

Схемы аналитического описания рукописного материала и проблема конвертации различных ИПС (на примере «Реперториум» Института литературы БАН и «Ведь» НИОР БАН)

А. Милтенова, В. Г. Подковырова

Институт литературы Болгарской академии наук, София, Болгария;
Библиотека Российской академии наук, Санкт-Петербург, Россия.

Информационно-поисковые системы (далее – ИПС) описания рукописных материалов предназначены для введения в широкий научный оборот новых данных и поддержки научных исследований в области кодикологии, палеографии, текстологии и ряда других гуманитарных дисциплин, базирующихся на рукописном материале. О сложности разработки системы описания рукописного материала в автоматизированных системах писалось неоднократно [Крушельницкая 1999; Подковырова и др. 2004], но, несмотря на эти сложности, к настоящему времени уже опубликовано несколько подобных схем описания [Miltenova, and et. 2000; Уханова 2004; Каширина и др. 2005 и др.]. Целый ряд ИПС описания рукописного материала разработан и уже функционирует в крупных хранилищах рукописей (ГИМ, Москва, БАН, Петербург и др.) и научных центрах (Институт литературы БАН, София, Екатеринбургский государственный университет и др.).

Однако до сих пор не решены проблемы разработки и принятия четких стандартов описания рукописного материала в подобных системах. Каждый научный центр за многие годы изысканий и практической деятельности по описанию разработал свои схемы. Например, в Отделе рукописей БАН с 1903 года, то есть за более чем сто лет его существования, опубликованы 20 томов научных описаний. Но даже внутри одного хранилища при составлении аналитического описания нет полного единообразия в его структуре. Это вызвано как разницей самих описываемых объектов, так и развитием терминологии и других аспектов научных знаний. Еще большая разница обнаруживается в деталях между описаниями в разных ИПС.

Информационный массив «Реперториум» реализован в последние десять лет (1997–2007 гг.). Модель описания в XML формате создана в

результате совместной работе ученых из Болгарии (Милена Добрева, Институт математики и информатики Болгарской академии, Анисава Милтенова, Институт литературы и Андрей Бояджиев, Софийский университет), из Голландии (Бейренд ван Дийк и Гари Гейлорд, Университет в Гронингене) и из США (Дейвид Бирнбаум, Университет в Питсбурге). Проект содержит около 350 описаний славянских рукописей XI–XVII вв. Компьютерное описание рукописей делается по модели, созданной на основе TEI (Text Encoding Initiative), и развивается в следующих направлениях: 1) стандартизирование описания рукописей как памятников культуры; 2) орфография, палеография и язык древних славянских памятников; 3) стандартизирование терминологии описания текстов; 4) схемы описания отдельных типов сборников; 5) идентификация конкретных текстов памятников письменности, их текстологических признаков и греческих соответствий; 6) связь между аналитическим описанием и полнотекстовой базой данных. Уже существует работающая поисковая система – проект *Repertorium Workstation* проф. Дейвида Бирнбаума из Питсбурга, США (URL: <http://clover.slavic.pitt.edu:8080/exist/mss/index.xml>).

Отсутствие единого стандарта как в схеме – структуре полей описания, так и в программном обеспечении пока не позволяет соединять данные о рукописях в единую базу данных. Если мы рассмотрим системы «Реперториум» (Ин-т литературы БАН, София) и «Ведь» (НИОР БАН, Санкт-Петербург) с точки зрения программной основы, то увидим, что они могут быть выставлены в Интернете без дополнительных доработок, поскольку обе системы созданы как платформонезависимые и с применением web-языков. Это уже немаловажный аспект их функционирования и гарант принципиальной возможности создания программы для конвертирования данных. Однако без принципиального единства структуры полей описания никакая конвертация невозможна.

В целом схемы аналитического описания рукописей не принципиально отличаются друг от друга. Обе схемы обязательно имеют две основные

части: внешнее описание рукописного материала (каталожные сведения) и описание его содержания (аналитическое описание).

Предложим в качестве примера сопоставительную таблицу части схем описания в двух представляемых ИПС (табл. 1)¹:

Таблица 1

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
<i>Внешнее описание</i> (каталожные сведения)	<i>I. Cataloguing Information</i>
<i>История записи</i>	<i>II. Technical documentation</i>
<i>Место хранения</i> <i>в настоящее время</i>	<i>2. Location of the Manuscript</i>
Название хранилища	2/2 The name of the city where the manuscript is kept
	2/3 Information on repository (library, archive, etc.)
Страна	2/1 Country
<i>Инвентарный учет</i>	
Инвентарный номер	2/5 Catalogue Number(s) [old and current]
Название по инвентарной книге	<i>1. The Manuscript Name</i>
Дата проверки наличия	<i>II Technical documentation</i>
Дата обнаружения отсутствия	<i>II Technical documentation</i>
Стоимость	<i>II Technical documentation</i>
<i>Данные о рукописи</i>	
Фонд	<i>2/6 Information on Collection(s)</i>
Шифр рукописного отдела	<i>2/4 Signatures or Shelf-marks</i>
Прежний шифр в БАН	
Прежний шифр до поступления в БАН	

¹ Полные схемы описания см. [Miltenova, and et. 1998; Miltenova, and et. 2000; Boyadziev, and et. 2003; Каширина и др. 2005].

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
Владелец или организация, где хранилась рукопись под прежним шифром	10 History of the manuscript [owner, commissioner, collector] Origin Provenance Acquisition
Примечание к прежнему шифру (история рукописи)	
<i>Состав единицы хранения</i>	<i>0 The Main Element</i>
Форма единицы хранения	
Копия – источник текста	7/3/6 The manuscript source
Фрагмент (по отношению к единице хранения)	7/3/8 MsPart (description of every fragment of manuscript separately)
Дополнительные элементы	[Additional information]
Примечания о составе единицы хранения	5/1 General Remarks. The elements <scribe> and <scribeList>
Научное название рукописи в целом (памятника) на языке описания	7/3/2 Standardized name of the manuscript given by the researcher(s)
Название по тексту рукописи (самоназвание)	7/3/3 The original name [title] of the manuscript
<i>Автор текста рукописи в целом</i>	
Полное имя автора текста рукописи в целом	7/3/4 The Author of the manuscript
Автограф (написана ли рукопись самим автором)	autograph

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
<i>Писец</i>	Scribe [scribes]
<i>Заказчик</i>	10 Owner, commissioner, collector (it
<i>Владелец</i>	is the part of the history of the
<i>Вкладчик</i>	manuscript)
<i>Прочие</i>	
<i>История рукописи</i>	Manuscript creation [manuscript date and manuscript place]
Место написания рукописи	6/3 The Place of Manuscript's Origin
Место хранения	6/1 General Remarks
История бытования рукописи	7/3/5 Texttological, linguistic or other data related with the manuscript [texts]
<i>Датировка</i>	6/2 Manuscript's Date
Примечания	6/1 General Remarks
<i>Количественные характеристики</i>	
Формат	3/5 Dimensions
Количество листов, страниц (в случае пагинации) или ставов (для свитка)	3/3 Number of leaves or folia
Размеры	3/5/1 Height 3/5/2 Width 3/5/3 Depth 3/5/4 Thickness

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
<i>Материальная основа письма (на чем написано)</i>	
Материальная основа	3/4 Material Description of the parchment Description of different kind of parchment in the manuscript Description of the paper Description of different kind of paper in the manuscript
<i>Материал письма (чем написано)</i>	
Материал письма основного текста	3/12 Ink(s)
<i>Алфавит</i>	
Алфавит	8 Alphabet
<i>Язык</i>	
Язык, на котором написана рукопись	
Извод	
Оригинал или перевод	7/3/7 The text translation
<i>Лингвистические особенности</i>	5/4 Description of Language characteristics
Фонетические особенности	
Наличие надстрочных знаков	
Морфологические особенности	
Синтаксические особенности	
Другие	

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
Примечания	5/3 Description of Orthography /3/1 General Remarks 5/3/2 The Element <letters> 5/3/2/a The Orthography of the jers 5/3/2/b The Orthography of the Juses 5/3/2/c The Orthography of the Letters for Jotated Vowels 5/3/2/d Description of the Orthography of the Other Letters
<i>Тип письма рукописи</i>	5/2/b Paleographical characteristics of the scribe's writing system
Тип письма	5/3/3 Ligatures
Примечания к типу письма	5/5/7 Other Signs
<i>Почерк</i>	5/5/6 Supralinear Signs
Кол-во разных почерков	Paleography [of every scribe]
Примечания	
<i>Переплет</i>	3/9 Binding
<i>Общая характеристика</i>	3/9/1 Boards
<i>переплета</i>	3/9/2 Box
	3/9/3 Cover
	3/9/4 Edges
	3/9/5 Fly leaves
	3/9/6 Fastening
	3/9/7 Furnishing
	3/9/8 Pastedowns
	3/9/10 Spine
Тип переплета	
Техника переплета	

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
Выступают ли доски за блок кодекса	
Примечания о переплете: наличие каптала, тип шитья и т.д.	3/9/9 Sewing and attachment to the boards
<i>Тиснение</i>	
<i>Накладные элементы</i>	
<i>Застежки</i>	
<i>Количественная характеристика организации текста на листе</i>	
Кол-во столбцов на листе	3/6 Columns
Число строк	3/7 Lines per page
<i>Нумерация листов кодекса</i>	3/11 Foliation and Pagination
<i>Нумерация тетрадей</i>	3/10 Quires 3/10/1 Description of the Quires 3/10/1/d Observations on the pricking
Тип нумерации тетрадей	3/10/1/f Markings
Материал письма (чем написано)	3/10/1/b Gregory Rule or the rule of parchment arrangement in the quire
Нарушения нумерации тетрадей	3/10/1/a Quire Composition
<i>Записи и вставки (дополнения)</i>	7/4 Description of the notes 7/4/1 General remarks. The <noteDesc> element
<i>Записи и вставки в текст</i>	
Листы	7/4/2/a The location of the note
Тип записи или вставки	7/4/2/e The formula of the note

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
Текст записи	7/4/2 The description of each note
Время внесения вставок или дополнений	7/4/2/b The date of the note
Тема записи	7/4/2/c Note contents
Имя писца	7/4/2/d The author of the note
Примечания к записи или вставке	7/4/2/g The elements in <noteltem> from the element <articleContentDesc>
<i>Криптография (Тайнопись)</i>	5/2/g Cryptography
<i>Художественные украшения</i>	3/13 Decoration
Тип художественного украшения	3/13/a Miniatures 3/13/b Headpieces 3/13/c Borders 3/13/d Tailpieces 3/13/f Wash 3/13/g Linefillers 3/13/h Initials 3/13/i Heraldry
Особые характеристики художественных украшений	7/4/2/h The decoration in the element <noteltem>
Примечания о художественных украшениях	5/2/c The description of the shape of each sign
<i>Характеристика бумаги</i>	
<i>Особенности бумаги</i>	
Тип бумаги (исходя из способа отлива)	3/4c Description of the paper
Примечание о цвете и структуре бумаги	3/4d Description of different kind of paper in the manuscript
<i>Водяные знаки</i>	3/4d/1 Watermark
Листы	3/4d/1/d Page range

Название поля (группы полей) ИПС «Ведь»	Название поля (группы полей) ИПС Repertorium
Научное название ВЗ	3/4d/2 List of watermarks
Описание изображения ВЗ	3/4d/1/a Motif 3/4d/1/b Countermark
Библиография	
Наличие расширенного описания ВЗ	3/4d/1/c Disposition
<i>Штемпели</i>	
<i>Сохранность рукописи</i>	4/ Damages, Restorations, and scientific notes on the manuscript condition
<i>Библиография рукописи</i>	8/ The Bibliographical Information in the <manuscriptLocation> element

Как видно из таблицы, большая часть параметров описания совпадает или очень близка, что при необходимости взаимной конвертации баз данных позволяет соединить их в едином поле.

Что касается аналитического описания содержания, то здесь дело обстоит еще более просто, несмотря на то, что именно описание текстов отдельных памятников положено в основу «Реперториума», а в ИПС «Ведь» основным объектом описания является рукопись. Но при этом описание текстов памятников существует как неотделимая часть описания рукописи (табл. 2).

Таблица 2

<i>II. Описание содержания рукописи (по листам)</i>	7 The Manuscript Contents
<i>II.1. Описание содержания</i>	
Листы	7/3/1 The location of the text [ff.]
Название по данным рукописи (самоназвание)	7/2/3 The original title of text on the level of the whole manuscript and of the level of separate articles

Научное название памятника на языке описания	7/3/2 The standardized name of the text given by the researcher
Инципит	Incipit
	Edition of the entire text [or of an important part of the text]
Эксплицит (концовка)	Explicit
Редакция памятника	7/3/8 Special elements for textual criticism
Примечания к названию и содержанию	7/3/8 Special elements for textual criticism
<i>II.2. Имя автора памятника</i>	Author
<i>II.3. Имя автора редакции памятника</i>	Editor
<i>II.4. Предметно-тематическая принадлежность памятника (Тематика содержания рукописи Предметно-тематический указатель)</i>	7/3 Textological comments on each text: archetype, protograph, antigraph, apograph; Sources of texts (Greek, Latin, etc. with quotation of reference books or editions in bibliography)

Сопоставление системы полей описания рукописных документов двух предложенных ИПС демонстрирует единство логической основы, позволяющей при необходимости создать программу конвертации всей информации или ее части из одной системы в другую, т.е. возможность создания единой БД рукописного материала на основе двух систем описания этого материала. Так как разработанные в БАН в Санкт-Петербурге и в Институте литературы в Софии системы уже опробованы на основе значительных массивов рукописей XI-XVIII веков, разнообразных по составу (богослужебные книги, сборники и др.), работа в направлении конвертации является одной из самых актуальных перспектив совместной деятельности ученых двух стран.

Список литературы

- Каширина и др. 2005 – *Каширина, Е. В.* Схема описания рукописного материала для автоматизированной системы Отдела рукописей БАН / Е. В. Каширина, В. Г. Подковырова // Материалы и сообщения по фондам отдела рукописной книги БАН. – СПб. : Изд-во Библиотеки РАН, 2005. – С. 377-394.
- Крушельницкая 1999 – *Крушельницкая, Е. В.* Проблемы отражения археографического, палеографического и кодикологического описания славянских рукописей в системе электронной базы данных / Е. В. Крушельницкая // Опыты по источниковедению. Древнерусская книжность: археография, палеография, кодикология. – СПб., 1999. – С. 223-239.
- Подковырова и др. 2004 – *Подковырова, В. Г.* Автоматизированная информационная система описания-хранения-поиска материалов из фондов Отдела рукописей БАН как особый тип локальной энциклопедии / В. Г. Подковырова, Е. В. Каширина // Энциклопедия как форма универсального знания: от эпохи Просвещения к эпохе Интернета. – СПб., 2004. – С. 300-307.
- Уханова 2004 – *Уханова, Е. В.* Компьютерные базы данных ОР ГИМ: принципы создания и перспективы развития / Е. В. Уханова // Археографический ежегодник за 2003 г. – М., 2004. – С. 79-96.
- Boyadziev, and et. 2003 – *Boyadziev, A.* A Unified Model for the Description of the Medieval Manuscripts? / A. Boyadziev, A. Miltenova, D. Radoslavova // Computational Approaches to the Study of Early and Modern Slavic Languages and Texts. – Sofia, 2003. – P. 113–125.
- Miltenova, and et. 1998 – *Miltenova, A.* Repertorium of Medieval Slavic Literature: Computer and Philological Standards / A. Miltenova, A. Boyadzhiev, S. Velev // Palaeobulgarica. – 1998. – XXII, N 2. – P. 50–69.
- Miltenova, and et. 2000 – *Miltenova, A.* An Electronic Repertory of Medieval Slavic Literature and Letters: a Suite of Guidelines / A. Miltenova,

A. Boyadzhiev // Medieval Slavic Manuscripts and SGML: Problems and Perspectives. – Sofia, 2000. – P. 44–68.

Schemes for analytical description of manuscripts and problems of conversion in various digital libraries (using as examples Repertorium and Vedi)

Anissava Miltenova, Vera G. Podkovyrova

Institute of Literature of the Bulgarian Academy of Sciences, Sofia, Bulgaria;

Library of Russian Academy of Sciences, St. Petersburg, Russia

The paper presents the process of standardization of analytical descriptions of Slavic manuscripts using informational technologies.