ПИСЬМЕННОЕ НАСЛЕДИЕ
И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Anissava Miltenova

Institute for Literature, Bulgarian Academy of Sciences (Sofia)

Scripta Bulgarica:
Digital Library of Medieval Bulgarian Literature

The project "SCRIPTA BULGARICA: Digital Library of Medieval Bulgarian Literature" (http://www.scripta-bulgarica.eu/en) is intended to create an electronic thesaurus and a library of Old Church Slavonic written tradition integrating three main types of information: a) terminology in the field of palaeoslavic studies — both in the existing scientific literature and in the successfully functioning digital corpuses; b) encyclopaedic articles about translated and original works outlining the most popular genres and genre forms in hagiography, homiletics, hymnography, chgronicles, narrative works and others; c) transcribed texts in Unicode from Slavonic manuscripts representing the translated literature from Byzantine Greek; d) articles about the most frequently translated Byzantine authors (e.g. Gregory the Theologian, St. John Chrysostom, St. Basil the Great, etc.), whose works (genres and genre forms) modeled and formed Orthodox Christian written culture. The articles for terminology include glossary in 7 languages: Latin, Greek, Old Church Slavonic, Russian, Serbian, English, German. Main information is in Bulgarian. The SCRIPTA BULGARICA project also includes a unique for such initiatives in the humanities connection between the generated standardized apparatus for description, research, publication and translation of medieval texts with the keywords and the bibliography. This allows the integration of the apparatus of scientific metadata in a uniform entity and the advanced stage of the project will see the development of an ontology of terms.

Keywords: thesaurus, terms in palaeoslavic studies, integrated digital corpus, Unicode, Drupal.

1. State of the Art in Bulgaria and in Europe

Bulgarian language in the Middle Ages (Old Church Slavonic) and medieval Bulgarian literature are studied in all universities in Bulgaria and in over 30 countries in Europe and in the world — so there is a constant interest to them. Over the recent years this interest has been growing, European society is trying to find more information about Slavic culture and especially for the Bulgarian one. The materials that represent medieval Bulgarian manuscripts

in Europe are sporadic, they are scattered in various sites, often in obsolete or outdated digitized formats (e.g. encoded in Latin: at CCMH http://www.helsinki.fi/slaavilaiset/ccmh and at TITUS http://wwwtitus.uni-frankfurt.de/texte/etcs/slav/aksl/). Moreover, some of the documents are kept in inaccessible repositories across Europe (and outside Europe), some of them are being kept in bad conditions and are not possible available to users. The preservation of these unique, rare and often fragile documents in digital form and by means of multimedia resources is a problem that requires an urgent response. The present project aims to contribute in the 21st Century to a better integration of medieval Bulgarian Literature in the culture of Europe to which it belongs.

The present digital representation of Bulgarian literature, especially from the Middle Ages is incomplete and insufficient. There is a lack of websites completed with Bulgarian original and translated texts taken from manuscripts and accompanied by appropriate metadata. There is no access to a complete panorama of the history of Bulgarian writing and writers. In a number of Bulgarian institutions, like libraries and universities digitization of the most important Medieval Bulgarian codices has been undertaken but these projects primarily include images of the manuscripts and there is no information on the contents and no access to the original texts. These projects are highly fragmented, not only because there is not enough funding for larger initiatives, but because of the lack of a comprehensive national strategy in Bulgaria in this respect.

2. Relevance of the problem in Bulgaria and in Europe

Recently, the interest in Bulgarian written cultural heritage has been tangibly increasing. In 2013, a year dedicated to the Moravian mission of SS. Cyril and Methodius, there were a lot of conferences and exhibitions which undoubtedly contributed to the growth of this interest. The forthcoming anniversaries of the death of St. Methodius in 2015 (1130 years) and of the death of St. Clement of Ohrid in 2016 (1100 years) are also of paramount importance.

The digital libraries of written cultural and historical monuments accompanied with scholarly commentaries and other well-structured complementary cultural and historical information, have a key role in modern education and programs for lifelong learning. Such digital archives are fact in the United States, for example, the fundamental projects in the field of classical philology: Thesaurus Linguae Graecae — initiated in 1972 at the University of California (Irvine); *Isocrates* — since 1985 at Brown University (Providence); *Perseus* — since 1988, at Tufts University (Medford, Massachu-

setts). These projects continue to this day. At the same time there are no such projects in Bulgaria and elsewhere that are to present the Slavonic and the medieval Bulgarian writing. Their main role would be to introduce before the specialists as well as before the general public the achievements of the medieval Bulgarian literature and cultural heritage. They would be actively used by a wide range of users and the methods they are built up upon would have a great influence in the field of the digitization of manuscripts. Today in the world there are more than 120 projects offering corpuses of works by the early Christian and medieval writers, as well as catalogues of manuscripts (see Robin Cover Pages and Text Encoding Initiative: http://www.tei-c.org / index.xml). Some of the most important Slavonic manuscripts have been included in the register of UNESCO "Memory of the World" (e.g. the 10th-century *Codex Suprasliensis*) and libraries worldwide include Bulgarian and Old Church Slavonic codices in EUROPEANA. Nevertheless, a complete presentation of the Slavonic and the Bulgarian written and cultural heritage is not available to the European users — neither as accessible metadata, nor as professionally transcribed sources.

This project aims to create an open digital library which is to present in a new, complete and comprehensive way the medieval Bulgarian and the ancient Slavic written heritage to Bulgarian and to European citizens by including descriptions of manuscripts, full-texts corpus and terminology thesaurus.

3. Project objective and tasks

The project "SCRIPTA BULGARICA: Digital Library of Medieval Bulgarian Literature" (http://www.scripta-bulgarica.eu/bg) financed by the National Science Fund (NSF) (contract No. ДФНИ-КО2/5 of 12.12.2014) is intended to create an electronic thesaurus and library of medieval Bulgarian texts (accessible through a special portal) integrating three main types of information:

a) Terminology in the field of Palaeo-Slavic studies — both in the existing scientific literature and in the successfully functioning digital corpuses. The project responds to the need for an easily accessible reference source of standardized terminology in several West European and Slavonic languages: English, German, Russian and Bulgarian, as well as their relevant terms in Old Church Slavonic, Greek and Latin;

b) Encyclopaedic articles about translated and original works outlining the most popular genres and genre forms in hagiography, homiletics, hymnography, chronicles, narrative works and others;

c) Transcribed texts from Slavonic manuscripts representing the translations from Byzantine Greek, original works and compilations of Bulgarian authors;

d) Articles about the most frequently translated Byzantine authors (e.g.
Gregory the Theologian, St. John Chrysostom, St. Basil the Great,
Athanasius of Alexandria, Ephrem the Syrian, Theodore the Studite,
and Hesychius of Jerusalem, among others), whose works (genres and
genre forms) modeled and formed Orthodox Christian written culture.

The portal was developed with the joint efforts of specialists from the
Institute for Literature and Sofia University. It steps on the experience and
materials accumulated as a result of several national and international proj-
ects of the Department of Medieval Bulgarian Literature at the Institute for
Literature in the last ten years. These projects include a Repertorium of Old
Bulgarian Literature and Letters (1995–2016) — within the framework of the
international Repertorium initiative, launched by the Department, which is
its basic organization. The international projects of the Department to create
digital resources in Palaeoslavonic studies were implemented in partnership
with the British Library (London) (1995–2012), the Library of the Russian
Academy of Sciences in St. Petersburg (2004–2010), the universities in Go-
thenburg and Uppsala (Sweden) (2008–2015), the Slovenian Academy of
Sciences and Arts, and the Institute of the Lithuanian Language in Vilnius,
among others. The Department has a traditional partnership with Sofia Uni-
versity, both in the case of the Repertorium initiative and in the development
of the Slavica Sanctorum digital encyclopaedia. As a result of these projects,
digital resources were created, and materials were collected which constitute
a full text information corpus uniting linguistic, archeographic, paleographic,
codicological and textological data.

The Scripta Bulgarica project also includes a unique for such initiatives
in the humanities connection between the generated standardized apparatus
for description, research, publication and translation of medieval texts with
the keywords and the bibliographic descriptions. This allows the integration
of the apparatus of scientific metadata in a uniform entity and the advanced
stage of the project will see the development of an ontology of terms.

4. Working Team

The working team of the project "SCRIPTA BULGARICA" brings together
specialists in the field of medieval studies and ancient languages and cultures
(trained in transcribing Old Church Slavonic manuscripts, text-critical analy-
sis, paleography and electronic editions of the codices with the help of infor-
mation technology): Prof. Anisava Miltenova (project leader), Assoc. Prof.
A. Stoykova, Assoc. Prof. R. Stankova, Asist. Prof. I. Kuzidova-Karadjino-
va, Asist. Prof. M. Petrova-Taneva, Asist. Prof. D. Petrova-Bankova, Asist.
Prof. R. Koycheva, Dr. S. Petrova-Hristova, Dr. G.. Nikolova (postdoctoral),

Dr. I. I. Iliev (postdoctoral), Dr. D. K. Petrova, S. Stefanov (postdoctoral), L. E. Stammler, D. Dimova, M. Vaptsarova, K. Hristova (PhD students). The software construction and technical support of the project portal SCRIPTA BULGARICA (architecture, analysis and data processing, searching modules, etc.) are carried out by Prof. Krassen Stefanov, Mihail Peltekov, Atanas Georgiev, Alexander Grigorov and PhD students at the faculty T. Zafirova-Malcheva, D. Miteva, E. Peltekova, colleagues from the Faculty of Mathematics and Informatics.

Assoc. Prof. Andrei Boyadzhiev is our associate from the Department of Cyril and Methodius Studies, Faculty of Slavonic philologies, Sofia University. The team of consultants in the project also includes Prof. Dorotei Getov (Institute for Literature, BAS), Assoc. Prof. Adelina Angousheva (University of Manchester) and Prof. David Birnbaum (a specialist in the field of medieval studies and informatics at the University of Pittsburg, USA).

5. Background

The ongoing projects in the Department of Old Bulgarian literature in the last years are of great importance for this initiative.

The project Repertorium of Old Bulgarian Literature and Letters was initiated in the Department of Old Bulgarian Literature, Institute of Literature, BAS. The Repertorium is a unique informational massive in Slavic studies that incorporates archeographic, palaeographic, codicological, textological and literary-historical data concerning the original and translated medieval texts, distributed in Balkan Cyrillic manuscripts (11th–17th cc.).

Now in the Repertorium are 161 encoded detailed manuscript descriptions, and 140 encoded short descriptions (http://aso.obdurodon.org/), together with metadata and bibliography. The Total content items in corpus are 3375 and unique content items are 1575. The methodological solution developed within the project enables fundamentally new submission of manuscripts in the XML format. The Repertorium allows an easy identification of individual texts and of manuscript compositions, investigation of the sources of the texts, study of the typology of the manuscript. The final product goes beyond a base of electronic records, it enables rethinking of the historical literary process (see: http://repertorium.obdurodon.org/). The project seta number of international initiatives and projects going: "Creating a digital library of Slavic manuscripts, incunabula and a reference database" — a joint project with the University of Gothenburg and the University of Uppsala, Sweden; "Machine-Readable Description and Searchable Catalogues of Cyrillic Manuscripts" (2005–2011) — a joint project with the British Library, London, for electronic description of the collection of Slavic manuscripts

there; "Medieval manuscripts and the possibility of presenting them using communication technologies" (2007–2010) — a joint project with the Research Centre of the Slovenian Academy (Ljubljana), as well as projects with the Institute for Serbian language (Belgrade), with the Institute of Lithuanian language (Vilnius) and the like. The Repertorium working team has been a member of the consortium Text Encoding Initiative since 2004 (see http://www.tei-c.org/Membership/) and an originator of the Commission on Data Processing of manuscripts and early printed books and on electronic publishing of medieval texts to the International Committee of Slavonic Studies (2003). The project is harmonized with the European initiatives CLARIN and DARIAH.

The Department of Old Bulgarian Literature in the Institute of Literature won a competition of ASO (Austrian Ministry of Science, Seventh Framework Programme) (2007–2008) with the project "SLOVO: Towards a Digital Library of South Slavic Manuscripts." Scholars from Austria, Slovenia, Serbia, Macedonia and Bulgaria were included in the project. An operational research network was established between institutions from five countries and aimed at studying Balkan monastery culture at risk. During the project an international conference was held in Sofia (21 to 26 February 2008) with participants from Bulgaria, Macedonia, Slovenia, Serbia, Croatia, Bosnia and Herzegovina, Britain, Germany, USA and Russia. There an establishment of a common platform for electronic publications in Cyrillic and Glagolitic was discussed. One of the outcomes of the project was an experimental Internet portal (http://slovo-aso.cl.bas.bg/).

The project "Metadata and electronic catalogs: multilingual electronic assistant researcher and teacher" (2004–2012), developed a common terminology database (accessible via Moodle). The project responded to the need in the European Information Space to build an easily accessible directory of standardized terminology in medieval studies in several languages: English, German, Russian, Serbian and Bulgarian. Scholars and specialists from BAS and Sofia University annotated 574 terms.

In 2010–2012 at the Institute of Literature was implemented the project "Codex Suprasliensis, Cyrillic manuscript of the tenth century: the creation of an electronic corpus, workshop and exhibition", financed by the UNESCO. The corpus includes the real text from the manuscript — deciphered, encoded and searchable, supplied with critical apparatus and a parallel Greek text. Digital images of the manuscript (kept in three libraries in three different countries: the National and University Library in Ljubljana, Slovenia; the Russian National Library in St. Petersburg, Russia; and the National Library in Warsaw, Poland) are available on line in parallel with the Old Bulgarian

deciphered texts (http://csup.ilit.bas.bg). During the project numerous problems and principles of electronic editions of medieval Bulgarian monuments were solved the (see: http://suprasliensis.obdurodon.org/). The project created a necessary basis for the construction of an extensive body of sources that can be used by a wide audience of users for various purposes.

In recent years, the working team issued five collections of research (in English, cf. bibliography). The results on this topic gain a high reputation in Europe and the world and are seen as a valuable Bulgarian contribution to this particular specialized area. No doubt that these results give opportunity Bulgarian, Balkan and Slavic heritage to become visible and available in the European information space. The contributions of the team from the Institute for Literature, BAS, at European level are: 1) As a result of experimental work together with foreign scholars the Glagolitic and the Old Cyrillic scripts were adopted in UNICODE (2007–2008); 2) a computer model for analytical description of Slavic manuscripts was completed, normalized to the requirements of the rules (P5) of the TEI consortium; the very TEI rules (P5) concerning the mediaeval manuscript description are developed by the active participation of Bulgarian scholars; 3) In cooperation with ISO and the European Digital Library an international standard for electronic edition of the old Slavonic texts and their presentation in the Internet was developed (2010–2012).

The Department of the Old Bulgarian Literature at the Institute of Literature, BAS, is the winner of a grand prize "Pythagoras" (Ministry of Education) for the best performing research team (2009).

6. Working program

The Main Project Aim is a creation of a universal Internet portal in which medieval Bulgarian texts, descriptions of South Slavic (and especially Bulgarian) manuscripts, and terminological data bases to be processed through innovative methods of comparison of various facts from the manuscript tradition on various levels: contents (macro- and micro-structure), language, codicology and paleography, etc. The project is envisaged to become a base for research and education in the field of the medieval Bulgarian literature, history, language, religion and culture and to allow an easy and fast access to the mediaeval texts — in original and in translation, as well as to the respective research metadata on the manuscript books and on the particular texts.

The steps toward the achievement of aims are as follows:

1. Creation of the Portal named SCRIPTA BULGARICA structured in accordance with the up-to-date requirements. The Portal will accommodate deciphered medieval Bulgarian texts (up to 1000 standard

pages), real texts (for the present they are 112) with concominant metadata, and a thesaurus of about 800 terms (for the present they are 474):

1. creation of an algorithm for statistical processing of medieval Bulgarian texts; algorithm's implementation;
2. creation of a search module; an algorithm by identity and by similarity; algorithm's implementation;
3. creation of an algorithm for automatic transition of the Bulgarian metadata to the English ones and vice versa.

2. Creation of a multilingual reference system of glossaries and thesauruses of terminology in palaeoslavistics:

   1. definition of terminological apparatus; collection of terms from bibliographic sources;
   2. completing the terminological catalogues with data from the original sources (book titles, genres, genre forms, etc.) and from analytic descriptions (terms in codicology, paleography, textology, Bible studies, etc.);
   3. comparison between the Bulgarian terms with their foreign counterparts and completion of the data base;
   4. computer processing of the glossaries and the thesauruses.

3. Study, collection and systematization of digitized material of texts and analytic manuscript descriptions, aiming at integration within a common corpus:

   1. creation of an inventory of manuscripts that are either Bulgarian of origin or connected with the Bulgarian history and culture and are kept in Bulgarian repositories or elsewhere;
   2. creation of analytic manuscript descriptions accompanied by relative metadata and bibliography;
   3. linkage of the descriptions to the digitized texts and texts corpora.

4. Creation of conditions for a sustainable development of the Portal by enlisting last-year bachelor's students, master's students, PhD students, and young researchers from the BAS and from the University of Sofia:

   1. testing of the electronic library and evaluation of its work by inquires;
   2. experimenting the system of small projects of bachelor's students, master's students, PhD students and young researchers (reading, analyses, annotating, translations, work with electronic sources) based on the electronic library;

3. involving bachelor's, master's and PhD students in building up new skills, such as unaided thinking, experimental work with texts and manuscripts; special attention will be paid to the work of the students in collaboration with each other; in application of knowledge in decisions concerning identifying and discovering new aspects in the studied artifacts; stimulating students to discover new artifacts (unknown and not yet described manuscript collections); to develop skills of presenting the results and discussing alternative decisions, as well as skills in written and oral research analysis.

5. Dissemination of the outcomes through the project's presentation and popularization in schools, universities, research and cultural organizations and among the general public (demonstrations, talks, lectures, seminars, etc.).

7. Type of activities, tasks, and methods

7.1. Tasks

A prime task of the project is to make a collection of digitized materials concerning the medieval Bulgarian written heritage easy-accessible to users of various needs — experts and specialists, as well as general users. In the National libraries in Bulgaria and other Balkan countries, such as Romania, Slovenia, Serbia, and Macedonia digitization of valuable manuscript collection has been in progress, but most of the digitized materials have not been yet deciphered. There are not enough original texts made available to the general public and for research and education. One of the project goals is to provide such deciphered texts supplied with translation in Modern Bulgarian and with commentaries.

This task will be accomplished by designing and building up an Internet Portal which architecture will unify a system of electronic libraries and resources for an effective, efficacious and user-friendly knowledge exchange in the field of the medieval Bulgarian writing. Core scenarios for detailed specifications of this architecture are to be created and analyzed. The hub of the system will be a storage for metadata describing all possible resources. The metadata and the resources in this storage will be extracted from the extant sources: electronic libraries created by the project partners, other electronic libraries concerning the written heritage, as well as other appropriate resources. The core storage will also contain a specially developed ontology called Ontology for medieval Bulgarian Literature (OOBL). The OOBL will be built up after a detailed analysis of the previous experience gained by the project team. This ontology will also be broadly discussed among the scholars in the field of medieval Bulgarian literature.

Another main task of the project will be a creation of a model for an interactive communication with users that will provide various forms of search and find, of offering and delivering appropriate digital resources for study, education, and general knowledge in the field of the mediaeval literature. On the basis of the model a prototype will be built. The latter will provide possibility for an automatic personalization of the interface that is to correspond at most to the users' characteristics, their roles in the society and their history of using the system. In order to allow the users to add their own resources and to create their own metadata on certain resources, an intelligent editor will be developed. It will be available in an individual working mode within the Portal. Thus, any user will be able to share resources within the system and to contribute to its development and enrichment. There will also be a mode for a mass usage through specialized software. Apart from that, experienced users who would like to incorporate automatically large bodies of metadata will be able to do that in a specially provided mode. The system will be made available in two language versions — Bulgarian and English, and with options of adding further language versions.

This way there will be accomplished the main aim of the project — to make a collection of digitized materials concerning the medieval Bulgarian written heritage easy-accessible and applicable in research and educational programs (for bachelor's, master's and PhD students, for teachers and lecturers, for research teams in the field of cultural studies and education).

7.2. Methods

The project is not intended towards a creation of new technologies and standards. It will follow the best practices in the field's methodology and digital resources, applied in some major projects such as European Digital Library, The Minerva Project, and The eContentPlus Program of the Ministries of Culture of the 25 EU member countries; especially in respect to the Convention of the Common European standardized model applicable for metadata description and edition of written sources. The project is consistent with the principles formulated by the Interedition (COST Action IS0704), a project for electronic editions and their use, in which activities the working team from the Institute of Literature took an active part.

The written sources will be described by means of the XML (Extensible Markup Language) that allows mark-up of certain text parts — freely chosen and with no range limitation, which, on its own turn allows fast and efficient search. The XML methodology for manuscript description in the oXigen ("<oXygen/>") interface follows the TEI P5 model which is a world-wide accepted standard. The main methodological principles in this respect are

as follows: 1) standardized file-format; multiple use for various educational purposes (the data are separated from the processing); 3) electronic text universality (independence from local platforms); 4) possibility for adequate storage of the corpus; 5) well-structured modules in the manuscript description allowing adequate search.

In its licensing policy concerning the texts' electronic versions the project hold on the Creative Commons BY-NC-SA license (http://creativecommons. org/licenses/by-nc-sa/3.0/legalcode) that allows a free of charge use for educational and research purposes.

The Ontology will be created by means of the Protégé Editor and will be incorporated into the main corpus of the library.

7.3. Type of activities

a) Research (search, cataloguing, annotation) of written sources at libraries and archives;

b) Transcribing of original Old Church Slavonic texts from manuscripts and conversion of existing electronic collections that are to be integrate into the SCRIPTA BULGARICA Portal;

c) Elaboration of the model for electronic manuscript description in XML-format, encoding the data and the relative bibliography, evaluation of the model and verification of the results;

d) Unification of the existing archives into a well-structured informational corpus which has to provide an easy access and data use;

e) Transfer and application of the project's knowledge, ideas, and methods into education, which activity envisages participation of last-year bachelor's students, master's students, and PhD students in the research and in the deciphering of the sources; experimental work using methods of interactive education and analysis;

f) Within the frames of the project, there will be held an international seminar and working meetings with leading researchers in the field of application of computer technologies in palaeoslavistics, such as Prof. David Birnbaum from the University of Pittsburg, Prof. Predrag Matejic, a Director of the Hilandar Research Centre at Ohio University, and others.

g) Testing the quality of the digital content and the search tools of the Portal by groups of students, scholars, and citizens through inquires and Internet forum in the SCRIPTA BULGARICA Portal;

h) Dissemination of the project's results into education on various levels — for bachelor's students, master's students, post-graduate students, and PhD students — through facultative courses, seminars,

lectures, and project presentations in Veliko Turnovo, Plovdiv, and Shumen.

We use *Drupal* system as a web portal for the Scripta Bulgarica project. Drupal provides flexible system with many ready available modules, design templates, and powerful search engines and it is also very good software development framework providing easy and user friendly interface. The system is easy to be used for linguistic experts without solid software engineering background.

8. Types of reference sources

The Scripta Bulgarica digital portal includes the following types of reference sources: a Thesaurus of terms, Articles about works, Source texts (written documents), Authoritative lists of Byzantine authors and Bibliography.

8.1. Thesaurus of terms

The Thesaurus is a reference source covering the categories listed above: hagiography, hymnography, rhetoric, textology, paleography and codicology, linguistics, macro-genres, and micro-genres. The working team adopted the following structure of an article:

1. header (term);
2. a short definition after the header. It is divided into three: term content, including the entire mass of meanings of the item designated by the term; term meaning — a small part of its content sufficient to identify it; term history — if it has changed its meaning it is noted separately.
3. identical terms in eight languages — Old Church Slavonic, Byzantine Greek, Latin, Russian, Serbian, English, German (or French); it is desirable to quote the author who has used it after the term in each language (with reference to bibliographical information);
4. bibliography — this information links the individual headers with references about the usage of the terms, the articles about works, source texts and authority files;
5. keywords, similar terms, references.

If synonyms exist (or different derivatives of a term) in contemporary academic literature, these are cited additionally. Where the term is lacking in Medieval Bulgarian, but it is a scientific term created as a consequence of research tradition, then this is mentioned in particular. The work on the complementing of these terms is divided into two stages: in the first stage, this information is amassed as a simple, unstructured list. The accumulation of these will lead to their hierarchisation and inclusion in ontological schemes.

канон

Описание | Глосарий

Един от основните полистрофични жанрове на православното богослужение. Възниква като допълнение към *библейските песни*. Изпълнява се на утринната и на някои други видове богослужение. Състои се *от песни*, всяка от които съдържа по един *ирмос* по няколко *тропара*, последният от които най-често е *богородичен*. Песните в състава на *канона* може да са 2, 3, 4, 8 или 9, което определя жанровата му разновидност (съответно *двупеснец, трипеснец, четирипеснец, осемпесенец и пълен канон*). Според едни автори пълният *канон* (от 9 *песни*) е възникнал в края на VII век и предшества по-кратките разновидности на жанра, а според други автори, напротив, най-кратките форми (към които се прибавя дори и т. нар. еднопеснец) са най-старинни.

Жанровата система на *химнографията* познава както отделния *канон*, така и цикълът от *канони*, предвидени за изпълнение в определен богослужебен период или свързани по друг начин (напр. по един за всеки от осемте *гласа*). В старобългарската *песенна поезия* такива са великопостният цикъл от 30 *трипеснеца* и 5 *четирипеснеца* от Константин Преславски, няколко осмогласни цикъла *канони* от св. Климент Охридски, предназначени за *Октоиха*, и много други.

Библиография:

Василик ВВ. Происхождение канона (Богословие, история, поэтика). Санкт Петербург: Санкт-Петербургского университета; 2006 с. 306.

Йовчева М. **Оригинални произведения на старобългарската химнография.** Приложение. В: Милтенова А. История на българската средновековна литература. История

Описание | Глосарий

стб.:
  канонъ
  (Kurz 1973)

  правило
  (Попов 1985: 150)

лат.:
  canon
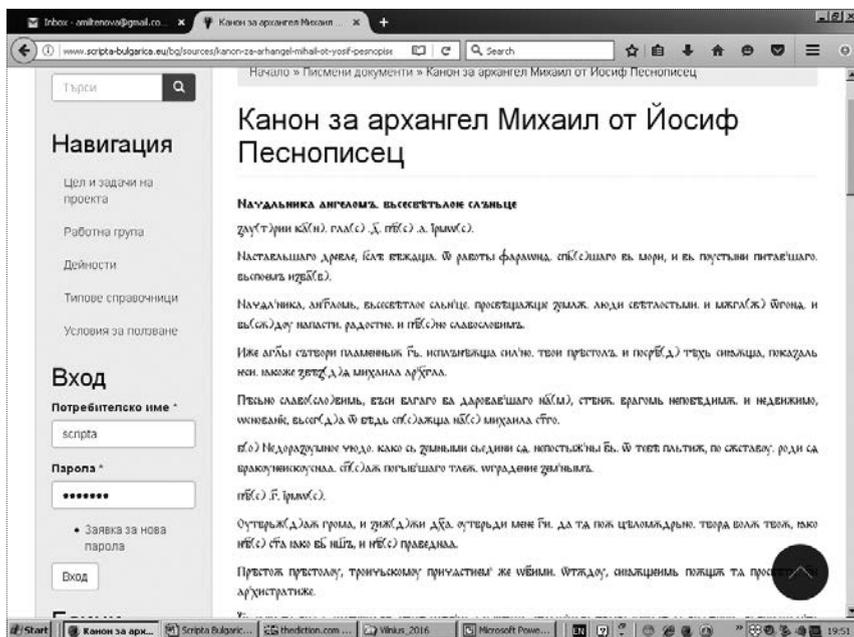 (Kurz 1973; Goar 1960: 332)

грц.:
  κανών
  (Wellesz 1961: 198; Getov 2007: 502)

рус.:
  канон
  (Kurz 1973; Василик 2006)

срб.:
  канон
  (Трифуновић 1990: 117)

англ.:
  canon
  (Wellesz 1961: 141; Getov 2007: 290, 502)

  kanon

8.2. Articles about works

These are articles representing meta-data relevant to medieval Bulgarian texts included in the digital library. They include: the title of the work, a short description of its origins, source data about it (copies); its association with a certain genre or genre form; its significance for the history of Medieval Bulgarian literature. The titles of the work in Old Church Slavonic and the modern European languages are also given. The bibliographic information includes: publications, translation in Modern Bulgarian (if any) and litera- ture. The keywords indicate relations (genres and genre forms) relevant to the work. The scientific titles are compiled by the principle of ontologies, i.e. first the general title is entered, followed by the specific one, e.g. Miracles of St. Nicholas of Myra, Miracle about the carpet; or Series of stories about the Holy Tree. Story about the third tree. The principles of the titles are followed and they are compared to those included in the project Repertory of Old Bul- garian Literature, Institute for Literature, BAS.

8.3. Source texts transcribed from the mansucripts in UNICODE

The texts are transcribed in machine-readable format that allows easy export of medieval documents in the Web and their use for digital processing.

8.4. Authority files about Byzantine authors

The authority files provide related information about the texts and their authors. The system of authority files is well-known in library practice and the same formats are used in the methodology of their creation. The files include a title (name of author in Modern Bulgarian, Old Church Slavonic, Greek, Latin, English and German, e.g. Ефрем Сирин, Ефрем Сириец (с. 306–373); Old Bulgarian: Ефремь (бл̃жен'наго Ефрема, с̃тго о̃ца нашег Ефрема); Greek: Ἐφραίμ ὁ Σῦρος; Latin Ephraem Syrus; English Ephrem the Syrian; German Ephräm der Syrer); a short review of the biographical data about him and about his works, particularly focused on those translated in Old Church Slavonic; bibliography.

Terms of use: free access through Creative Commons (Creative Commons Attribution-NonCommercial-ShareAlike 2.0 Generic license).

9. Expected outcomes, sustainability, and effect of the project SCRIPTA BULGARICA

9.1. Expected outcomes:

a) access to new, unstudied or so far unavailable sources on the history of medieval Bulgarian Literature — in electronic form (about 1000 pages);

b) development of new methods of access to cultural and historical heritage kept at more than 50 libraries and archives (including monasteries on Mount Athos and Mount Sinai). Most of these sources are not accessible via Internet.

c) analytic descriptions in English of more than 200 manuscripts kept at Bulgarian and at foreign archives and libraries;

d) new ontology of terminology in mediaeval letters;

e) new internet Portal providing access to all existing resources;

f) creation of a module for user feedback.

9.2. Exploitation plan

9.2.1. During the time of the project an exhaustive plan for further use and applying of its outcomes will be created. Apart from that, there will be made sufficient efforts to attract the general public attention towards the project's aims through disseminating the project's outcomes. The project results will be published in Internet, in scholarly, educational, and popular editions, as well as in the mass media in order to promote them not only among scholars but also among teachers, students, journalists, museum specialists, local administrations and cultural societies and associations.

The sustainability of the project activities is part of their own nature, i.e.:

a) free access to the project's outcomes through Internet resources which are to provoke interest and to motivate new studies; the project is open of type, which involves its further enrichment and development;

b) the project's inter-disciplinary and multi-disciplinary approach will involve students in learning new paradigms and research approaches; the project's scheme is universal of character and is applicable to any sort of sources;

c) the project aims at setting up a network of young researchers and academic stuff in Bulgaria and abroad which guarantees further joint research and exchange of knowledge, literature and materials.

9.2.2. The human factor in the project's sustainability. The fact that the team will include academic stuff and students from the University of Sofia assures rapid application of the project's results in education. The participation of PhD students and young researchers is a precondition for further use of the collected materials and further development of the shared ideas about the sources and the methods of electronic publishing.

9.2.3. The financial support by the Bulgarian Scientific Fund is crucial for the project's development. Nevertheless, extra funding will be seek from international programs for the preservation of the Balkan and the Slavic cultural heritage, programs of the local communities, administration, munici-

palities and municipal centers and museums for preservation of the cultural heritage and traditions in various Bulgarian regions.

The project aims at stimulating the interest of civil and religious non-government organizations that would gain their own educational and popularization aims through this project.

Bibliography

Birnbaum D., Miltenova A., Slevinski S. (ed.), 2003: *Computational Approaches to the Study of Early and Modern Slavic Languages and Texts*. Sofia.

Miklas H., Miltenova A. (ed.), 2008: *Slovo: Towards a Digital Library of South Slavic manuscripts. Proceedings of the international conference, 21–26 February 2008, Sofia, Bulgaria*. Sofia.

Miltenova A., 2000: Bulgarian Studies at the Dawn of The 21st Century, *A Bulgarian-American Perspective. Sixth Joint Meeting of Bulgarian and North American Scholars, Благоевград, 30.05–02.06.*1999. Sofia.

Miltenova A., Birnbaum D. (ed.), 2000: *Medieval Slavic Manuscripts and SGML: Problems and Perspectives*. Sofia.

Miltenova A. (ed.), 2012: *Rediscovery. Codex Suprasliensis, Bulgarian manuscript of 10th century*. Sofia.

Miltenova A., Radoslavova D., Pancheva E. (ed.), 2006: *Computer Applications in Slavic Studies. Proceedings of Azbuky.Net International Conference and Workshop, 24–27 October 2005, Sofia, Bulgaria*. Sofia.

Анисава Милтенова

SCRIPTA BULGARICA:
цифровая библиотека средневековой болгарской литературы

Целью проекта "SCRIPTA BULGARICA: Digital Library of Medieval Bulgarian Literature" (http://www.scripta-bulgarica.eu/en) является создание электронного тезауруса и библиотеки памятников старославянской письменности. Предполагается интеграция четырех основных типов информации: a) палеославистической терминологии (как представленной в научной литературе, так и успешно функционирующей в цифровых корпусах); b) энциклопедических статей о переводных и оригинальных произведениях (особое внимание уделяется наиболее популярным жанрам и жанровым формам агиографии, гомилетики, гимнографии, хроник, нарративным произведениям и т.д.); c) наборных текстов (в Unicode) славянских памятников, представляющих собой переводы византийской грекоязычной литературы; d) статей о наиболее часто переводившихся византийских авторах (Григории Богослове, св. Иоанне Златоусте, св. Василии Великом и др.), чьи работы (жанры и жанровые формы) оказали влияние на формирование православной христианской письменной культуры. Раздел терминологических статей включает глоссарий на 7 языках: латыни, греческом, старославянском, русском, сербском, английском, немецком. Основ-

ная информация приведена на болгарском языке. Проект SCRIPTA BULGARICA
оснащен уникальной для подобных инициатив в гуманитарных науках связью авто-
матически генерируемого стандартизированного научного аппарата описаний, ис-
следований, публикаций и переводов средневековых текстов с ключевыми словами
и библиографией. Такая связь позволяет: интегрировать аппарат научных метадан-
ных в унифицированную словарную статью; на продвинутой стадии проекта при-
ступить к разработке онтологии терминологии.

Ключевые слова: тезаурус, термины палеославистики, интегрированный циф-
ровой корпус, Unicode, Drupal.

Anisava Miltenova

SCRIPTA BULGARICA:
skaitmeninė bulgarų viduramžių literatūros biblioteka

Projekto "SCRIPTA BULGARICA: Digital Library of Medieval Bulgarian Litera-
ture" (http://www.scripta-bulgarica.eu/en) tikslas — sukurti senosios bažnytinės slavų
literatūros biblioteką ir tezaurą. Numatoma integruoti keturis pagrindinius informacijos
tipus: a) paleoslavistikos terminiją (tiek vartojamą mokslinėje literatūroje, tiek sėkmin-
gai funkcionuojančią skaitmeniniuose korpusuose), b) encilopedinius straipsnius, skirtus
vertimams ir originaliems kūriniams (ypač dėmesio skiriama populiariausiems žanrams
ir jų formoms: ageofrafijai, homiletikai, kronikoms, naratyvams ir kt.) c) transliteruotus
(Unicode) senosios bažnytinės slavų literatūros paminklus — Bizantijos autorių kūrinių
vertimus; d) straipsnius apie dažniausiai verčiamus Bizantijos autorius (Grigalių Nazia-
rietį, Joną Auksaburnį, Bazilijų Didįjį ir kt.), kurių kūriniai (žanrai ir jų formos) turėjo
įtakos stačiatikių rašytinės kultūros formavimui. Terminologinių straipsnių skyriuje patei-
kiamas glosarijus 7 kalbomis: lotynų, graikų, senąja slavų, rusų, serbų, anglų, vokiečių.
Pagrindinė informacija pateikiama bulgarų kalba. Projektas *Scripta Bulgarica* turi unika-
lų (kol kas nebūdinga tokio tipo humanitariniams projektams) automatiškai generuojamo
standartizuoto aprašų, tyrimų, kūrinių tekstų publikacijų ir vertimų mokslinio aparato ir
reikšminių žodžių bei bibliografijos ryšį, tai leidžia ne tik integruoti mokslinių metaduo-
menų aparatą į žodyno straipsnį, bet ir tolesnėse projekto vystymo stadijose pradėti rengti
terminų ontologiją.

Reikšminiai žodžiai: tezauras, paleoslavistikos terminai, integruotas skaitmeninis
korpusas, *Unicode, Drupal*.

Анисава Милтенова, д-р филологических наук, профессор, заведующая Сектором
древнеболгарской литературы, Институт литературы, Болгарская академия наук, София,
Болгария.

Anissava Miltenova, Ph.D., Professor, Head of the Department of Old Bulgarian Literature,
Institute of Literature, Bulgarian Academy of Sciences, Sofia, Bulgaria.

Anisava Miltenova, filologijos mokslų daktorė, profesorė, Senosios bulgarų literatūros sekto-
riaus vedėja, Bulgarijos mokslų akademijos Literatūros institutas, Sofija, Bulgarija.

E-mail: amiltenova@gmail.com