <orthNote> with the value of type="jotVowel".

```
<orthography>
<orthNote type="jotVowel">The letters <c xml:id="cu">e</c> and
<c xml:id="cu">ie</c> are both used in intervocalic position.</orthNote>
</orthography>
```

#### *The orthography of other letters or signs*

The distribution of all other letters and special signs with the exception of musical notation and cryptographic systems follow the model as described here for jotated vowels. The same model could be used for the description of non-Slavic orthographic features, as well.

#### *Description of grammar*

The description of grammatical characteristics could be divided into separate units, but in this case, unlike with orthography, there may not be a standard list of descriptive criteria, and the researcher has to introduce the structure using the available means. Consider the following example:

```
<grammar>
<list>
<item xml:id="phonology">Adjective forms prefer the archaic <m xml:
lang="cu">-ц-</m> to <m xml:lang="cu">-ст-</m>. Cases with
<m xml:lang="cu">-ст-</m> are rare. Uncontracted forms of pronominal
adjectives and of verbs of the -je-conjugation in Ind. Praes. Act. and of verbs in
Imperfect are much more frequent than contracted ones. Verbal forms(Ind. Praes. Act.,
p.3, -je-conjugation) with assimilation, such as <w xml:lang="cu">бѣглат</
w>, <w xml:lang="cu">рѣдоуотъ</w> are frequent. In Imperativus there
is variation of forms with <c xml:lang="cu">ѣ</c> and with <c xml:
lang="cu">и</c>, such as <w xml:lang="cu">покажѣте</w> 290.17, but
also <w lang="cu">покажите</w>
80.5.</item>
<item xml:id="verbs">The manuscript uses non-sigmatic
Aoristus (e.g., <w xml:lang="os">вѣнидъ</w>, <w xml:
lang="cu">сѣнидъ </w>, <w xml:lang="cu">идомъ</w>, <w
```



`xml:lang="cu">отидѧ</w>)` and first sigmatic Aoristus of verbs with stems ending in a consonant (e.g., `<w xml:lang="cu">привѣс</w>`, `<w xml:lang="cu">вѣвѣсом</w>`, `<w xml:lang="cu">вѣзньса</w>`). Verbs with infinitive stems ending in a vowel often preserve the archaic ending `<m xml:lang="cu">-тъ</m>` in Ind. Aor. Act, p. 2 and 3, Sg., e.g., `<w xml:lang="os">обитѣ</w>`, `<w xml:lang="cu">оумрът</w>`. Three cases of omitted ending `<m xml:lang="cu">-тъ</m>` for Ind. Praes. Act., p. 3. Sg. are found (`<w xml:lang="os">е</w>` instead of `<w xml:lang="cu">естѣ</w>`, `<w xml:lang="cu">нѣ</w>` for `<w xml:lang="cu">нѣст</w>`, and `<w xml:lang="cu">бѣд</w>` for `<w xml:lang="cu">бѣдет</w>`).  
</item>  
<item xml:id="nomina">The old suffix `<m xml:lang="cu">-нѣ</m>` is written frequently as `<m xml:lang="cu">-нѣ</m>`, e.g., `<w xml:lang="cu">знамение</w>`, but this orthography is not consistently followed. In other forms `<m xml:lang="cu">-ѣ</m>` is preferred (cf. `<w xml:lang="cu">дѣѣ</w>`, 24r.20).  
</item>  
</list>  
</grammar>

### *Description of lexical features*

Lexical features can be described similarly to grammatical ones. Whether this takes the form of a list with items or of just one unit depends on the encoder's vision and the length of the material:

<lexis>The manuscript uses lexis that is believed to be typical of the earliest Gospel translation from Greek, cf. the use of the archaic lexemes `<foreign xml:lang="cu">баль, жаль, миса, шюи, братръ, небесьскъ, неприѣзнь, пропаги, иночадъ, врьтъ, искрънъ, грѣсти</foreign>` could prove this hypothesis. Also, in word-formation the manuscript prefers the old suffix `<m xml:lang="cu" type="suffix">ице ѧ</m>` and not `<m xml:lang="cu" type="suffix">ждь</m>`. Still, the following forms are considered to be relatively later: `<foreign xml:lang="cu">благодать, написание, драхлъ</foreign>`.  
</lexis>

Both examples from the electronic description of *Codex Marianus* by Margaret Dimitrova, Andrej Bojadžiev and Elisaveta Mussakova

### *The manuscript contents*

The description of the manuscript contents in the Repertorium electronic description model is distributed over two main parts, one pertaining to the manuscript as a single body of texts and the other to each text individually.

### *The content of the whole manuscript*

The content of the whole manuscript is described in the element `<msContents>`. Nothing in this element is mandatory, but the element itself is required, and the encoder might start by defining the range of texts in terms of the research tradition. There are three special attributes for this element. The first, `@type`, can record whether the manuscript collection of texts is original (*original*) or represents either an earlier *compilation* or a *translation*. The second, `@style`, can situate the manuscript within a broad typology, viz., choosing between the values *narrative* and *non-narrative*. The third, `@class`, can be used to provide a convenient typology for all the texts in the manuscript. The usage of these three attributes with their values can be illustrated with this example:

```
<msContents class="vitas" type="translation"
  style="narrative">
...
</msContents>
```

Here the basic information about all the texts in the manuscript is given. In this case, the main classification of the texts falls under the category of *Vitas of saints* that are translations and that are narrative.

### *Several levels of the representation of whole manuscript contents*

#### *Unstructured view*

If we are compiling just a short description of the manuscript without information about each text, the most convenient method is to use just a paragraph with short information:

```
<msContents class="NT.Gospels" type="translation"
  style="narrative">
<p>The Aprakos Evangelion is without beginning and end. It starts with the reading for the Saturday of the 3rd week after Lent (Mk 8:30-31), continues with the Sunday after the 4th week of the Lent (Mk 9:17), and then the reading for the Saturday of the 5th week after Lent follows. After that there are texts for Palm Sunday and for the Monday and Tuesday of Holy Week.</p>
</msContents>
```

From the electronic description of MS Zogr. 282  
by Andrej Bojadžiev

#### *Structured view*

For more elaborate cataloguing purposes it is necessary to use special elements to describe the textual characteristics of the manuscript content in full. These elements are:

`<summary>`: a short characteristic of all of the texts  
`<titlePage>`: the title page of the manuscript if it exists

<title>: the title of the manuscript set of texts as given by the research tradition  
<author>: the author or authors of the whole text collection  
<respStmt>: responsibility statement – used for information about the editors or translators of the collection  
<note>: notes about the manuscript’s texts: e.g., missing parts of the collection or layout of the whole text collection  
<filiation>: text historical data; place of this collection in the textual traditions, e.g., with respect to protograph, antigraph, apograph or various sources of text compilation and/or translation  
<textLang>: information about the language of the text collection in relation of its textual tradition (not the language of the scribe or copyist)  
<sampleText>: a snippet of original text in the language and script of the manuscript; typically includes the main manuscript heading, rubric, incipit and explicit of the whole manuscript  
<msItemStruct>: description of each text in the manuscript

*A summary of the texts*

A summary should contain a brief phrase characterizing in some way the collection of texts in the manuscript.

```
<summary>Contains DICTIONARIVM SLAVONICOLATINUM OPERI  
AMBROSII CALEPINI Servata Verborum integra. SERIE CONFORMATUM  
MOSCHOVIAE. Anno 1695.</summary>
```

From the electronic description of BL Add. MSS 12069  
by Dilyana Radoslavova

This summary could also contain data about the quantity of the text in the manuscript, e.g.:

```
<summary>The MS contains 34 texts.</summary>
```

*The manuscript title page*

This is a very rare tradition in medieval manuscripts, and is more common in early modern books. Where it occurs, it is best to encode just the main information and to give a link to the image of the page:

```
<titlePage>  
  <docTitle>  
    <titlePart>The title as it appears on the title  
page</titlePart>  
  </docTitle>  
  <figure>  
    <figDesc>A description of the image</figDesc>  
    <graphic url="sometitlePage.png">A link to the  
image</graphic>  
  </figure>  
</titlePage>
```

*The authors, editors and compilers of the text collection  
in the manuscript*

In most cases manuscripts contain texts from different authors and their tradition of edition and compilation could be very variable. There are, however, also codices that contain more or less a fixed collection of texts by or attributed to a single author. In such cases, such as the *Pandects of Antiochus*, *Pandects of Nikon of Montenegro* or *Chrysorrhoeas* collection with writings of John Chrysostom (Thomson 1982; Miltenov 2010; Милтенов 2011) we can provide information about the author and sometimes about the editor/compiler of the whole set of texts, as in this example:

```
<author>John Chrysostom</author>
<respStmt>
  <resp>Τρανσλατορ</resp>
  <name>monk Antonij</name>
</respStmt>
<respStmt>
  <resp>Compiler</resp>
  <name>Vladislav Grammatik</name>
</respStmt>
```

From the electronic description of MS Rila 3/6  
by Anissava Miltenova

Here we have three kinds of information: 1) about the original author of the texts (John Chrysostom), 2) about the translator of his works (monk Antonij), and 3) about the compiler of this miscellany (Vladislav Gramatik). All the information other than authorship is contained in the element `<respStmt>`, which has two parts: `<resp>` (responsibility), here a brief statement in the form of words as *editor*, *compiler*, *scribe* etc., could be made and information about the person in elements `<name>` or `<persName>` (personal name, cf. Burnard & Bauman 2007: `respStmt`).

*A note in the description of texts*

The element `<note>` in the description of texts is restricted to cases that deal with extra-textual historical features, such as missing parts of the texts:

```
<note>Most of the texts have some missing parts. Cf.
Manuscript Condition </note>
```

Here, a reference to the physical condition of the manuscript is given. In other cases the element `<note>` could be useful to further refine the text attribution of the whole codex and/or to provide bibliographic data:

```
<note>Sprostranov was the first to attribute the whole collection to the single compiler
<bibl><ref
target=" ../bibliography/scyr.xml#Sprostanov1902">Sprostranov
1902: 52-56</ref></bibl></note>
```

*Information about text history for the whole manuscript*

The text historical information on the level of the whole manuscripts is part of

the content of the series of element <filiation>. This element has one important attribute, @type, the most common values of which are :

*antigraph*: information about the text collection's antigraph

*apograph*: information about the text collection's apograph

*litRedaction*: information about the textual redaction or recension of the whole manuscript

*protograph*: information about the text collection's protograph

*source*: the source of the whole text collection

*translation*: data about the translation with respect to number and/or origin.

Examples of usage are given below. On the level of the whole manuscript these features are invoked only when conclusions can be drawn for the whole text collection about a common source, protograph, literary redaction, etc.

*The original text on the level of the whole manuscript*

Some collections of texts have a common title or heading, incipit and explicit. In this cases, the element <sampleText> could be used. In all other cases the element <sampleText> should be used to illustrate particular texts from the manuscripts, rather than the whole body of the codex.

*The characterization of each text*

Information about each individual text resembles closely the description of the whole text collection. All features of such a description are given as part of the element <msItemStruct> (*a structured manuscript item*, Burnard & Bauman 2007: msItemStruct). This element has the same three attributes, with the same functions, as <msContents>: @type, @style and @class. In addition to these attributes it may be convenient to use unique identifier for each text as part of the XML document (attribute @xml:id), as well as perhaps @defective with one of the values: *true*, *false*, *unknown*, *inapplicable*. In the Repertorium framework the value of the @xml:id in <msItemStruct> is a combination of the fixed string *ACD* (abbreviation for *Article Content Description*) and the place number of the text in order in the manuscript. A combination of all of the important attributes of <msItemStruct> can be illustrated by this hypothetical example:

```
<msItemStruct defective="false" type="compilation" style="non-narrative" xml:id="ACD1">
```

...

```
</msItemStruct>
```

*The text location*

The exact place of the text occurs in the manuscript is given as part of the element <locus>. In the Repertorium we use this element together with the attribute @n (number) to represent the number of a text as part of a whole text collection:

```
<locus n="1">Ir-25v</locus>
```

The contents of the element here is the exact ordinal location in the manuscript where the text is to be found.

When the one and same text is dispersed across different parts of manuscript a <locusGrp> element is used:

```
<locusGrp n="7">  
  <locus n="7.1">154r-155v</locus>  
  <locus n="7.2">90r-90v</locus>  
</locusGrp>
```

From the electronic description of MS NBKM 309  
by Anissava Miltenova

In this example the text is falsely reordered in two different places in the manuscript, so that the beginning part of the text is bound later than the ending part. In such cases it is mandatory to give the exact order of the leaves. The attribute @n gives here the exact number of the text (7), which is restored to logical order with the help of sub-numbers (7.1 and 7.2).

*The name of the text given by the researcher*

In the approach of Repertorium Initiative, there is a strict distinction between text supplied by the researcher (the so-called “scientific title”) and the original title of a text as it appears in the manuscript copy. Both are important, the original text because it is part of the archaeographic object and the scientific title because it enables researchers to identify and attribute texts consistently, even when the titles may be spelled differently in different manuscripts.

Consider the following example:

```
<msItemStruct xml:id="ACD6" style="narrative"  
  type="translation">  
<locus n="6">200r-203v</locus>  
<title lang="en">Interpretation of the Acts of Apostles. </title>  
<author>St. Hyppolitus of Rome. </author>  
<note>Interpretation on the 9th Proverb: “Wisdom has built her house”.</note>  
<filiation type="source" subtype="Greek"><bibl>Migne, PG 10,  
col. 952-953</bibl></filiation>  
<filiation type="translation"><num>1?</num></filiation>  
<filiation type="protograph">Bulgarian</filiation>  
...  
</msItemStruct>
```

If the title were just transferred automatically from the original text of the manuscript, the text would read “About the 12 Apostles. Where each of them preached”. Such a title would not make it possible to identify other copies of the same text, where the heading in those other manuscripts will probably be different, and it would also not be possible to compare this text with its Greek source. In this case the element <title> should provide a standardized title that is compatible with other Slavonic and non-Slavonic traditions.

*The author of the text*

The author of the text is given as content of the element `<author>`. This element contains the name of the author of the text as attributed by the encoder, and it may not agree with the name of an alleged author included in the manuscript itself. The usage of this element and as well as the element `<respStmt>` is exactly as the illustrated as part of the element `<msContents>` above.

*Special data related to the text*

Data about the text not directly related to the text tradition but nonetheless important for a full description may be encoded as part of the content of the element `<note>` inside `<msItemStruct>`. In the example given above, the content of this element is:

```
<note>Interpretation on the 9th Proverb: "Wisdom has  
built her house"</note>
```

In this example, the element's content explains the title given by the researcher. Without the knowledge that this fragment is part of a larger text, which has its own tradition, it would be impossible to undertake further steps in the investigation. This element may also be used to explain text losses and interpolations, which may provide keys to understanding the textual tradition.

*Dates for texts sung or read on particular occasions*

As part of the same element `<note>` inside the `<msItemStruct>` element we could enter fixed or movable dates presented in the manuscript copy or known by the encoder in comparison with other sources. Consider the following example:

```
<title xml:lang="en">Vita of John the Hesychast</title>  
<note place="inline">  
<date when="--03-31" type="church" evidence="internal"  
subtype="fixed">31st of March</date>  
</note>
```

From the electronic description of the Codex Suprasliensis  
by Maya Petrova-Taneva

Here we have all possible information about this date in the Codex Supraliensis. The commemoration of St. John the Hesychast in this manuscript is on March 31, formalized as 03-31. The date is used in the church calendar (`type="church"`), is presented in the manuscript (`internal`, i.e., not supplied or conjectured by the encoder) and is part of the `fixed` (non-movable) feast days.

*The language of the text*

Unlike the element `<scribeLang>`, which is intended to give information about the language of the scribe and of the manuscript parts, `<textLang>` can provide material for the linguistic analysis of the textological features of a particular text. Consider the example:

```
<textLang>Only in this text we could find the older forms of the stemmatic
```

Aoristus ... </textLang>

This observation could lead to closer comparison of this text with other parts of the contents, and may help determine whether its origin is different from theirs.

*The text source*

Usually the source in the Slavonic tradition is a Greek or Latin original text that served as the source of translation. In the example above the excerpt:

```
<filiation type="source" subtype="Greek"><bibl>Migne, PG 10, col. 952-953</bibl></filiation>
```

means not that the given text is a direct source for the translation, but that it is cited as a parallel to the Slavic text in the described manuscript.

*The text translation*

If the text is translated from other sources, then in the element <filiation> with the type="translation" the encoder may enter information about the translation, e.g., the number of the translation and the place. A question mark can be used to indicate uncertainty, e.g.:

```
<filiation type="translation"><num>1?</num></translation>
```

*Special parts for textual criticism*

As with the description of the whole manuscript content, at least four values could be given for @type in the element <filiation>: *protograph*, *antigraph*, *apograph*, and *litRedaction*. These elements can provide a context for researching the transmission of the text by situating it within the larger tradition.

```
<filiation type="litRedaction">The manuscript uses archaic lexis and is considered to belong to the archaic group of Apostoloi representing the Cyrillo-Methodian translation. See <bibl><ref>Воскресенский 1879: 111-112</ref></bibl></filiation>
```

From the electronic description of the Manuilev Apostle by Margaret Dimitrova, Andrej Bojadžiev and Elisaveta Mussakova

A combination of several <filiation> elements with different values of @type can specify the exact place of a text copy, as in this example:

```
<filiation type="protograph">Bulgarian</filiation>
<filiation type="antigraph">Middle Bulgarian</filiation>
<filiation type="litRedaction">2</filiation>
```

From the electronic description of MS Odessa, State Library 1/112 by Anissava Miltenova

*The text location and textual criticism*

A comparison of different copies of one and the same text may lead to interesting results. For example MS NBKM 309 and MS Slav. Dujčev 17 share almost the same content, the latter manuscript being the copy of the former. During restoration the content of the MS NBKM 309 was rearranged, leading to disordered folia. As a



result the content of both is now best described by comparison (Милтенова 1989). In the process of electronic cataloguing the contents of the both manuscripts have been reconstructed as part of the process of identifying a possible common ancestor. In electronic form this task could be fulfilled by attribute values in elements `<msItemStruct>` and `<locus>`. In both manuscript descriptions `@xml:id` begins with value *ACD1*. However, the values of the attribute `@n` on the element `<locus>` differ. The first text in MS NBKM 309 is:

```
<msItemStruct xml:id="ACD1" type="translation">
  <locus n="3">109r-110r</locus>
  <title>Calendologion</title>
  ...</msItemStruct>
```

The same text in the description of Slav Dujčev 17 has the same number in `<locus>` but comes in the second place:

```
<msItemStruct xml:id="ACD2" type="translation">
  <locus n="3">3v-6v</locus>
  <title>Calendologion</title>
  ...
</msItemStruct>
```

One text is missing in MS NBKM 309 but it is present in Dujčev 17, where it bears a designation number 2:

```
<msItemStruct xml:id="ACD1" type="translation">
  <locus n="2">2r-3v</locus>
  <title xml:lang="bg">Thunder book according to ascending and descending
  moon</title>
  ...
</msItemStruct>
```

From the electronic descriptions of MSS NBKM 309  
and Slav. Dujčev 17 by Anissava Miltenova

In both manuscripts the first item from the reconstructed cycle is missing, and therefore an element `<locus>` with `n="1"` is absent as well.

Thus, we can have three different points of view about the place of a particular text: 1) as part of the manuscript; 2) as part of the content shared with another manuscript, and 3) as part of the reconstructed cycle of texts.

#### *Decorations for a particular text*

In principle decorations should be described as part of `<physDesc>`. In some cases, however, parts of the cataloguing could be entered as the element `<decoDesc>` and with the same XML content as `<msItemStruct>`. This could be especially convenient in cases when the manuscript copy is compared with other copies of the same text, which could lead to conclusions about specific text formation and layout.

*Description and edition of the colophon*

Unlike later or non-genuine inscriptions and notes, a colophon is considered part of the genuine content of the manuscript. As such it is part of the `<msItemStruct>` element and its description shares the same features as other text. The only exception to this rule is that the text of the colophon is entered as a special element `<colophon>`:

```
<msItemStruct defective="false" type="original"
style="narrative" xml:id="ACD15">
<locus n="15">77v</locus>
<author>Vladislav Gramatik</author>
<colophon xml:lang="cu">Исписа се сиа книга у Нагоричину Младо(м). въ
дому Николѣ спа(н)чевика. и почѣ се писати, м(с)ца ноемвриа. въ кѣ. дѣнь. на възведение
приачистые вѣе въ сѣаа сты(х). и дописа се тогож(д)е м(с)ца ное(м). в гредущее лѣто. дѣ.
днь. у пе(к). дне(м). Въ лѣто ꙗꙗꙗѣ. Владислав диак писа книгу сию. ѿ новаго бръда.
</colophon>
</msItemStruct>
```

From the electronic description of MS Rila 3/6  
by Anissava Miltenova

Sometimes a manuscript contains more than one authentic scribal colophons. In this case all are described and edited in the corresponding places of the description:

```
<msItemStruct type="original" xml:id="ACD6"
class="colophon">
<locus>f. 154v</locus>
<author>Nikita</author>
<note place="inline"><date
evidence="conjecture">1640</date></note>
<colophon>+ Изволеніемъ оца и поспышеніе(м) сиа и съвършеніемъ стго дѣа
аминъ. исписа(х) азъ многогрѣшныи никита сїо книгѣ. сав(в) мита(р) и фадисена. и слово
въ не(д)лю о блѣднемъ. и слово въ не(д)лю месопѣ(о)нї о скончани мира. и слово о изгнани
адамовѣ. и Фулами избрании. и слово о чѣтнѣмъ дрѣвѣ. и кѣпї ю петѣкв велковѣ. да работи
за еговѣ дшѣ и за родители его. велка и мѣри его не(д)ли. и приложи ю оу стго архѣаггла. оу
махаала сръбска ѣ етрополе, вѣ да прости. </colophon>
</msItemStruct>
...
<msItemStruct xml:id="ACD10" type="original"
class="colophon">
<locus>f. 203r</locus>
<author>Daniil of Etropole</author>
<note><date evidence="internal">1640</date></note>
<colophon>+ Изволеніемъ оца и поспышеніе(м) сиа и съвършеніе(м) стго дѣа,
аминъ. исписа(х) азъ еромона(х) данїи(л) сїо книгѣ. и кѣпї ю жива(н) ѿ село етрополе и
приложи ю ѣ стго архѣаггла михаила. оу цѣвѣ махаала сръбска.
```

да слѣжи(т) за негоуѣ дѣиѣ и за подрѣжѣе егово дафинѣ докле трае. и кои се ще покѣси да ю извадитъ из црѣвѣ или да ю посвоитъ. да бѣде(т) проклетъ ѿ га бѣ и ѿ прѣ(с)тѣю его мѣры и ѿ тѣи бѣгосны(х) ѿцѣ. при цара тѣрска сѣлтанъ нбрах(и). и при еп(с)кпа ловецка кѣ(р) сѣмевна. въ лѣто ѿ адама. зрѣи </colophon>

</msItemStruct>

From the electronic description of MS NBKM 1388  
by Dilyana Radoslavova

The description and edition of the colophon could include data about the formula of the text in terms of diplomatic structures, e.g., the form and parts of *protocollo*, *testo*, *eschatollo*:

```
<formula>protocollo(invocatio)+ ...</formula>
```

Then in the text the various parts could be encoded with the help of the element <seg> with a corresponding value of the attribute @type:

```
<colophon> <seg type="protocollo">+ Изволенѣмъ оца и поспѣшенѣ(м) сѣа  
и сѣврѣшенѣемъ стѣго дѣа аминъ</seg> ...</colophon>
```

### *The original text from the manuscript*

#### *Headings, incipita and explicita as part of text descriptions*

The Repertorium uses a special element <sampleText> as a wrapper around several elements used for the identification of text items in a manuscript. Its function is to make a distinction between the language of the description (metadata used to describe the manuscript) and the original text. The element itself uses one global TEI attribute, @xml:lang, which represents the language of the content, and for medieval Slavonic codices its value is almost always *cu*. The optional attribute @script commonly contains the value *Cyrs* (for old Cyrillic) or *Glag* (for Glagolitic).

```
<sampleText xml:lang="cu" script="Cyrs">
```

```
  <head>Мѣца тогожде · ѿ · днѣ · словъ · ѿ рождѣствѣ ивана прѣче и ѿ оумрътѣю  
  ѿца его захарѣе ѿче бл҃сви · </head>
```

```
  <incipit>Въ лѣто четвѣртодесетное црѣтвующоу ироду възыскание сѣтвори  
  изъвити млдѣнце, иже въ иерѣлимѣ</incipit>
```

```
  <explicit>.црѣтвова же архела дѣ лѣтѣ испльнь. въ нас же црѣтвующоу боу  
  нашему иѣоу хѣоу · емоу же слава(а) съ ѿцѣемъ и съ стѣмъ дѣхомъ нѣна и присно и вѣкы  
  вѣкым(ъ) аминъ.</explicit>
```

```
</sampleText>
```

From the electronic description of MS NBKM 1039 by Dessislava Atanasova

#### *Samples from the original text of the manuscript*

In some older catalogues there is a special section that contains a sample of the manuscript text. When creating an electronic version of such materials, <sampleText> is a convenient way to preserve these text examples.

### *Edition of rare or un-edited texts*

In some manuscripts there may be texts that are rare or perhaps completely unknown to researchers, and the edition of such texts in the description of a manuscript could serve as a preliminary publication. In such cases an element `<div>` (division, Burnard & Bauman 2007: `div`) should be used with an appropriate value in the attribute `@xml:lang`.

### *Original texts in other parts of the electronic description*

Other parts of the description might also include examples and citations from the manuscript source. These could be entered in two ways:

1. As part of the content of the element `<foreign>` with an appropriate value in element `@xml:lang`. In this case `<foreign>` means in a different language in comparison to the rest of the text.

2. As part of a linguistic segmentation model. In the TEI framework it is possible to choose among the elements `<c>` (character), `<cl>` (grammatical clause), `<m>` (grammatical morpheme), `<pc>` (punctuation character), `<phr>` (grammatical phrase), `<s>` (sentence-like text division), `<seg>` (arbitrary segmentation) or `<w>` (grammatical word), cf. Burnard & Bauman 2007: `model.segLike`. As with the elements `<sampleText>` and `<foreign>`, these should all have attribute `@xml:lang` with the corresponding value.

### *Bibliographic references in <msItemStruct>*

The last element in the content of `<msItemStruct>` is for bibliographic references, which are restricted to the attribution of text and its edition. They come in two variants: 1) one bibliographic reference in the element `<bibl>`, or 2) a list of more than one reference (element `<listBibl>`). In the framework of the Repertorium project all bibliographic references for all manuscript descriptions are gathered in one place, and the XML files that contain the descriptions of the individual manuscript hold only pointers to these external reference lists. These pointers could look like the following:

```
<bibl><ref target="../bibliography/i.xml#IvanovJ1932">Иванов
1932</ref></bibl>
or
<listBibl>
  <bibl><ref target="blah-blah.org"></ref></bibl>
  <bibl><ref target="blah-blah.org"></ref></bibl>
<!-- Here more bibliographic references -->
</listBibl>
```

### *Description of texts with complicated structure*

There are many texts in the medieval tradition that are composed, in turn, from other texts or text parts, and the Repertorium model, building on TEI, provides a method for describing such texts as part of the larger ones. This task could be accom-

plished by nesting `<msItemStruct>` inside another `<msItemStruct>` element:

```
<msItemStruct>
  <msItemStruct>...</msItemStruct>
  <msItemStruct>...</msItemStruct>
</msItemStruct>
```

The elements can be synchronized as in this example:

```
<msItemStruct xml:id="ACD15">
  <locus n="15">283r-311v</locus>
  <title xml:lang="bg">Физиолог</title>
  <title xml:lang="la">Physiologus</title>
  <note place="inline">
    <bibl><ref target=" ../bibliography/scyr.xml#
StojkovaA1994">Стойкова 1994</ref>
  </bibl>
  <bibl><ref
target=" ../bibliography/acyr.xmlAleksandrovA1893">Александров
1893</ref>
  </bibl>
  </note>
  <filiation type="protograph">Bulgarian</filiation>
  <filiation type="redaction">Pseudo-Basil's</filiation>
  <sampleText xml:lang="cu">
    <head>Начело бѣствословнику имѣиѣ изъвблѣченїа скотѣмъ
животнымъ</head>
    <explicit>Тако въ неприлежанїи ходеще ниже мѣтѣвою. ниже
вдѣнїемъ въспоминають чѣда своя. вънемагающе бжїю помощь о грѣсѣхъ ихъ. </explicit>
  </sampleText>
  <msItemStruct xml:id="ACD15.1">
    <title xml:lang="bg">За лѣва</title>
    <title xml:lang="en">About the Lion</title>
    <sampleText xml:lang="cu">
      <head>Пръвое в лѣвъ </head>
      <incipit>Начнѣмъ рѣци в лѣвъ црѣв звѣремъ. теже какѣвъ скотѣ естъ. иво
Іаковъ блѣваше Іоуду глѣаше. скимень лъвовъ.</incipit>
    </sampleText>
  </msItemStruct>
  ...
</msItemStruct>
```

From the electronic description of MS Pant. monastery,  
Mount Athos 22 by Ana Stojkova

In this example the main text is called „Physiologus” and it contains more than 40 small stories. The first one is „About the lion”. The numbers and identifiers are set

in such a way that it is possible from their value to see the text hierarchy (cf. values ACD15 and ACD15.1). Every text could have its own standardized title, as given by the research tradition, along with its own incipit and/or explicit, as found in the manuscript (cf. also Стойкова 2011).

### *The history of the manuscript*

#### *General remarks*

The location of the manuscript date and place differs across descriptive models. The deviations are caused not by the particular research culture (e.g., Byzantine Greek, Slavonic, Latin, Armenian, Georgian, or Hebrew), but, rather, by the more general assumptions of local traditions. Very often the date of a manuscript is represented as part of the heading of the description, together with the place of origin and its name. In the Repertorium framework historical information about a manuscript is unified within the element `<history>`.

#### *Date and place of origin of the manuscript*

This element could contain a short overview of the history of the manuscript in the form of one or more paragraphs (element `<p>`) or it could contain a structured description of its origin, provenance and acquisition in a library or archive, possibly introduced by some kind of summary (element `<summary>`). The original date and place of the manuscript are containing in the element `<origin>`:

```
<origin>
  <origDate evidence="internal"
cert="high">1658</origDate>
  <origPlace evidence="internal"
cert="high">Etropole</origPlace>
</origin>
```

From the electronic description of MS NBKM 78  
by Dilyana Radoslavova

In this example the elements `<origDate>` and `<origPlace>` both have attributes `@evidence` and `@cert` (certainty). The value *internal* means that we know from the manuscript about the date and place of origin. Additionally we can provide the attribute `@cert` with a value such as, for example, *high*, which means that we believe that the data in manuscript is the true one. Other possible values of `@cert` are *low*, *medium* or *unknown*.

#### *Notes and additional materials about the history of the manuscript's possession and acquisition*

Notes representing data about manuscript possession and acquisition, as well its edition, should be encoded as part of the manuscript history in the element `<accMat>` (accompanying material, Burnard & Bauman 2007: `accMat`):

```
<history>
  <accMat>
    <p>Consist of notes, one dated 1926, by <persName>François Dimitry
de Rochefort</persName>, alias <persName>Dimitrij Adol'fovič</persName>, indicating
his relationship to the Oznobišin family and saying what the MS is. </p>
  </accMat>
...
</history>
```

From the electronic description of MS British Library Add.  
MSS 43508 by Dilyana Radoslavova

### *Provenance*

All facts concerning the provenance of the codex are described as part of the element `<provenance>`. In this element it is possible to encode all names and dates as one and more sentences:

```
<provenance>The MS belonged to <persName type="owner">Engelbert
Kämpfer</persName></provenance>
```

From the electronic description of British Library Sloane  
MS 2910b by Dilyana Radoslavova

This part of description should provide all facts and events prior to the manuscript's acquisition in the library or archive (Burnard & Bauman 2007: `provenance`).

### *Acquisition*

This element documents the way an archived item entered a particular library or collection, together with a possible reference to the sources:

```
<acquisition>Purchased by <persName>Harley</persName> from
<persName>Richard Jones</persName>, bookseller, in 1715 <bibl><ref
target="bibliography/w.xml#WrightCE1966">Wright
(1966)</ref></bibl></acquisition>
```

From the electronic description of British Library Harley  
Ms 3519 by Dilyana Radoslavova

### ***Additional information in the description***

This final part of the electronic manuscript catalogue contains:

- administrative information about the description and manuscript custody
- facts about copies of the manuscript
- a bibliographic list of known publications of or about the manuscript

#### *Administrative information*

The administrative information in this model consists of a recording of changes in the electronic description (`<recordHist>`), a statement of the manuscript's availability (`<availability>`), a sketch of custodial events (`<custodialHist>`), and optional extra or complementary data (`<note>`).

### *Recording the history of the electronic descriptio*

The element `<recordHist>` can be used to document all changes in the electronic description. In this element, as elsewhere in the TEI, the encoder has two ways to submit data – as one or more paragraphs (`<p>`) or as a sequence of specialized elements, which in this case are `<source>` and `<change>`. The `<source>` is not repeatable, but because there could be many changes in the electronic description, `<change>` is optional and repeatable.

```
<recordHist>
  <source>Derived from the second volume of Benjo Tsonev's description of
  manuscripts in the Sofia National library</source>
  <change when="2012-02-02"><name>Andrej Bojadžiev</name>
  changed the electronic description to reflect the last modifications in the
  model</change></recordHist>
```

### *Availability of the manuscript*

Unlike the element `<availability>` as part of the `<teiHeader>` (cf. below), the appearance of this element in the context of administrative information refers to the availability of the manuscript itself. In most of the cases access to the archival collection is restricted, which is reflected in an entry like the following:

```
<availability><p>Restricted. Used by permission and by appointment with the
curator</p></availability>
```

The encoder might wish to consult the archive itself for the appropriate formulation of rules for access to the manuscript.

### *Custodial facts*

Here too, the encoder has two ways to submit data – as one or more paragraphs (`<p>`) or as one or more `<custEvent>` (*custodial event*, Burnard & Bauman 2007) elements. This element might include facts about restoration, conservation of the manuscript, participation in exhibitions, etc., e.g.:

```
<custodialHist>
  <custEvent type="exhibition" from="2011-05-11"
to="2011-06-11">Part of the exhibition "Slavic Manuscripts from the Historical
Museum in Kjustendil"</custEvent>
</custodialHist>
```

Every event in the history of the manuscript can be encoded as a separate `<custEvent>` element containing the attribute `@type`, which further specifies its kind. The attributes `@from` and `@to` give information about the date span.

### *The note in the administrative information*

Other data that cannot be formalized in advance could become part of the element `<note>` as part of the administrative information (`<adminInfo>`).



### *Copies of the manuscript*

Information about various known copies of the manuscript and their characteristics are encoded as part of the element `<surrogates>`. Consider the following example:

```
<surrogates>
<p>
  <label>Microform</label>
  <idno>Central Library of BAS, MF-11</idno>
</p>
<p>
  <label>Digital images</label>
  <list>
    <item><locus>f. 184</locus>
      <idno type="file">C5246-07</idno></item>
    <item><locus>188</locus><idno type="file">X5247-07</idno>
  </item>
  ...
  </list>
</p>
</surrogates>
```

Some manuscript may have multiple copies. This information is recorded as a list of items, each of which has information about the place in the manuscript the picture is taken from and identifying information about the file and its name.

### *A bibliographical list*

The last item in the electronic manuscript description is a compiled bibliographic list of references. This reference list should include only publications about the described manuscript, such as earlier descriptions or editions or special research related to the palaeographic, codicological, text historical or linguistic features of the manuscript. All other types of bibliographic references should be mentioned in the appropriate places in the description.

### *Encoding parts of the manuscript*

The element `<msPart>` (i.e. *manuscript part*) should be used in the cases when the manuscript is a) a convolute; b) a palimpsest, c) a supplemental physically different part in the manuscript or a d) supplemental and additionally written literary part of the manuscript. Several cases can be distinguished:

- When the codex is re-bound and later end leaves and paste-downs are inserted by the rebinding process this information should be entered in the corresponding sections of the element `<binding>`;
- The same is true in cases when the parchment or paper is reinforced as part of the rebinding process.

- However, when a later leaf or leaves is inserted in order to supplement the text and this part is obviously different, the element <msPart> should be used to describe the codicological, palaeographic, language and text features of this inserted part.

- The element msPart is not used to distinguish parts of the text written by the particular scribes. This information is part of the element scribe and its sub-elements.

- The element msPart is not used when the manuscript is now kept in various repositories or in one repository under different shelfmarks. This information is entered in element msIdentifier and its sub-element altIdentifier.

The element <msPart> consists of the same elements for identification and the description of physical features, language, palaeography and text as mentioned above for an entire manuscript.

The identification part has several variants:

1. With the element <msIdentifier> – usage should be restricted to cases where the physically different part of the manuscript is kept in another repository or within the same repository but under a different shelfmark.

2. With the element <altIdentifier> – could be used to mention a special name of the manuscript part, e.g., in the case of palimpsests.

3. The most obvious and convenient way to describe physically distinct part of the manuscript is to use the elements <locus> and <locusGrp>.

### *Description of palimpsests*

The description consists of a common cataloguing part and two or more <msPart> elements, as in this example:

```
<msDesc>
  <msIdentifier> ... Here comes the common information for all parts, e.g.,
country, repository, shelfmark, etc. </msIdentifier>
<physDesc>
  <objectDesc form="codex" >
    <supportDesc material="parchment">
<p>The manuscript is a palimpsest consisting of ... parts</p>
      <!-- Common data about the both parts -->
    </supportDesc>
  </objectDesc>
</physDesc>
  <msPart n="1" type="upperLayer">
    <altIdentifier type="laterPart">
      <msName>e.g., Boyana Gospel</msName>
    </altIdentifier>
    <physDesc> ... </physDesc>
    <scribeDesc> ... </scribeDesc>
    <msContents> ... </msContents>
    <history>
      <origin>
```

```

                <origDate></origDate>
                <origPlace></origPlace>
            </origin>
        </history>
        <additional> ... </additional>
    </msPart>
    <msPart n="2" type="palimpsest">
        <altIdentifier>
            <msName>e.g., Boyana Palimpsest</msName></altIdentifier>
            <physDesc></physDesc>
            <scribeDesc> ... </scribeDesc>
            <msContents> ... </msContents>
            <history>
                <origin>
                    <origDate></origDate>
                    <origPlace></origPlace>
                </origin>

                <additional></additional>
            </history>
        </msPart>
    </msDesc>

```

Somewhere at the beginning of the physical description should be a short statement about the fact that the manuscript in question is a palimpsest. Sections about manuscript cataloguing, history, etc. are almost always common for all data with one exception: the original dates of writing and places of origin differ, and this information should be provided explicitly. In cases involving restoration and references there could be a need for additional data sets. Parts are further specified with the help of @ type.

### *Description of convolutes*

From the point of view of electronic encoding there is not much difference between palimpsests and convolutes, even though, unlike palimpsests, convolutes can be formed from paper and are defined not in terms of layers, but as a horizontal sequence of parts.

```

<msDesc>
    <msIdentifier> ... Here comes the common information for all parts, e.g.,
country, repository, shelfmark, etc. </msIdentifier>
    <physDesc>
        <objectDesc form="codex" >
            <supportDesc material="parchment">
<p>The manuscript is a convolute consisting of ... parts</p>
                <!-- Common data about the both parts -->

```

```

        </supportDesc>
</objectDesc>
</physDesc>
  <msPart n="1" type="original">
    <locus>
      ...
    </locus>
    <physDesc> ... </physDesc>
    <scribeDesc> ... </scribeDesc>
    <msContents> ... </msContents>
    <history>
      <origin>
        <origDate></origDate>
        <origPlace></origPlace>
      </origin>
    </history>
    <additional> ... </additional>
  </msPart>
  <msPart n="2" type="convoluted">
    <locus> ... </locus>
    <physDesc></physDesc>
    <scribeDesc> ... </scribeDesc>
    <msContents> ... </msContents>
    <history>
      <origin>
        <origDate></origDate>
        <origPlace></origPlace>
      </origin>

    <additional></additional>
    </history>
  </msPart>
</msDesc>

```

In principle, convolutes do not require special manuscript names or shelfmarks. They do, however, require a special section pointing to the place(s) where the parts are to found. In the example above the element `<locus>` fulfills this role.

#### *Description of additional folia or parts of the manuscripts*

The use of the `<msPart>` element might be appropriate when there are additional parts in the manuscript that are not to be considered as convoluted, or do not form a part of binding (paste-downs or end-leaves), and which might intersect with the provenance of the manuscript. Such parts should be described as separate element `<msPart>` with the value of the attribute `@type` *additional*:

```

<msPart n="1" type="additional">
  <locus>
    ...
  </locus>
  <physDesc> ... </physDesc>
  <scribeDesc> ... </scribeDesc>
  <msContents> ... </msContents>
  <history>
    <origin>
      <origDate></origDate>
      <origPlace></origPlace>
    </origin>
  </history>
...
</msPart>

```

Examples of such usage could include an additional note on a separate leaf, or text from a later scribe.

#### *Additionally written part(s) of a manuscript*

A special case could be a separate text written on the same folia as the original manuscript but containing an entire separate literary work. This addition is best described inside a special `<msPart>` element with its own location, name of the scribe, if available, and, of course, with a description of the content. Very often these texts are unique in the tradition and they deserve a special edition inside the description. In such cases an element `<div>` with its sub-elements should be provided:

```

<msPart>
  <locus>bottom margins</locus>
  ...
  <msContents>
    <div xml:lang="cu">

```

Пишемъ азъ ереи Петръ, да се (знае) какво врьме вьше. (Бех) попь на село Мирково и вьше попь Стоянь, заедно дръжѣхме село Мирково ... при злочьстиваго цара сульгань Мехмеда, да го богъ убие и да го порази, че той расипа христѣнски вьре и велики цркви и велики градове ... Бьше честита земѣ Загоре, та па се яви бдйракъ от запать, та доде до срѣдь демле от Бечъ, от Будимь, от Янекъ, дори от Злате явлъке. Та се провьвише турци и прѣзморци, татаре и каталане и хиньдие и миньдие, и опашата вьра, та ги вдигѣ царь Мехметъ на войске и самси трьгнѣ, Кара кеѣя му бе везирь. Та трь... този лѣто изгине войска турска, исврахе ги латинска войска. Ёдва утече царь Мехметъ и Кара кеѣя назать. И да знае, кой чете да види, че беше от Бога, че даде Богъ жито сьпо, сирене сьпо, а на Загоре рож ходеше гроше. Та беше зло и невола и зими зли. Да не даде богъ веке такьвись години зли въ лето зр.с.й. (1690). Та пакъ излезе сень паша от Прѣзьморе, та доде на Цариградь при цара Мехмеда. Та му даде царь Мехметъ Софие да дръжи пашалжкъ. Та отиде сень паша на войске на Будимь, та са би сась алдманци този лето. А той нема цю да стори на алдманци, ами доде, та събра сете войскѣ турска, та са врьне цару Мехмеду на Цариградь, та удари цара Мехмеда на Цариградь. Та се

стори чүдо вовь Цариградь, изгинехе турци аване, тефтерьдаре, и капицие, и востанцие и цара Мехмеда затвори, а то стори брата му царь Сулимань. Та се стори жито този лето берекеть, а Сень паша се врзне на Софие. Доде до(й)де на Софие, а той расипа аване, каде мене. Та че доде на Софие та се истроваха аване, и кадие, и мольбата от страхае Сенови. Та седе на Софие месець, та че пусти царь Сулимань улака до него да иде на Бельградь да се бие, а той не ще да иде. А царь стори нифрамъ на него да се бие. Та затворихе клисурите и друмовете да го не пустеть. Та седехе по месець по деръвенете турци и рая и кадие, та да не утече Сень паша. А той трзгне тогись на Бельградь да отиде, а то дадохе алманци, та превзехе Белиградь и опрехе до Нишь. И Сень се запръхе на Нишь, а тие отидохе на Босне да пръвземать Босне. А Сень пусти мойске, та оплени Кипровець и изьгуби и пороби...

```
</div>
</msContents>
<additional>
  <listBibl>
    <bibl type="edition">Дринов 1882</bibl>
    <bibl type="edition">Цонев 1910: 447–448</bibl>
  </listBibl>
</additional>
</msPart>
```

*A complicated example: a palimpsest kept in various repositories*

One particularly complicated case for description is a Sinai Slavic Miscellany that now is kept in two libraries, where it is divided into 5 parts. The manuscript, written on parchment and paper, has palimpsest parts in Cyrillic, Glagolitic and Greek. The description begins with the heading:

```
<msDesc>
<head>Codex Sinaiticus, 12th c., the end-13 c., and 13 c.</head>
...
</msDesc>
```

In this case we have separate `<altIdentifier>` elements following immediately after the head:

```
<altIdentifier type="partial"><country>Egypt</country>
<geogName>
<geogFeat> Mount</geogFeat><name>Sinai</name></geogName>
<repository>Monastery "St. Catherine"</repository><idno
type="shelfmark">Sin 34</idno></altIdentifier>
<altIdentifier type="partial"><country>Egypt </country>
<repository>Monastery "St. Catherine"</repository><idno
type="shelfmark">Sin 18/N</idno></altIdentifier>
<altIdentifier type="partial"><country>Russia</country>
<settlement>Sankt Peterburg</settlement><repository>Russian
National Library</repository><idno type="shelfmark">Gr.
70</idno></altIdentifier>
```

```
<altIdentifier type="partial"><country>Russia</country>
<settlement>Sankt Peterburg</settlement><repository>Russian
National Library</repository> <collection>Osnovnoe
sobranie</collection><idno type="shelfmark">Q.p.I.63
<idno></altIdentifier>
<altIdentifier type="partial"> <country>Russia</country>
<settlement>Sankt Peterburg</settlement><repository>Russian National
Library</repository><collection>Osnovnoe
sobranie</collection><idno type="shelfmark">Q.p.I.64
</idno></altIdentifier>
```

Then come physical characteristics like number of leaves, dimensions, and layout characteristics that do not belong to one particular part, followed by a general overview about the manuscript condition. After this follow the characteristics of all 7 scribes who wrote the upper layer of the text. Consider the following example:

```
<scribe n="E"><locusGrp><locus>Q.п.I.63, 1r-3r</locus>
<locus>Sin. 34, 63v-71v</locus><locus>Sin. 34, 77r-81r
</locus></locusGrp>
<palaeoCharact script="cyrs"><p>Uncial</p></palaeoCharact>
<scribeLang><summary>Middle Bulgarian with Russian
features</summary><orthography><orthNote type="jer"
subtype="front">only
<c xml:lang="cu">Ѣ</c> </orthNote><orthNote type="jus"
subtype="nonConsis"><c xml:lang="cu">Ѧ</c> replaced by
<c xml:lang="cu">Ѧ</c>; <c xml:lang="cu">Ѧ</c>
replaced by <c xml:lang="cu">Ѧ</c>;</orthNote><orthNote
type="otherLetters"><c xml:lang="cu">Ѧ</c>;
<c xml:lang="cu">Ѧ</c> and
<c xml:lang="cu">Ѧ</c></orthNote></orthography>
... </scribeLang></scribe>
```

Here the place where scribe E wrote is identified and attributed, followed by a short overview of his palaeography and language.

A description of the content applies to all parts together, and begins with general information about the origins of all of the texts:

```
<msContents>
<filiation type="protograph">Middle Bulgarian</filiation>
<filiation type="antigraph">Russian</filiation>
...
</msContents>
```

Since the texts now are in different repositories, their sequence and locations are especially important. Consider the following example:

```
<msItemStruct xml:id="ACD1">
  <locusGrp>
```

## A. Bojadžiev

```
<locus n="Sin.18N.1">Sin. 18/N, 1v-8v</locus>
  <locus n="Gr.70.2">Gr. 70, 54r-98r (54, 59, 60, 57,
58, 55, 56, 61-98)</locus>
</locusGrp>
...
<title xml:lang="la">Pandectae</title>
<author xml:lang="la">Nicon monachus</author>
...
</msItemStruct>
```

The cataloguing information of the common part ends with data about the manuscript's origin, history and administrative information. Then begins a description of each of the palimpsest's parts, for example:

```
<msPart type="palimpsest" subtype="Glagolitic"
corresp="QpI64">
<locus>ff. 1-8</locus>
<physDesc><objectDesc><layoutDesc><layout>one
column</layout></layoutDesc></objectDesc></physDesc>
<scribeDesc><scribe><palaeoCharact script="glag">
<p>Round
Glagolica</p></palaeoCharact></scribe></scribeDesc>
<msContents><p>Office menaion or Troparia from Octoechos
(1958)</p></msContents>
<history><origin><origDate>11th c.</origDate></origin>
</history></msPart>
```

All examples in this section are from the electronic description made by Anissava Miltenova and Yavor Miltenov

### *Recommended minimum metadata content in manuscript description files for the Repertorium*

The metadata information contained in the <teiHeader> element conforms completely to the TEI Guidelines *Minimal and recommended headers* (Burnard & Bauman 2007: Chapter 2.6). The only mandatory element is the description of file (<fileDesc>), which, in the framework of Repertorium, should give the following information:

- Title of the document
- Author of the document
- Editor and/or other person responsible for the content of the document
- Publisher
- Place of publication
- Date
- Internal Repertorium identifier
- Availability of the document
- Information about the primary source for writing the document
- Documentation of possible document revisions



### *The title and creators of the documents*

This information comes at the beginning of the <fileDesc>:

```
<teiHeader>
  <fileDesc>
    <titleStmt><title>Codex Suprasliensis: reading menaion for March,
10th century</title><author xml:id="MP">Maya Petrova-
Taneva</author><editor xml:id="AM">Anisava
Miltanova</editor><respStmt>
<name xml:id="AB">Andrej Bojadziev</name><resp>edition of XML
code</resp></respStmt>
    </titleStmt>
  </fileDesc>
</teiHeader>
```

The first part of the TEI file contains the <titleStmt> (i.e., title statement), which unifies information about the document title (<title>) and creators. In this example there are three types of creators: author, editor and the person who edited the XML code. Each of these creators is identified by the value of the @xml:id attribute.

### *Information on document publication*

The second type of information deals with the publication of the document. It includes the name of the publisher, place and date of publishing, internal Repertorium identifier and availability (available, restricted, etc.). A sample of such information could look as follows:

```
<publicationStmt><publisher>e.g., Institute of Literature,
BAS</publisher><pubPlace>e.g., Sofia</pubPlace><date>e.g.,
1998</date><idno type="Repertorium">RC-
MPCodexSuprasliensis</idno>
<availability status="free">
<p>The intellectual content of this file is copyrighted by the creators under
the Creative Commons Attribution-No Derivative Works 3.0 under
legislation in ... </p>
</availability>
</publicationStmt>
```

The Repertorium identifier is formed from the abbreviation RC, meaning Repertorium code, followed by an en dash, the first letters of the author's (i.e., encoder's) name, and an abbreviation of the repository and shelfmark. When the manuscript has a special name or is divided into several repositories, a unique name is used.

The content of the works in the Repertorium are published under a *Creative Commons* license that allows users to share (to copy, distribute and transmit the work) under the condition that they must attribute the work in appropriate way and may not alter, transform, or build upon it (<http://creativecommons.org/licenses/by-nd/3.0/>). This

license is valid for the content written by the researchers. For TEI licensing and citation guidelines for software and text materials users may consult the TEI: Licensing and Citation part of TEI Guidelines (<http://www.tei-c.org/Guidelines/access.xml>). The XML additions and changes to the TEI models are governed by that same policy.

### *The source of the document*

The last part of the <fileDesc> element deals with sources. In the case of a manuscript description that is intended to serve as the sole content of the electronic file, this information may be very brief. A thorough explanation of the sources for the description could then be part of the <recordHist> element in the administrative part of the file (<adminInfo>). A short example of a bibliographic reference could be:

```
<sourceDesc><p><ref type="edition"
target="../bibliography/zbib.xml#ZaimovJ_CapaldoM1982">
Займов, Капалдо 1982</ref></p></sourceDesc>
```

If, however, the description is to be accompanied by an edition of the manuscript text, then the <sourceDesc> should contain all elements of the description, beginning with the <msDesc>. In that case the edited text of the manuscript would constitute the body of the document:

```
<TEI><teiHeader> ...
<sourceDesc><msDesc>...</msDesc></sourceDesc> ...
</teiHeader>
<text><body> ... </body></text></TEI>
```

### *Non-bibliographic aspects of the document*

This information forms part of the <profileDesc> element (Burnard & Bauman 2007: profileDesc). Here the following aspects are encoded:

- exact date of the first document edition (day, month and year)
- language usage in the document
- classification of the described text

The following example illustrates the use of these characteristics:

```
<profileDesc>
  <creation><date when="1998-12-19">19.12.1998</date>
</creation>
<settlement>Sofia</settlement></creation>
<langUsage>
  <language ident="ch">Old Slavic</language>
  <language ident="bu">Bulgarian</language>
  <language ident="en">English</language>
</langUsage>
<textClass>
  <keywords scheme="Repertorium"><list>
  <item>Reading menaion for March</item><item>March</item>
  <item>Menologion</item><item>Vitae</item><item>Homilies<item>
  </list></keywords></textClass>
</profileDesc>
```

The creation (date and place) information is for internal usage. The form of the date as part of the attribute is given as the sequence yyyy-mm-dd (year-month-day). Most of the files in the Repertorium are in Bulgarian, English and Old Bulgarian (Old Slavic, Old Church Slavonic). Latin is also used widely, especially in some manuscript names, e.g., Codex Zographensis. In this place in the description it is encoded as:

```
<language ident="la">Latin</language>
```

The Repertorium uses its own metadata scheme for metadata, encoded with `<keywords scheme="Repertorium">`. The metadata itself is always represented as a list of keywords.

### *Descriptions of possible document changes*

Changes or revisions are recorded as part of the element `<revisionDesc>`. Each change is documented separately, starting with the most recent, e.g.:

```
<revisionDesc>
  <change who="#AB" when="2010-07-10">Conversion to new
  Repertorium scheme. Included head element and its content. Small editorial changes.
  Rearranged revisionDesc: now the last change comes to the first place. </change>
  <change who="#AB" when="2009-04-17">Removed constructions
  &lt;author&gt;anonumous&lt;/author&gt;</change>
  <change who="#AB" when="2009-04-17">Added copyright
  statement</change>
  <change who="#AB" when="2009-04-13">Added Repertorium
  Identifier</change>
  <change who="#AB" when="2009-04-15">Proofreading</change>
</revisionDesc>
```

Here the value of the attribute `@who` points to one of the content creators identified elsewhere. The date in the attribute is in year-month-day format. This list of revisions is intended to reflect primarily the technical side of the document. The documentation of editorial changes of a manuscript description might more properly be encoded in the `<recordHist>` as part of the administrative information (`<adminInfo>`).

### *Usage of some elements common with TEI P5 declarations*

#### *Divisions*

In the TEI approach there are two ways to structure a given text at a high level: numbered and unnumbered text divisions. The Repertorium uses only the latter approach. In the description of manuscripts in the Repertorium project the element `<div>` is used only in cases when an unknown or rare text or text variant is discovered and merits inclusion as a preliminary edition.

#### *Paragraphs*

Paragraphs (`<p>`) may appear in various places in the description—essentially anywhere where there is not enough data for a structured description or where such a description is otherwise unnecessary or undesirable. A possible example might be par-

agraphs in the element `<condition>`, where different types of damage or loss could be grouped together, as in this example from the description of *Codex Suprasliensis*:

```
<condition>
<p xml:id="missedparts">The beginning and the end of Ljubljana part are
missing. Two leaves have been lost from the middle of the Vita of St Paul the Simple,
after f. 86v. There are lacunae between ff.78v-79r, 84v-85r. </p>
<p xml:id="replacements">f. 38 (containing part of the Vita of SS Terentius,
Africanus and Pompilus) is bound in the wrong place and should be read after f. 88v.
</p>
<p xml:id="trimmededges">The edges of the folia in the Ljubljana part were
trimmed during the restoration of this part of the MS, with the result that the parts that are
now in Saint Petersburg and Warsaw are slightly larger (<dimensions unit="mm"
type="folia"><height>313</height>
<width>250</width></dimensions></p>
</condition>
```

### *Bibliographies and references*

Bibliographies provide a choice between one simple bibliographical reference (element `<bibl>`) and a list of such references (element `<listBibl>`). Both elements are described in TEI Guidelines (Burnard & Bauman 2007). In the Repertorium framework neither of these should contain a full bibliographical record; Repertorium practice is that these elements contain only pointers to full bibliographic information, which is encoded at the end of a description or, when the description is part of the larger project, in a separate file.

### *Conclusion*

The Repertorium Initiative provides detailed, structured additions to the TEI XML manuscript description model. It complements the TEI model in some parts through the rearrangement of known elements, the addition of new ones, and the extension of the base of specialized attributes, all with an eye toward intra-project consistency and support for particular details of the Slavonic philological tradition.

### *REFERENCES*

- Altheim, Murray & Shane McCaron (eds.). 2001. XHTML<sup>™</sup> 1.1 – Module-based XHTML. W3C Recommendation 31 May 2001. <http://www.w3.org/TR/2001/REC-xhtml11-20010531/> (12.02.2012).
- Berglund, Anders (ed.). 2006: Extensible Stylesheet Language (XSL) Version 1.1. W3C Recommendation 05 December 2006. <http://http://www.w3.org/TR/2006/REC-xsl11-20061205/> (12.02.2012).
- Berners-Lee, T., Fielding, R., Masinter, L. (eds.). 2005: Uniform Resource Identifier (URI): Generic Syntax. *RFC 3986*, IETF, January 2005. <http://www.ietf.org/rfc/rfc3986.txt/> (12.02.2012).

## *Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions*

- Birnbaum, David J. 2009. An XML user steps into, and escapes from, XPath quicksand. Balisage: The Markup Conference 2009, August 11 - 14, 2009. <http://www.balisage.net/Proceedings/vol3/print/Birnbaum01/BalisageVol3-Birnbaum01.html> (12.02.2012).
- Birnbaum, David J. & Anissava L. Miltenova (eds.). 2000. *Medieval Slavic Manuscripts and SGML: Problems and Perspectives*. Sofia: Marin Drinov Publishing House.
- Birnbaum, David J., Andrej T. Boyadzhiev, Milena Dobрева, Anissava L. Miltenova (eds.). 1995. *Computer Processing of Medieval Slavic Manuscripts. Proceedings. First International Conference, 24-28 July 1995, Blagoevgrad, Bulgaria*. Sofia: Marin Drinov Publishing House.
- Boag, Scott, Don Chamberlin, Mary F. Fernández, Daniela Florescu, Jonathan Robie, Jérôme Siméon (eds.). 2007. XQuery 1.0: An XML Query Language. W3C Recommendation 23 January 2007. <http://www.w3.org/TR/2007/REC-xquery-20070123/> (12.02.2012).
- Bojadžiev, Andrej. 2000. Codicology, Palaeography, and Language in the Computer Description of Medieval Slavic Manuscripts. In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov, 2000. 69-90.
- Bojadžiev, Andrej, Anissava Miltenova & Dilyana Radoslavova 2003. A Unified Model for the Description of Medieval Manuscripts? *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings of the „Electronic Description and Edition of Slavic Sources” conference, 24-26 September 2002, Pomorie, Bulgaria*. Sofia: Boyan Penev Publishing Center, Institute of Literature. 113-135.
- Boyadzhiev, Andrej. 2009. Constructing Repertorium 3.0. *Scripta & e-Scripta* 7. 13–48.
- Bos, Bert, Håkon Wium Lie, Chris Lilley, Ian Jacobs (eds.). 2008. *Cascading Style Sheets, Level 2 CSS2 Specification*. W3C Recommendation, 11 April 2008. <http://www.w3.org/TR/2008/REC-CSS2-20080411/> (15 July 2010).
- Bray, Tim, Dave Hollander, Andrew Laymon, Richard Tobin (eds.). 2006a: *Namespaces in XML 1.0 (Second edition)*, W3C Recommendation, 16 August 2006. <http://www.w3.org/TR/xml-names/> (15 July 2010).
- Bray, Tim, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler & François Yergau (eds.). 2008: *Extensible Markup Language (XML) Version 1.0* (Fifth edition), W3C Recommendation, 26 November 2008. <http://www.w3.org/TR/REC-xml/> (12.02.2012).
- Burnard, Lou & Syd Bauman (eds.). 2007: *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. <http://www.tei-c.org/Guidelines/P5/> (12.02.2012).
- Burnard, Lou, Syd Bauman & Steven DeRose (eds.) 2001: *TEI P4: Guidelines for Electronic Text Encoding and Interchange*. <http://www.tei-c.org/Guidelines/P4/index.xml> (12.02.2012).
- Burnard, Lou, Richard Gartner & Peter Kidd. 1997. The Cataloguing of Western Medieval Manuscripts in Bodleian Library: a TEI approach with appendix describing a TEI-conformant manuscript description. <http://www.users.ac.uk/~lou/wip/MS/> (12.10.2010).
- Clark, James (ed.). 1999: XSL Transformations (XSLT) Version 1.0, W3C Recommendation 16 November 1999. <http://www.w3.org/TR/1999/REC-xslt-19991116/> (12.02.2012).
- Clark, James (ed.). 2002: Relax NG Compact Syntax. OASIS Committee Specification 21 November 2002. <http://www.relaxng.org/compact-20021121.html> (12.02.2012).
- Clark, James & Steve DeRose (eds.). 1999. XML Path Language (XPath) Version 1.0.

- W3C Recommendation 16 November 1999. <http://www.w3.org/TR/1999/REC-xpath-19991116> (12.02.2012).
- Clark, James & Murata Makomoto (eds.). 2001. Relax NG Specification. OASIS Committee Specification 3 December 2001 <http://www.relaxng.org/spec-20011203.html> (12.02.2012).
- Cleminson, Ralph. 1998. Computer Applications and the Study of Early Cyrillic Printed Books. *Palaeobulgarica* 22 (2). 104–115.
- Cleminson, Ralph. 2003. A Good Servant but a Bad MASTER: Uses and Abuses of Standards in Manuscript Description. In A. Miltenova, D. J. Birnbaum & S. Slevinsky (eds.), *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings of the „Electronic Description and Edition of Slavic Sources” conference, 24-26 September 2002, Pomorie, Bulgaria*. Sofia: Boyan Penev Publishing Center, Institute of Literature. 105-111.
- Common Data Hub: ISO 3166 Country Codes: [http://www.commondatahub.com/live/geography/country/iso\\_3166\\_country\\_codes](http://www.commondatahub.com/live/geography/country/iso_3166_country_codes) (12.02.2012).
- DeRose, Steven, Ron Daniel Jr., Eve Maler, & Jonathan Marsh (eds.). 2003. Xpointer xmlns() Scheme. W3C Recommendation 25 March 2003 <http://www.w3.org/TR/2003/REC-xptr-xmlns-20030325/> (12.02.2012).
- DeRose, Steven, Eve Maler & David Orchard (eds.). 2001. XML Linking Language (XLink) Version 1.0. W3C Recommendation 27 June 2001. <http://www.w3.org/TR/2000/REC-xlink-20010627/> (12.02.2012).
- Dobрева, Milena. 2000. A repertory of the old Bulgarian literature: Problems concerning the design and use of a computer supported model. In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov. 91–98.
- Dutschke, C. W., Mirella Ferrari & R. H. Rouse (eds.). 1987. *Medieval and Renaissance Manuscripts in the Claremont Libraries*. University of California Press.
- Gagova, Nina. 1998. TSM: The element NOTE: Improvement of the encoding possibilities: a philological proposal. In Milena Dobрева (ed.), *Text Variety in the Witnesses of Medieval Texts*. Sofia: Institute of Mathematics and Informatics. 161-164.
- Gagova, Nina. 2000. Study of groups of manuscripts chosen by socio-cultural criteria (Manuscripts belonging to rulers libraries from the fourteenth and fifteenth centuries). In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov. 131–169.
- Gillisen, L. 1977. *Prolégomènes à la codicologie. Recherches sur la construction des cahiers et la mise en page des manuscrits médiévaux*. Gand.
- Grosso, Paul, Eve Maler, Jonathan Marsh & Norman Walsh (eds.). 2003a. XPointer Framework, W3C Recommendation 25 March 2003. <http://www.w3.org/TR/xptr-framework/> (12.02.2012).
- Grosso, Paul, Eve Maler, Jonathan Marsh & Norman Walsh (eds.). 2003b. XPointer element() Scheme, W3C Recommendation 25 March 2003. <http://www.w3.org/TR/xptr-element/> (12.02.2012).
- ISO 10646: International Organization for Standardization, “ISO/IEC 10646:2003. Information technology – Universal Multiple-Octet Coded Character Set (UCS)”, 2003.



*Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions*

- ISO 15924: International Organization for Standardization, “ISO 15924:2004. Information and documentation -- Codes for the representation of names of scripts”, January 2004.
- ISO 3166-1: International Organization for Standardization, “ISO 3166-1:1997. Codes for the representation of names of countries and their subdivisions -- Part 1: Country codes”, 1997.
- ISO 639-1: International Organization for Standardization, “ISO 639-1:2002. Codes for the representation of names of languages -- Part 1: Alpha-2 code”, 2002.
- ISO 639-2. International Organization for Standardization, “ISO 639-2:1998. Codes for the representation of names of languages -- Part 2: Alpha-3 code, first edition”, 1998.
- Jones, L. W. 1946a. Pricking Systems in New York Manuscripts. *Miscellanea G. Mercati*, vol. VI (*Studi e Testi*, 126). Città del Vaticano. 80–92.
- Jones, L. W. 1946b. Pricking Manuscripts: The Instruments and Their Significance. *Speculum* 21. 389-403.
- Krushelnitskaya, Ekaterina V. 2003. Description of Medieval Slavic Sources in the Electronic Catalog of the Department of Manuscripts in the National Library of Russia: Goals, Research Problems, and Prospects. In A. Miltenova, D. Birnbaum & S. Slevinski (eds.), *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings of the „Electronic Description and Edition of Slavic Sources” conference, 24-26 September 2002, Pomorie, Bulgaria*. Sofia: Boyan Penev Publishing Center, Institute of Literature. 157–188.
- Leroy, J. 1976. *Les types de réglure des manuscrits grecs*. Paris: Institut de recherche et d’histoire des textes (Bibliographie-Colloques-Travaux préparatoires).
- Marsh, Jonathan & Richard Tobin (eds.). 2009. *XML Base (Second Edition)*. W3C Recommendation, 28 January 2009. <http://www.w3.org/TR/2009/REC-xmlbase-20090128/> (12.02.2012).
- Miklas, Heinz & Anissava Miltenova (eds.). 2008. *Slovo: Towards a Digital Library of South Slavic Manuscripts*. Proceedings of the International Conference 21-26 February 2008, Sofia, Bulgaria. Sofia: Bulgarian Academy of Sciences, Institute of literature.
- Miltenov, Yavor. 2010. A New Source for Studying the Symeonian Zlatostruy Collection. *Scripta & e-Scripta* 8-9. 387–404.
- Miltenova, Anissava. 2008. Metadata and Electronic Catalogs of Slavic Manuscripts: A Multilingual Web-based Terminological Resource for Mediaeval Slavic Studies. *Scripta & e-Scripta* 6. 83–104.
- Miltenova, Anissava, A. Boyadzhiev & St. Velev 1998. Repertorium of Medieval Slavic Literature: Computer and Philological Standards. *Palaeobulgarica* 22 (2). 50-69.
- Miltenova, Anissava & Andrey Boyadzhiev. 2000. An Electronic Repertory of Medieval Slavic Literature and Letters: A Suite of Guidelines. In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov. 43-68.
- Miltenova, Anissava, David Birnbaum & Sarah Slevinski (eds.). 2003. *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings of the „Electronic Description and Edition of Slavic Sources” conference, 24-26 September 2002, Pomorie, Bulgaria*. Sofia: Boyan Penev Publishing Center, Institute of Literature.
- Moussakova, Elissaveta. 2003. The Element “Decoration” in Electronic Manuscript Catalogs.

- In A. Miltenova, D. J. Birnbaum & S. Slevinski (eds), *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. Proceedings of the „Electronic Description and Edition of Slavic Sources” conference, 24-26 September 2002, Pomorie, Bulgaria*. Sofia: Boyan Penev Publishing Center, Institute of Literature. 137–156.
- PDF Reference: PDF Reference Sixth Edition. Adobe Portable Document Format 1.7 November 2006. Adobe Systems Incorporated. [http://www.adobe.com/devnet/acrobat/pdfs/pdf\\_reference\\_1-7.pdf/](http://www.adobe.com/devnet/acrobat/pdfs/pdf_reference_1-7.pdf/) (12.02.2012).
- Peterson, David, Shudi (Sandy) Gao, Ashok Malhotra, C. M. Sperberg-McQueen & Henry S Thomson (eds.). 2009: W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes. W3C Working Draft 30 January 2009. <http://www.w3.org/TR/xmlschema11-2/> (12.02.2012).
- Philips, A., Davis, M. (eds) 2006c: *BCP 47*, which consists of Philips, Davis 2006a and Philips Davis 2006b
- Philips, A., Davis, M. (eds). 2006a: Tags for Identifying Languages, *BCP 47, RFC 4646*, September 2006. <http://tools.ietf.org/html/rfc4646> (12.02.2012)
- Philips, A., Davis, M. (eds). 2006b: Matching of Language Tags, *BCP 47, RFC 4647*, September 2006. <http://tools.ietf.org/html/rfc4647> (12.02.2012)
- Philips, A., M. Davis (eds). 2006c: *BCP 47* (RFC 5646), which consists of Philips, Davis 2006a and Philips Davis 2006b. <http://tools.ietf.org/html/bcp47> (12.02.2012).
- Radoslavova, Dilyana. 1998. SGML Tagging of Complex Texts in Old Slavic Liturgical Manuscripts. In Milena Dobрева (ed.), *Text Variety in the Witnesses of Medieval Texts*. Sofia: Institute of Mathematics and Informatics. 164–166.
- Radoslavova, Dilyana. 2000. Problems Concerning the Computer Description and Processing of Slavic Manuscripts of Euchological Type with the Help of TSM (A typology of Bulgarian *trebnici* in the 17<sup>th</sup> century. Preliminary results). In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov. 170-204.
- Szirmai, J. A. 1999. *The Archaeology of Medieval Bookbinding*. Ashgate.
- Shudi (Sandy) Gao, C. M. Sperberg-McQueen & Henry S. Thompson (eds.). 2012. W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures. W3C Proposed Recommendation 19 January 2012. <http://www.w3.org/TR/xmlschema11-1/> (12.02.2012).
- Sperberg-McQueen, C. M. and Burnard, Lou (eds.). 1994: *TEI P3: Guidelines for Electronic Text Encoding and Interchange*. <http://www.tei-c.org/Vault/GL/P3/index.htm> (12.02.2012).
- Stojkova, Ana. 1998. Variety and Standardization of Old Slavonic Texts (Problems of Creating a Data Base). In Milena Dobрева (ed.), *Text Variety in the Witnesses of Medieval Texts*. Sofia: Institute of Mathematics and Informatics. 147–153.
- Stojkova, Ana. 2000. Problems of the Composition of Late Medieval Non-Liturgical Miscellanies and the Place of Hagiographic Works in them. In A. Miltenova & D. J. Birnbaum (eds.), *Medieval Slavic Manuscripts and SGML. Problems and Perspectives*. Sofia: Marin Drinov. 108-130.
- Thomson, Francis. 1982. Chrysostomica Palaeoslavica. A Preliminary Study of the Sources of the Chrysorrhoeas (Zlatorstui) Collection. *Cyrrillomethodianum* 6. 1-65.
- Unicode 5.1. *The Unicode Standard, Version 5.1*, Unicode Consortium. 2008. Addison-Wesley



- Professional. <http://www.unicode.org/> (15 July 2010).
- Walsh, Norman (ed.). 2002. XML Catalogs. OASIS Committee Specification 06 August 2001. <http://www.oasis-open.org/committees/entity/spec-2001-08-06.html/> (12.02.2012).
- Александров, А. 1893. *Физиолог*. Казань.
- Дринов, Марин. 1882. Български летописен разказ от края на XVII в. *Периодично списание* 3. 8–12
- Милтенев, Явор. 2011. Славянският превод слова Иоанна Златоуста De sancto hieromartyre Phoca. *Vyzantinoslavica* 59 (1–2). 311–368.
- Райков, Божидар, Ст. Кожухаров, Х. Миклас и Хр. Кодов. 1994. *Каталог на славянските ръкописи в библиотеката на Зографския манастир в Света гора*. София: CIBAL.
- Стойкова, Ана. 1994. *Физиологът в южнославянските литератури*. София: БАН.
- Стойкова, Ана. 2011. *Славянският Физиолог. Византийска редакция: Електронно издание на текста и сравнително изследване*. <http://physiologus.proab.info/> (12.02.2012).
- Цонев, Беньо. 1910. *Опис на ръкописите и старопечатните книги на Народната библиотека в София*. София: Народна библиотека.

## **Appendices**

### *Appendix 1. Renamed Repertorium 2.1 Elements and their TEI P5 and Repertorium 3.0 correspondences*

#### *Repertorium 2.1*

#### *TEI P5 and Repertorium 3.0*

---

|                                  |                                 |
|----------------------------------|---------------------------------|
| <articleContentDesc>             | <msItemStruct>                  |
| <articleName>                    | <title>                         |
| <articleAuthor>                  | <author>                        |
| <translation>                    | <filiation type="translation">  |
| <source>                         | <filiation type="source">       |
| <protograph>                     | <filiation type="protograph">   |
| <antigraph>                      | <filiation type="antigraph">    |
| <apograph>                       | <filiation type="apograph">     |
| <litRedaction>                   | <filiation type="litRedaction"> |
| <articleData>                    | <note>                          |
| <articleData<br>type="language"> | <textLang>                      |
| <miscObservat>                   | <condition>                     |
| <catalogueStmnt>                 | <msIdentifier>                  |
| <manuscriptName>                 | <msName>                        |
| <repositCountry>                 | <country>                       |
| <repositCity>                    | <settlement>                    |
| <repositSignature>               | <idno type="shelfmark">         |
| <catalogNr>                      | <idno type="catalogue">         |
| <codicology>                     | <physDesc>                      |
| <dim>                            | <dimensions>                    |
| <manuscriptCreation>             | <origin>                        |

*Repertorium 2.1*

*TEI P5 and Repertorium 3.0*

---

|                    |   |
|--------------------|---|
| <manuscriptDate>   | <origDate>  |
| <manuscriptPlace>  | <origPlace>   |
| <contentDesc>      | <msContents>  |
| <repro>            | <surrogates>  |
| <numFolio>         | <extent>  |
| <numPage>          | <extent>  |
| <quireStructure>   | <collation>   |
| <otherDecoration>  | <decoNote>  |
| <pageRange>        | <locus>, <locusGrp>                                   |
| <markings>         | <catchwords>, <signatures>                            |
| <decoration>       | <decoDesc>  |
| <churchCal>        | <date> in the content of <note>,<br>cf. <articleData> |
| <noteFormula>      | <formula>   |
| <noteAuthor>       | <author>  |
| <noteContent>      | <summary>   |
| <noteData>         | <note>  |
| <location>         | <locus>, <locusGrp>                                   |
| <locNumber>        | <locus n="...">                                       |
| <position>         | <locus>   |
| <compositionQuire> | <formula>   |
| <miniatures>       | <decoNote type="miniatures">                          |
| <headpieces>       | <decoNote type="headpieces">                          |
| <borders>          | <decoNote type="borders">                             |
| <tailpieces>       | <decoNote type="tailpieces">                          |
| <wash>             | <decoNote type="wash">                                |
| <linefillers>      | <decoNote type="linefillers">                         |
| <initials>         | <decoNote type="initials">                            |
| <sup>              | <c>, <phr>, <w> or <seg> with<br>attribute rend="sup" |
| <ligature>         | <c> or <g>  |
| <watermarkList>    | series of <watermark> elements                        |
| <writing>          | <summary>   |

*Appendix 2. Renamed Repertorium elements without TEI P5  
correspondence*

| <i>Repertorium 2.1</i> | <i>Repertorium 3.0</i> | <i>TEI P5 possible mapping</i> |
|------------------------|------------------------|--------------------------------|
| <orthogrCharact>       | <orthography>          | #PCDATA in <textLang>          |
| <letterForm>           | <signForm>             | #PCDATA as part of <p> in      |
| <listScribe>           | <scribeDesc>           | <handNote><br><handDesc> (?)   |

*Appendix 3. Unique Repertorium elements*

| <i>Repertorium 3.0</i> | <i>TEI P5 possible mapping</i>                              |
|------------------------|---|
| <bindingNote>          | <p> in <binding>  |
| <countermark>          | #PCDATA in <support>  |
| <crypto>               | #PCDATA in <handNote>                                       |
| <grammar>              | #PCDATA in <textLang>                                       |
| <gregoryRule>          | #PCDATA in <layout>   |
| <lexis>                | #PCDATA in <textLang>                                       |
| <motif>                | #PCDATA in <support>  |
| <orthography>          | #PCDATA in <textLang>                                       |
| <pagination>           | <foliation> ?   |
| <palaeoNote>           | #PCDATA or <c> + explanation in<br><handNote><br><handDesc> |
| <palaeoCharact>        | <handNote><br><handDesc>                                    |
| <quire>                | #PCDATA in <collation> +<br><formula>                       |
| <ruling>               | #PCDATA in <layout>   |
| <sampleText>           | -   |
| <scribe>               | -   |
| <scribeLang>           | <textLang> in <msContents>                                  |

*Appendix 4. Obsolete Repertorium Elements*

<antigraph>, <apograph>, <articleAuthor>, <articleContentDesc>, <articleData>, <articleName>, <borders>, <cadels>, <catalogNr>, <catalogueStmnt>, <churchCal>, <codicology>, <compositionQuire>, <contentDesc>, <decoration>, <disposition>, <headpieces>, <initials>, <ligature>, <linefillers>, <litRedaction>, <location>, <locNumber>, <manuscriptContentDesc>, <manuscriptCreation>, <manuscriptDate>, <manuscriptName>, <manuscriptPlace>, <markings>, <miniatures>, <miscObservat>, <neighbour>, <noteAuthor>, <noteContent>, <noteData>, <noteFormula>, <numbColumn>, <numberTexts>, <numbLines>, <numFolio>, <numPage>, <otherDecoration>, <overview>, <pageRange>, <paperList>, <parchmentList>, <position>, <protograph>, <quireStructure>, <repositCity>, <repositCountry>, <repositSignature>, <repro>, <scribeList>, <source>, <sup>, <tailpieces>, <translation>, <wash>.

*Appendix 5. TEI modules used in the Repertorium and their changes*

```
<moduleRef key="header" except="appInfo application
biblFull handNote scriptNote typeNote" />
<moduleRef key="linking" except="alt altGrp join joinGrp
link linkGrp" />
```

## A. Bojadžiev

```
<moduleRef key="core" except="postBox postCode teiCorpus"
/>
<moduleRef key="tei" />
<moduleRef key="textstructure" except="argument byline
closer dateline div1 div2 div3 div4 div5 div6 div7 opener
postscript salute signed" />
<!-- optional modules -->
<moduleRef key="analysis" />
<moduleRef key="certainty" />
<moduleRef key="figures" />
<moduleRef key="msdescription" except="handDesc msItem
scriptDesc typeDesc" />
<moduleRef key="namesdates" except="climate terrain" />
<moduleRef key="transcr" except="handNotes" />
```

### Appendix 6. Deleted TEI models

```
<classSpec ident="model.physDescPart" type="model"
mode="delete" module="msdescription" />
```

The Repertorium uses a different model for physical description

### Appendix 7. Changed TEI element declarations

*Element additions*

```
<elementSpec ident="additions" module="msdescription"
mode="change">
<classes mode="change">
<memberOf key="att.typed" mode="add" />
<memberOf key="att.dateable" mode="add" />
</classes>
</elementSpec>
```

Allows inclusion of typology and attributes for formalizing dateable additional parts, e.g., @when, @to, @from, etc.

*Element altIdentifier*

```
<elementSpec ident="altIdentifier" module="msdescription"
mode="change">
<content>
<rng:zeroOrMore><rng:ref name="msName" /></rng:zeroOrMore>
<rng:zeroOrMore>
<rng:zeroOrMore>
<rng:ref name="model.placeNamePart_sequenceOptional"/>
</rng:zeroOrMore>
<rng:optional><rng:ref name="institution" />
</rng:optional>
<rng:optional><rng:ref name="repository" /></rng:optional>
```

```
<rng:zeroOrMore><rng:ref name="collection"
/></rng:zeroOrMore>
<rng:optional><rng:ref name="idno" />
</rng:optional>
<rng:optional><rng:ref name="note" /></rng:optional>
<rng:optional><rng:ref name="altIdentifier" />
</rng:optional></rng:zeroOrMore></content>
</elementSpec>
```

1. Includes `<msName>` to allow naming a separate sections or parts of the manuscript;
2. Makes `<altIdentifier>` a nested element, thus allowing designation of collections and shelfmarks inside other collections, or inclusion of alternative identifiers (e.g., former) for archives with composite structure.

*Element binding*

```
<elementSpec ident="binding" module="msdescription"
              mode="change">
<content>
<rng:oneOrMore>
<rng:choice>
<rng:ref name="model.pLike" />
<rng:ref name="bindingNote" />
<rng:ref name="condition" />
<rng:ref name="decoNote" />
</rng:choice>
</rng:oneOrMore>
</content>
</elementSpec>
```

Includes element `<bindingNote>` with specialized features not necessarily related to decorative parts of the binding.

*Element collation*

```
<elementSpec ident="collation" module="msdescription"
              mode="change">
<content>
<rng:optional><rng:ref name="p" /></rng:optional>
<rng:zeroOrMore><rng:ref name="quire" /></rng:zeroOrMore>
<rng:optional><rng:ref name="signatures" /></rng:optional>
<rng:optional><rng:ref name="catchwords" /></rng:optional>
<rng:optional><rng:ref name="gregoryRule" /></rng:optional>
<rng:optional><rng:ref name="ruling" /></rng:optional>
<rng:optional><rng:ref name="pricking" />
</rng:optional></content>
</elementSpec>
```

Instead of unstructured elements on paragraph level provides several more specialized sections, such as `<quire>` for description of each quire, `<gregoryRule>`, `<ruling>` and `<pricking>`.

*Element history*

```
<elementSpec ident="history" module="msdescription"
              mode="change">
  <content>
    <rng:choice>
      <rng:oneOrMore><rng:ref name="model.pLike" />
    </rng:oneOrMore>
    <rng:group>
      <rng:optional><rng:ref name="summary" /></rng:optional>
      <rng:optional><rng:ref name="origin" /></rng:optional>
      <rng:optional><rng:ref name="accMat" /></rng:optional>
      <rng:zeroOrMore><rng:ref name="provenance" />
    </rng:zeroOrMore>
      <rng:optional><rng:ref name="acquisition" /></rng:optional>
    </rng:group></rng:choice>
  </content>
</elementSpec>
```

Includes element `<accMat>` in this place of the schema and removes it from the `<phys-Desc>` element. `<accMat>` (accompanying material) as described in the TEI P5 Guidelines is related with the manuscript history after its writing, unlike the element `<additions>`.

*Element layout*

```
<elementSpec ident="layout" module="msdescription"
              mode="change">
  <classes mode="change">
    <memberOf key="att.measurement" mode="add" />
  </classes>
</elementSpec>
```

Adds an attribute class `measurement`. This allows the usage of attributes `@unit`, `@quantity` and `@commodity`.

*Element material*

```
<elementSpec ident="material" module="msdescription"
              mode="change">
  <classes mode="change">
    <memberOf key="att.global" />
    <memberOf key="att.global.linking" />
    <memberOf key="att.global.analytic" />
    <memberOf key="att.global.facs" />
    <memberOf key="att.global.change" />
    <memberOf key="att.canonical" />
  </classes>
  <attList><attDef ident="usage">
    <desc>describes the usage of the material</desc>
    <datatype><rng:text />
  </attDef>
</attList>
```

## Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions

```
</attDef></attList>
```

```
</elementSpec>
```

Adds attribute @usage to allow a short characterization of material.

*Element msContents*

```
<elementSpec ident="msContents" module="msdescription"  
              mode="change">
```

```
<content>
```

```
<rng:choice>
```

```
<rng:oneOrMore><rng:ref name="model.pLike" />
```

```
</rng:oneOrMore>
```

```
<rng:group>
```

```
<rng:group><rng:optional>
```

```
<rng:ref name="summary" />
```

```
</rng:optional>
```

```
<rng:optional>
```

```
<rng:ref name="titlePage" />
```

```
</rng:optional>
```

```
<rng:optional><rng:ref name="title" /></rng:optional>
```

```
<rng:zeroOrMore><rng:ref name="author" /></rng:zeroOrMore>
```

```
<rng:zeroOrMore><rng:ref name="respStmt" />
```

```
</rng:zeroOrMore>
```

```
<rng:zeroOrMore><rng:ref name="model.noteLike" />
```

```
</rng:zeroOrMore>
```

```
<rng:zeroOrMore><rng:ref name="filiation" />
```

```
</rng:zeroOrMore>
```

```
<rng:optional>
```

```
<rng:ref name="textLang" /></rng:optional><rng:optional>
```

```
<rng:ref name="decoNote" /></rng:optional>
```

```
<rng:optional>
```

```
<rng:ref name="sampleText" />
```

```
</rng:optional>
```

```
<rng:optional>
```

```
<rng:ref name="div" />
```

```
</rng:optional>
```

```
</rng:group><rng:zeroOrMore>
```

```
<rng:ref name="msItemStruct" />
```

```
</rng:zeroOrMore></rng:group>
```

```
</rng:choice>
```

```
</content>
```

```
<attList>
```

```
<attDef ident="type">
```

```
<desc>describes the type of the text</desc>
```

```
<datatype><rng:text />
```

```
</datatype><valList type="closed">
```

```

<valItem ident="original"><desc>the text is
original</desc></valItem>
<valItem ident="compilation">
<desc>the text is compiled from variuos sources</desc>
</valItem>
<valItem ident="translation">
<desc>the text is translated</desc>
</valItem></valList></attDef>
<attDef ident="style">
<desc>describes the style of the text</desc>
<datatype><rng:text /></datatype>
<valList type="closed">
<valItem ident="narrative">
<desc>the text is narrative</desc></valItem>
<valItem ident="non-narrative">
<desc>the text is non-narrative, e.g., hymnography</desc>
</valItem></valList></attDef>
</attList></elementSpec>

```

1. Changes the order of elements in some places to reflect the tradition in describing Slavonic manuscripts, e.g., <title> comes before <author>; 2. Includes <respStmt> for situations where these are need for clarification; 3. Includes <sampleText> to distinguish original text from the description; 4. Includes <div> element for the edition of rare or unknown texts; 5. Provides two more attributes (@type and @style) for compatibility with the old Repertorium schema.

*Element msDesc*

```

<elementSpec ident="msDesc" module="msdescription"
mode="change">
<content>
<rng:zeroOrMore><rng:ref name="model.headLike" />
</rng:zeroOrMore>
<rng:ref name="msIdentifier" />
<rng:choice><rng:zeroOrMore><rng:ref name="model.pLike" />
</rng:zeroOrMore>
<rng:group>
<rng:optional><rng:ref name="physDesc" /></rng:optional>
<rng:optional><rng:ref name="scribeDesc" /></rng:optional>
<rng:optional><rng:ref name="msContents" /></rng:optional>
<rng:optional><rng:ref name="history" /></rng:optional>
<rng:optional><rng:ref name="additional" /></rng:optional>
<rng:zeroOrMore><rng:ref name="msPart" /></rng:zeroOrMore>
</rng:group></rng:choice>
</content>
</elementSpec>

```

Includes one more level of description for palaeographical and language features



(<scribeDesc>). Moves model.headLike to the first position to serve as a heading of the description.

*Element msIdentifier*

```
<elementSpec ident="msIdentifier"
  module="msdescription"
              mode="change">
  <content>
  <rng:zeroOrMore><rng:ref name="msName" />
  </rng:zeroOrMore>
  <rng:zeroOrMore>
  <rng:zeroOrMore><rng:ref name="p" /></rng:zeroOrMore>
  <rng:zeroOrMore>
  <rng:ref name="model.placeNamePart_sequenceOptional"/>
  </rng:zeroOrMore>
  <rng:optional><rng:ref name="institution" />
  </rng:optional>
  <rng:optional><rng:ref name="repository" />
  </rng:optional>
  <rng:zeroOrMore><rng:ref name="collection" />
  </rng:zeroOrMore>
  <rng:optional><rng:ref name="idno" /></rng:optional>
  <rng:optional><rng:ref name="note" />
  </rng:optional></rng:zeroOrMore>
  <rng:zeroOrMore><rng:ref name="altIdentifier" />
  </rng:zeroOrMore>
  </content>
</elementSpec>
```

1. Moving the element <msName> to the first place in the content; 2. Adding <note> to allow more specific information to appear here, e.g., bibliographic references.

*Element msItemStruct*

```
<elementSpec ident="msItemStruct" module="msdescription"
              mode="change">
  <content><rng:group>
  <rng:choice><rng:optional><rng:ref name="locus" />
  </rng:optional>
  <rng:optional><rng:ref name="locusGrp" />
  </rng:optional></rng:choice>
  <rng:choice><rng:oneOrMore><rng:ref name="model.pLike" />
  </rng:oneOrMore>
  <rng:group><rng:zeroOrMore>
  <rng:ref name="title" /></rng:zeroOrMore>
  <rng:zeroOrMore><rng:ref name="author" /></rng:zeroOrMore>
  <rng:zeroOrMore><rng:ref name="respStmt" />
  </rng:zeroOrMore>
```

## A. Bojadžiev

```
<rng:zeroOrMore><rng:ref name="model.noteLike" />
</rng:zeroOrMore>
<rng:zeroOrMore><rng:ref name="filiation" />
</rng:zeroOrMore>
<rng:optional><rng:ref name="textLang" /></rng:optional>
<rng:zeroOrMore><rng:ref name="decoNote" />
</rng:zeroOrMore>
<rng:zeroOrMore><rng:ref name="colophon" />
</rng:zeroOrMore>
<rng:optional><rng:choice><rng:ref name="sampleText" />
<rng:ref name="div" />
</rng:choice></rng:optional>
<rng:zeroOrMore><rng:ref name="msItemStruct" />
</rng:zeroOrMore>
<rng:zeroOrMore><rng:ref name="listBibl" />
</rng:zeroOrMore>
<rng:zeroOrMore><rng:ref name="bibl" />
</rng:zeroOrMore></rng:group>
</rng:choice>
</content>
<attList><attDef ident="type"><desc>describes the type of
the text</desc>
<datatype><rng:text /></datatype><valList type="closed">
<valItem ident="original"><desc>the text is
original</desc></valItem>
<valItem ident="compilation"><desc>the text is compiled
from variuos
sources</desc></valItem><valItem
ident="translation"><desc>the text is
translated</desc></valItem></valList></attDef>
<attDef ident="style"><desc>describes the style of the
text</desc><datatype>
<rng:text /></datatype><valList type="closed"><valItem
ident="narrative">
<desc>the text is narrative</desc></valItem>
<valItem ident="non-narrative">
<desc>the text is non-narrative, e.g.,
hymnography</desc></valItem></valList>
</attDef></attList> </elementSpec>
```

1. The same changes as for the element <msContents>. 2. Adds the element <locusGrp>.

### *Element msPart*

```
<elementSpec ident="msPart" module="msdescription"
mode="change">
```

## Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions

```
<classes mode="change"><memberOf key="att.typed" mode="add" />
</classes>
<content>
<rng:group><rng:group><rng:choice><rng:zeroOrMore>
<rng:optional><rng:ref name="altIdentifier" />
</rng:optional>
<rng:optional><rng:ref name="msIdentifier" /> </rng:optional>
<rng:optional><rng:ref name="locus" /></rng:optional>
<rng:optional><rng:ref name="locusGrp" /></rng:optional>
</rng:zeroOrMore></rng:choice>
</rng:group><rng:zeroOrMore><rng:ref name="model.headLike"
/></rng:zeroOrMore><rng:choice>
<rng:oneOrMore><rng:ref name="model.pLike" />
</rng:oneOrMore>
<rng:group><rng:optional><rng:ref name="physDesc" />
</rng:optional><rng:optional><rng:ref name="scribeDesc" />
</rng:optional>
<rng:optional><rng:ref name="msContents" /></rng:optional>
<rng:optional><rng:ref name="history" /></rng:optional>
<rng:optional><rng:ref name="additional" /></rng:optional>
<rng:zeroOrMore><rng:ref name="msPart" /></rng:zeroOrMore>
</rng:group></rng:choice></rng:group></content>
</elementSpec>
```

1. Reflects the same changes as <msDesc> element; 2. Adds the <locus> and <locusGrp> elements for parts of the manuscript that are palimpsests or convolutes; 3. Adds the special attribute @type.

### *Element musicNotation*

```
<elementSpec ident="musicNotation" module="msdescription"
mode="change">
<classes mode="replace"><memberOf key="att.global" />
<memberOf key="att.typed" mode="add" /></classes>
<content><rng:ref name="macro.specialPara" /></content>
</elementSpec>
```

Adds attributes @type and @subtype to allow typologization of the notation.

### *Element physDesc*

```
<elementSpec ident="physDesc" mode="change"
module="msdescription">
<content>
<rng:group>
<rng:zeroOrMore><rng:ref name="model.pLike" />
</rng:zeroOrMore>
<rng:group>
<rng:optional><rng:ref name="objectDesc" />
</rng:optional>
```

```

<rng:optional><rng:ref name="ink" /></rng:optional>
<rng:optional><rng:ref name="decoDesc" />
</rng:optional>
<rng:optional>
<rng:ref name="bindingDesc" /></rng:optional>
<rng:optional><rng:ref name="sealDesc" /></rng:optional>
<rng:zeroOrMore><rng:ref name="additions" />
</rng:zeroOrMore></rng:group></rng:group></content>
</elementSpec>

```

1. Element <accMat> is moved to history; 2. Elements for the description of palaeography are removed and a new model in <scribeDesc> has being used. 3. The element <ink> is added.

*Element supportDesc*

```

<elementSpec ident="supportDesc" module="msdescription"
              mode="change">
<content><rng:choice><rng:oneOrMore><rng:ref name="model.
pLike" /></rng:oneOrMore><rng:group>
<rng:optional><rng:ref name="support" />
</rng:optional>
<rng:optional><rng:ref name="extent" />
</rng:optional><rng:zeroOrMore><rng:choice>
<rng:ref name="foliation" /><rng:ref name="pagination" />
</rng:choice></rng:zeroOrMore>
<rng:optional><rng:ref name="collation" /></rng:optional>
<rng:optional><rng:ref name="condition" /></rng:optional>
</rng:group></rng:choice>
</content></elementSpec>

```

The element <pagination> is added.

*Element watermark*

```

<elementSpec ident="watermark" module="msdescription"
              mode="change">
<classes mode="change">
<memberOf key="att.global" />
<memberOf key="att.global.linking" />
<memberOf key="att.global.analytic" />
<memberOf key="att.global.facs" />
<memberOf key="att.global.change" />
<memberOf key="att.canonical" /></classes>
<content>
<rng:zeroOrMore><rng:choice>
<rng:text /><rng:ref name="model.phrase" />
<rng:ref name="countermark" />
<rng:ref name="motif" /></rng:choice></rng:zeroOrMore>
</content>

```

## Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions

```
<attList>
<attDef ident="tracing">
<desc>describes the methods of recording of some watermark.
Methods include: tracing, rubbing, etc. </desc><datatype>
<rng:text /></datatype></attDef></attList></elementSpec>
```

1. Adds attribute class <canonical> with to allow usage of @key and @ref. 2. Adds new attribute @tracing to characterize the way the watermark has been recorded; 3. Adds elements <countermark> and <motif>.

### Appendix 8. New elements

The Repertorium Initiative model introduces the following new elements in the TEI schema.

#### Element bindingNote

```
<elementSpec ident="bindingNote" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>The element allows structured description of
binding's characteristics in addition of decoration
and condition. These features could apply but are
not limited to the binding's parts: cover, plates,
techniques. </desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" />
<memberOf key="att.damaged" /></classes>
<content><rng:zeroOrMore><rng:choice><rng:text />
<rng:ref name="model.phrase" /><rng:ref name="decoNote" />
<rng:ref name="condition" /></rng:choice>
</rng:zeroOrMore></content></elementSpec>
```

#### Element countermark

```
<elementSpec ident="countermark" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
  mode="add">
<desc>Called also counter watermark. Allows a separate
description of countermark as part pf watermark element.
</desc>
<classes mode="change">
<memberOf key="att.global" /><memberOf key="att.typed" />
<memberOf key="att.damaged" /></classes>
<content>
<rng:zeroOrMore><rng:choice><rng:text />
<rng:ref name="model.phrase" /><rng:ref name="motif" />
</rng:choice></rng:zeroOrMore>
</content></elementSpec>
```

*Element crypto*

```
<elementSpec ident="crypto" ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Provides opportunity to describe the cryptographical
system(s) in the MS</desc>
<classes mode="change"><memberOf key="att.global" />
  <memberOf key="att.typed" /></classes><content>
<rng:ref name="macro.paraContent" />
</content></elementSpec>
```

*Element grammar*

```
<elementSpec ident="grammar" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
mode="add">
<desc>Describes grammatical features of the scribe's
language</desc>
<classes mode="replace"><memberOf key="att.global" />
</classes>
<content><rng:ref name="macro.specialPara" /></content>
</elementSpec>
```

*Element gregoryRule*

```
<elementSpec ident="gregoryRule" module="msdescription"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
prefix="re" mode="add">
<desc>Description of order of parchment's leaves in the
quires in terms of hair- or flesh-side arrangement. </desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" />
</classes><content><rng:ref name="macro.phraseSeq" />
</content></elementSpec>
```

*Element ink*

```
<elementSpec ident="ink"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Description of ink(s)</desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.measurement" /></classes>
<content><rng:ref name="macro.phraseSeq" /></content>
</elementSpec>
```

*Element lexis*

```
<elementSpec ident="lexis" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Description lexical features. </desc>
```

## *Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions*

```
<classes mode="replace"><memberOf key="att.global" />
</classes>
```

```
<content><rng:ref name="macro.specialPara" />
</content></elementSpec>
```

### *Element motif*

```
<elementSpec ident="motif" module="msdescription"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
prefix="re" mode="add">
```

```
<desc>Description of water- or countermark motif. </desc>
```

```
<classes mode="change"><memberOf key="att.global" />
```

```
<memberOf key="att.typed" /></classes>
```

```
<content>
```

```
<rng:zeroOrMore><rng:choice><rng:text /><rng:ref
name="model.phrase" /><rng:ref name="motif" />
```

```
</rng:choice></rng:zeroOrMore>
```

```
</content></elementSpec>
```

### *Element orthNote*

```
<elementSpec ident="orthNote" module="msdescription"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
prefix="re" mode="add">
```

```
<desc>Description of individual rules for letters or other
signs usage in the manuscript. </desc>
```

```
<classes mode="replace"><memberOf key="att.global" /><memberOf
key="att.typed" /><memberOf key="att.naming" />
```

```
<memberOf key="att.scoping" /></classes>
```

```
<content><rng:ref name="macro.specialPara" />
```

```
</content></elementSpec>
```

### *Element orthography*

```
<elementSpec ident="orthography" module="msdescription"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
prefix="re" mode="add">
```

```
<desc>Provided to describe the orthographical rules in the
writing of a scribe. </desc>
```

```
<classes mode="replace"><memberOf key="att.global" />
```

```
<memberOf key="att.typed" /></classes>
```

```
<content>
```

```
<rng:choice><rng:oneOrMore><rng:ref name="model.pLike" />
</rng:oneOrMore><rng:group>
```

```
<rng:optional><rng:ref name="summary" />
```

```
</rng:optional><rng:oneOrMore><rng:ref name="orthNote" />
```

```
</rng:oneOrMore></rng:group></rng:choice></content>
```

```
</elementSpec>
```

### *Element pagination*

```
<elementSpec ident="»pagination»" module="»msdescription»
```

```

ns=>http://www.ilit.bas.bg/repertorium/ns/3.0 prefix="re"
mode="add">
<desc>Describes pagination in some manuscript as opposite
to the description of foliation. Pagination is used in
later or Early Modern manuscripts.</desc>
<classes mode=>change><memberOf key=>att.global />
<memberOf key=>att.scoping /></classes><content>
<rng:ref name=>macro.specialPara />
</content></elementSpec>

```

*Element palaeoCharact*

```

<elementSpec ident="palaeoCharact"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Palaeographical characteristics</desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" /> <memberOf
key="att.handfeatures" /></classes>
<content><rng:choice><rng:oneOrMore>
<rng:ref name="model.pLike" /></rng:oneOrMore>
<rng:group><rng:optional><rng:ref name="summary" />
</rng:optional>
<rng:zeroOrMore><rng:ref name="palaeoNote" />
</rng:zeroOrMore><rng:optional>
<rng:ref name="musicNotation" /></rng:optional>
<rng:optional><rng:ref name="crypto" /></rng:optional>
</rng:group></rng:choice></content>
<attList><attDef ident="script" usage="opt"><datatype>
<rng:text /></datatype></attDef></attList></elementSpec>

```

*Element palaeoNote*

```

<elementSpec ident="palaeoNote"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
mode="add">
<desc>Palaeographical description of some letter
or sign</desc>
<classes mode="change">
<memberOf key="att.global" />
<memberOf key="att.typed" />
<memberOf key="att.handFeatures" />
</classes><content><rng:ref name="macro.phraseSeq" />
</content></elementSpec>

```

*Element pricking*

```

<elementSpec ident="pricking" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">

```



## *Guidelines to Repertorium Initiative XML Model for Manuscripts Descriptions*

```
<desc>Description of pricking appearance and techniques
applied</desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" /></classes>
<content><rng:ref name="macro.phraseSeq" />
</content>
</elementSpec>
```

### *Element quire*

```
<elementSpec ident="quire" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Description of a separate quire</desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" /></classes>
<content>
<rng:zeroOrMore><rng:ref name="p" /></rng:zeroOrMore>
<rng:zeroOrMore>
<rng:choice><rng:ref name="extent" /><rng:ref
name="signatures" /><rng:ref name="catchwords" />
<rng:ref name="gregoryRule" /><rng:ref name="ruling" />
<rng:ref name="pricking" /></rng:choice>
</rng:zeroOrMore></content>
<attList><attDef ident="status">
<desc>describes the status of the quire or folia: original,
added, missing</desc>
<datatype><rng:text />
</datatype></attDef></attList></elementSpec>
```

### *Element ruling*

```
<elementSpec ident="ruling" module="msdescription"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0"
  prefix="re" mode="add">
<desc>Ruling types and techniques goes here. </desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" /></classes>
<content><rng:ref name="macro.phraseSeq" />
</content></elementSpec>
```

### *Element sampleText*

```
<elementSpec ident="sampleText"
  ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
  mode="add">
<desc>Text sample including manuscript heading, incipit
and/or explicit of the original text</desc>
<classes mode="change"><memberOf key="att.divLike" />
<memberOf key="att.declaring" /><memberOf
```

```

key="att.global" /><memberOf key="att.typed" /></classes>
<content>
<rng:zeroOrMore><rng:choice><rng:ref name="head" /><rng:ref
name="rubric" /><rng:ref name="incipit" /></rng:choice>
</rng:zeroOrMore>
<rng:zeroOrMore><rng:choice><rng:ref name="model.pLike"
/><rng:ref name="div" /></rng:choice></rng:zeroOrMore>
<rng:zeroOrMore>
<rng:choice><rng:ref name="explicit" /><rng:ref
name="finalRubric" /></rng:choice>
</rng:zeroOrMore></content>
<attList><attDef ident="script" usage="opt">
<datatype><rng:text /></datatype></attDef></attList>
</elementSpec>

```

*Element scribe*

```

<elementSpec ident="scribe"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
mode="add">
<desc>Description of palaeography,
orthography and language of particular scribe</desc>
<classes mode="change"><memberOf key="att.global" />
</classes><content>
<rng:optional> <rng:choice><rng:ref name="locus" />
<rng:ref name="locusGrp" /></rng:choice> </rng:optional>
<rng:optional><rng:ref name="model.nameLike.agent" />
</rng:optional>
<rng:optional><rng:ref name="palaeoCharact" />
</rng:optional><rng:zeroOrMore> <rng:ref
name="scribeLang"/> </rng:zeroOrMore><rng:optional>
<rng:ref name="note"/></rng:optional>
<rng:zeroOrMore><rng:ref name="respStmt" />
</rng:zeroOrMore></content></elementSpec>

```

*Element scribeDesc*

```

<elementSpec ident="scribeDesc"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re" mode="add">
<desc>Repertorium model description of various scribes:
palaeography, orthography and language. </desc>
<classes mode="change"><memberOf key="att.global" />
</classes>
<content>
<rng:choice><rng:oneOrMore><rng:ref name="model.pLike" /></
rng:oneOrMore>
<rng:group><rng:optional><rng:ref name="summary" />
</rng:optional>

```

```
<rng:oneOrMore><rng:ref name="scribe" /></rng:oneOrMore>
</rng:group></rng:choice>
</content></elementSpec>
  Element scribeLang
<elementSpec ident="scribeLang"
ns="http://www.ilit.bas.bg/repertorium/ns/3.0" prefix="re"
mode="add"><desc>The language of particular scribe. </desc>
<classes mode="change"><memberOf key="att.global" />
<memberOf key="att.typed" /></classes>
<content><rng:choice>
<rng:oneOrMore><rng:ref name="model.pLike" />
</rng:oneOrMore>
<rng:group><rng:optional>
<rng:ref name="summary" /></rng:optional><rng:optional>
<rng:ref name="orthography" /></rng:optional>
<rng:optional><rng:ref name="grammar" /></rng:optional>
<rng:optional><rng:ref name="lexis" /></rng:optional>
</rng:group></rng:choice></content></elementSpec>
```

### *About the author*

**Dr. Andrej Bojadžiev** is Assoc. Prof. at the Sofia University „St. Kliment Ohridski“. His main fields of interests are Slavic historical linguistics, Slavic Cyrillic and Glagolitic palaeography, markup technologies, digital humanities, and electronic publishing.