

РЕПЕРТОАР НА СТАРОБЪЛГАРСКАТА ЛИТЕРАТУРА И КНИЖНИНА
С КОМПЮТЪРНИ СРЕДСТВА

(База данни от 200 ръкописа)

Приложението на информационните технологии в славянската филология, историография и архивистика изостава чувствително от други хуманитарни области, в които приложението на компютрите е навлязло преди няколко десетилетия. В областта на класическата филология напр. отдавна съществува богата база данни както от библиографски описания, така и от обработката на впечатляващ корпус от текстове с компютърни средства, която продължава и сега (напр. *Thesaurus Linguae Graecae* — разработван от 1972 г. в Университета в Калифорния, Ирвин, САЩ; *Isocrates* — разработван от 1985 г. в Университета Браун, Провиден, САЩ; *Perseus* — разработван от 1988 г. в Харвардския университет и Бостънския университет и др.). Резултатите от тези проекти могат да се ползват както чрез Интернет, така и чрез CD-ROM. Едва напоследък (от 1999 г.) в Монтфортския университет (Дания) с финансовата ходкрепа на Европейската Общност се осъществява проектът *Master*, който има за задача да разработи схема за база данни за ръкописи, старопечатни книги и архиви, писани с *само* латинска азбука. Този проект все още е в начален стадий.

В миналото са правени са опити славянските ръкописи да бъдат каталогизирани с библиотечни продукти за база данни (*ISIS*), какъвто представляваше проектът, замислен преди повече от десет години в *SIBAL* (БАН), но те изискват фиксиран характер на полетата, в които може да се кодира текст, а също така — поместват ограничено количество информация. В Хайделберг, Германия бе създаден комерсиалният продукт *IST* (фирма *TSAR*), който бе използван в книгата на Т. Черторицкая (*Vorläufiger Katalog Kirchenslavischer Homilien des beweglichen Jahreszyklus. Aus Handschriften des 11-16. Jahrhunderts vorwiegend ostslavischer Provenienz. 1994*). Оказа се, че този софтуер има редица недостатъци, непълноти и неудобства с оглед разработването на репертоар (напр. използването на съвременна *руска*

кирилица; наличието на сложен шифър за обозначаване на отделните текстове и др.).

Настоящата система за компютърно аналитично описание на средновековни славянски ръкописи, използвана в проекта “Репертоар на старобългарската литература и книжнина” (**TSM – Template for Slavonic Manuscripts**); бе осъществена от м. юни до м. август 1994 г. в гр. Питсбърг, САЩ, в процеса на съвместната работа на проф. д-р Дейвид Бирнбаум от Университета в Питсбърг, САЩ и на ст. н. с. II ст. д-р Анисава Милтенова, Секция за стара българска литература, Институт за литература при БАН. Първоначалната програмна реализация принадлежи на колектив в състав: Бейренд Дайк, тогава аспирант в Университета в Грьонинген, Холандия, Д-р Хари Гейлорд, преподавател по компютърни системи в Грьонинген и н. с. Милена Добрева, Институт по математика и информатика при БАН. Експериментирането на програмата чрез тестове продължи почти до края на 1994 г. През м. юли—август на 1995 г. бяха направени последни промени и уточнения в дефиницията на типа документ (DTD) при посещението на проф. Дейвид Бирнбаум в София в колектив: проф. Д. Бирнбаум, ст. н. с. А. Милтенова, н. с. А. Бояджиев. Проектът бе спонсориран от IREX, Вашингтон, чрез конкурс (1994-1995), както и от фондация “Отворено общество”, София, през периода 1994—1997. След 1997 г. проектът бе продължен в сътрудничество с Централно-Европейския университет — Будапеща, с координатор проф. Ралф Клеминсън от Университета в Портсмут, Великобритания. Той бе подкрепен и от Националния фонд за научни изследвания при МОН. Резултатите от проекта бяха публикувани в сборника “Medieval Slavic Manuscripts and SGML: Problems and Perspectives” (Sofia, “Prof. Marin Drinov” Publishing House, 2000).

Използуваното в настоящия проект описание *специално* е пригодено за стари славянски текстове на равнището на съвременната археография, палеография, кодикология и текстология; разработката на шрифтове за изписване на оригиналните текстове със стара кирилица принадлежи на н. с. Румян Лазов от Института по математика и информатика при БАН. През 1997-1998 г. бе разработена гъвкава търсеца програма възоснова на софтуера “Magic” от Станимир Велев (Институт за литература).

Изработването на Репертоара е многогодишна научна задача, първият етап от която е обработката на балканските кирилски ръкописи с небиблейско съдържание от X-XI до края на XIV в. В Секцията за стара българска литература, Институт за литература при БАН са обработени вече над 200 разнообразни по съдържание ръкописи чрез тази система в работен колектив от осем души. Привлечени са като експерти и постоянни участници колеги от Софийския университет.

TSM е създадена по стандарта Standard Generalized Markup Language (SGML), приет от Международния комитет по стандартизация (ISO). Този стандарт се базира на възможността да се включват специални “маркировки” в текстове на естествени езици, които заграждат определени части от текста и сигнализират какво представляват тези данни. Това улеснява значително извличането на данни от текста при неговата компютърна обработка. Работата със Standard Generalized Markup Language (SGML) навлиза в практиката през 90-те години чрез работата на група специалисти от Франция, САЩ, Холандия и Германия обединени в т. нар. Text Encoding Initiative (TEI), които са изпробвали приложението на множества от **маркери** върху различни класове текстове — поезия, драма, проза, речници. Този стандарт за пръв път тук се прилага при описанието на стари славянски ръкописи и за включване на *произволни* (свободни от ограничения) обеми от ненормализирани текстове от самите ръкописи в процеса на описанието. Той отговаря на следните принципи: 1. Стандартизация на файл-формата на документа; 2. Възможност за многократна употреба за различни цели (данните трябва да бъдат отделени от процеса на обработка); 3. Универсалност на електронния текст (независимост от локални платформи); 4. Възможност за адекватно съхраняване на ръкописите в електронна форма; 5. Възможност за добре структурирани части на различните аспекти при обработка на ръкописите (кодикология, палеография, ортография, текстология и т. н.).

Системата за кодиране на средновековен славянски текст (TSM) единодушно е приета като стандарт за електронна обработка от специалистите в страните, в които се съхраняват славянски ръкописи. Тя бе обсъдена на международна конференция в Благоевград, (24-28 юли, 1995) и

получи висока оценка. От тогава до днес проектът е демонстриран на редица международни форуми. През 1997 г. проектът бе представен на семинар в Централно-европейския университет в Будапеща пред участници от дванадесет страни. По време на 12-тия международен конгрес на славистите в Краков, 1998 г. се състоя кръгла маса на тема “Компютърна обработка и анализ на средновековни славянски ръкописи и старопечатни книги”, която бе посрещната с голям интерес. В резултат бе образувана специална комисия за компютърна обработка на средновековни ръкописи, старопечатни книги и архивни материали към Международния комитет на славистите.

Репертоарът на старобългарската литература и книжнина се осъществява като универсален информационен масив, който обединява археографски, палеографски, кодикологически, текстологически и значими литературно-исторически данни за оригиналните и за преводните средновековни текстове, разпространени в балканските кирилски ръкописи. Репертоарът обединява части от *реален текст* от конкретните ръкописни книги (напр. оригинално заглавие, начало и край на статията на старобългарски език; подглави или съставящи статии на сложните по композиция текстове и пр.) и *резултати от научното изследване на състава на ръкописите*. Специално място е отделено на *историята на отделните текстове* (напр. при преводните текстове се отбелязват вероятните им византийски първообрази, ако са установени такива, времето и мястото на превода, литературната редакция на текста, предполагаемият протограф, антиграф и т. н.) и на тяхното разпространение в други славянски литератури. Репертоарът предвижда и включването на библиография за ръкописите и за отделните текстове в тях (издания, изследвания, преводи на съвременни езици). Търсенето и справките в репертоара се осъществяват чрез индекси, които функционират по определени показатели: хронология, книгохранилище, тип на ръкописа, преписвач, място на създаване на ръкописа, взаимоотношение между тип ръкопис и тип статии според избрани ключови думи, автор на статии или на конкретен текст, дата от християнския календар, определено заглавие (оригинално), научно заглавие (на английски и на български), начало и край на конкретен текст, приписки и т.н. Със своето постепенно изграждане репертоарът на южнославянските

ръкописи предполага *не само* изследване на контекста, в който се поместват съчиненията на един или друг автор или проучване типологията на ръкописната книга — най-същественят очакван резултат се заключава във възможностите за *преосмисляне на историко-литературния процес*, в създаване на условия за *нов поглед* върху явленията, особено що се отнася да преводната литература през Средновековието.

TSM е структурирано от пет взаимосвързани дяла: описание на файла, научно описание на ръкописа като цяло, научно описание на отделните текстове в ръкописа, образци от текста на ръкописа, библиография¹.

I. Описанието на файла включва: името на автора на файла; името на редактора на файла; институцията, която има делегирани авторски права за публикуване на данни от описанието; град и държава, в която се осъществява описанието; дата на изработване на файла; източници на описанието (картотека, каталог, публикация или работа “*de visu*”); предвидено е отбелязването на името на ръкописа на български и на английски език; място на съхранение на ръкописа (име на архивохранилището и книгоохранилището, на сбирката); сигнатура — текуща и стара; каталожен номер; лице, на което е принадлежал ръкописът преди да постъпи в книгоохранилището; библиографски данни. Както се вижда, тук са събрани археографски данни, които обикновено се съдържат в кратките каталози на славянски ръкописи. Тази част от електронното описание защитава авторските права на изследователя, който е съставил файла, а също така — препраща към библиографията, свързана с *описанието* на конкретния ръкопис. Библиографските сведения могат да се включват вътре в електронното описание според **TSM**, но могат и да съществуват в отделен файл, към който може да се препраща с кодов номер. Към всяко поле от описанието е възможно да се добави коментар в свободен текст за допълнителни данни, напр. по отношение историята на ръкописа и др.

II. Научното описание на ръкописа като цяло дава възможност за подробна кодикологическа, палеографска и правописна характеристика на ръкописа. Хронологията на ръкописа (неговото датиране) и локализацията

му са включени като резултат от неговото палеографско и ортографско проучване. При случай, че датировката е заета от готов източник (каталог, описание или друга публикация), това изрично се указва в полето, определено за отговорността на учения по отношение на фиксираните данни. В тази част на електронното описание се съдържат още сведения за източника на ръкописа (ако той е известен), за идентификацията на превода и неговата хронология (ако паметникът е преводен и ако преводът е проучен), за евентуалния протограф, антиграф и апограф на целия ръкопис, за литературната редакция на паметника като цяло. Изброените елементи са предвидени да бъдат използвани тогава, когато ръкописът представя *един* паметник (напр. изборно евангелие, псалтир без съпътстващи статии и др.), като изследователят преценява, че няма нужда да описва по отделно всеки графично маркиран текст в него (напр. отделните псалми). Те могат да служат за описание също така на ръкопис, който включва едно единствено произведение, напр. Лествица на Йоан Синайски, Диоптра на Филип Монотроп, Житие на св. Василий Нови и т. н.

III. Научно описание на отделните текстове (статии) в ръкописа. Тази част от електронното описание е структурирана според обстоятелството, че повечето средновековни славянски ръкописи по своята същност представляват сборници (антологии) от текстове. Аргументите за подбора на възможните признаци, според които се изгражда научното описание на конкретен текст, са разгледани в самостоятелна публикация². Тук само ще подчертаем, че при формулирането им на първо място е било ръководещо съображението за рационалност на описанието — целта, която си поставя настоящият проект е не само събиране на текстове от славянското ръкописно наследство в единен свод, но представянето им в такъв вид, че да се улесни тяхното използване. В този смисъл изключително важен е въпросът за идентификацията на текстовете и стандартизирането на техните научни наименования, който ще бъде засегнат подробно по-долу в частта за документалния език на научното описание. В елементите са намерили място

¹ За характеристика на отделните признаци вж. в статиите по отделни въпроси.

² *Miltenova, A.* Computer-Assisted Analysis of the Macrostructure and Typology of Medieval Slavic Miscellanies. — In: Computer Processing of Medieval Slavic Manuscripts. First International Conference, Blagoevgrad, 24-28 July 1995. Proceedings. Sofia, 1995, 129—139.

както такива, които са почерпани от реалния текст (напр. името на автора или псевдоавтора на статията; данни за датата от църковния календар, към която е прикрепен текстът), така и признаци, които са резултат на оценката на изследователя — напр. класификацията според TextType (оригинален, компилация или превод), според TextStyle (наративен, ненаративен); описанието на източниците на текста, което е особено важно при компилативния тип творби и др. Именно в тази част на електронното описание е придвидено да се отбележат най-важните етапи от историята на отделните текстове, което е особено важно при проследяване на тяхното разпространение в ръкописите. Включват се също така данни за мястото на определен текст в ръкописа, за предхождащите и за следващите го текстове, което позволява да се изучава структурата на състава в сравнение с други сходни по съдържание ръкописи.

На този етап от проекта приемаме, че репертоарът на средновековната литература и книжнина се съставя въз основа на научното описание и реалния текст на отделните статии (текстови единици), съдържащи се в ръкописа, чието диференциране става въз основа на съвкупност от определени признаци. При това сме се стремили да спазваме следната последователност при структуриране описанието на текста:

- изграждане на конвенция за общ модел (модели) на текста.
- създаване на инварианти на този модел за композиции, структурирани на паратактичен принцип (напр. за текста на Физиолога, който е съставен от отделни кратки наративни части, за текста на компилациите от ексцерпти, които нямат вътрешни заглавия и др.).
- изграждане на непротиворечива йерархична структура на текстовите единици, които имат мозаечна структура или композиция на хипотактичен принцип (напр. химнографската композиция служба, чинопоследованията и др.).

IV. Образци от текста на ръкописа. Основното изискване тук е да се ексцерпира достатъчно количество текст, осигуряващо неговата бърза и безпогрешна идентификация. Образците са толкова, колкото изследователят е предвидил според избраното от него разчленение на структурата на ръкописа; научното описание на всяка статия (относително самостоятелна

текстова цялост, рамкирана от заглавие, начало и край) е свързано чрез вътрешни поредни кодове (ACD1, ACD2, ACD3 и т. н.) с реалния текст.

1). Заглавие на текста <TITLE> (ако той има такава, а не е фрагмент с липсващо заглавие). То често съдържа вътрешнолитературно жанрово определение на средновековния текст, информация за темата на творбата, сведения за функцията на текста (за неговото предназначение), понякога — името на неговия автор, в някои случаи — дата от християнския календар, за която е предназначен и т. н. Богатите данни, които съдържа заглавието на средновековния текст, го правят необходим елемент в репертоара.

2). Начало на статията <INCIPIT>. Тази част от текста има значение на диференциращ признак и спомага най-много за неговата идентификация. Началото е важно за установяване на превода и на литературната редакция на текста. Отрязъкът от текст, бележещ началото, трябва да бъде достатъчно голям (в рамките на 10-15 думи или даже на една синтаagma) за да могат да се правят изводи за същността му. Освен това, както е известно, началото на текста може да бъде база за различни наблюдения и статистики, напр. може да бъде показателно за неговата жанрова принадлежност.

3). “Тялото” <BODY> на статията е условно обозначение, което дава възможност да бъде обработен в електронна форма отрязък текст, който не е начало и не е край; при произведения със сложна структура или с композиционно зададена нееднородност (напр. редуване на ненаративни и наративни текстове в рамките на службата) елементът “тяло” може да бъде повторен няколко пъти както според паратактичен, така и според хипотактичен принцип — толкова, колкото е необходимо, за да бъдат описани всички съставлящи подразделения на текста (най-често — чрез елемента <DIV> и <P>)³. В тази част от схемата се включва реален текст и в случаите, когато заглавието, началото или края на текста липсват поради механична повреда на ръкописа, т. е. могат да се описват фрагменти от определена статия, която вече е идентифицирана в научното описание (III).

4). Край на средновековния текст <EXPLICIT>. Той може да съдържа (освен задължителната формула) данни във връзка с историята на

³ Вж. конкретните приложения на модела при различните типове ръкописи.

конкретното произведение, ето защо неговото регистриране е също толкова важно, колкото и на началото на текста.

V. Библиография. Тя е структурирана според международния стандарт, като има възможност да бъде изнесена в отделен, спомагателен файл към цялостния корпус от описания на ръкописи. С други думи повторението на едно и също заглавие в един или в няколко описания може да бъде избегнато чрез вътрешни препратки (<REF>). Предвидени са няколко места за вмъкване на библиография — цитиране източниците за каталожното описание (в рамките на описанието на файла, тук I); позоваване на изследвания и публикации върху целия ръкопис (в края на научното описание на ръкописа); позоваване на проучвания върху отделни статии (произведения, текстове) в състава на сборниците (в края на научното описание на всяка статия).

Важно е да се подчертае, че в схемата TSM е предвиден изчерпателен брой елементи, които обаче не е задължително да бъдат попълнени веднага от изследователя. TSM е построена така, че позволява повторно връщане към съответния ръкопис или отделен текст от друг специалист в следващ период от време, като съществува възможност новото допълване, редактиране или преоценка на научните факти изрично да бъдат упоменати. Трябва да се изтъкне, че настоящата схема е **отворена**, което може да се смята за едно от големите ѝ предимства в сравнение с други опити за съставяне на електронни каталози или бази данни.

Както подчертах, настоящият проект има за цел да отговори на съвременната необходимост от нови средства и подходи при установяване на акцентите в литературния живот през Средновековието и на взаимната обусловеност на литературните явления във византийско-славянската общност. Несъмнено литературното или историческото изследване се съпътствува от системна извороведска работа, която би позволила с най-голяма пълнота да бъде очертан фонът, на който възникват произведенията на класическата старобългарска литература, най-старите оригинални жития, служби, похвални слова и т. н. Отдавна се чувства необходимостта от **комплексно изучаване** на старобългарската книжнина — като съвкупност от славянски и от византийски текстове, събрани според достигналите до нас

преписи в средновековните ръкописи. Репертоарът си поставя за цел именно това, използвайки приложението на компютърните технологии.

Биха могли да се очертаят няколко възможни последователни степени на използването на резултатите от електронната обработка на славянски ръкописи като база за нови изследвания:

1. Описание на славянски ръкописи с компютърни средства ⇒ електронни индекси от типа "initia-desiderata" ⇒ конкорданс ⇒ изводи въз основа на сравнителния анализ на данните от палеографски, кодикологически, структурен и текстуален характер;
2. Описание на славянските ръкописи с компютърни средства ⇒ изграждане на база данни ⇒ създаване на електронен тезаурос с хипертекстова организация от типа на Bibliotheca Hagiographica Graeca;
3. Описание на славянските ръкописи с компютърни средства ⇒ присъединяване на "образи" — визуални образци от почерка, украсата и от цялостното оформяне на ръкописа, таблици и др. (чрез сканиране) ⇒ създаване на компютърен продукт от мултимедиен тип⁴.

Разработката на репертоар на старобългарската литература и книжнина ще предложи принципно нови решения на редица проблеми от развоя на историко-литературния процес в средновековна България, като наред с пълния обзор на авторското наследство на старобългарските писатели има възможност да се очертае максимално изчерпателно обема и характера на преводната литература. Репертоарът дава възможност за идентификация на отделни текстове и на цели състави, за изследване на изворите на текстовете, за сравнителен анализ на ръкописите и тяхната типология. Репертоарът е в състояние да регистрира и да сравнява структурите на сложни в жанрово отношение текстове, каквито напр. са химнографските композиции. Резултатите от съставянето на репертоар ще допринесат за разкриване на взаимната обусловеност на литературните явления, жанрове и стилове във византийско-славянската общност. Не на

⁴ Miltenova, A. L., D. J. Birnbaum, A. T. Bojadzhiev 1996: Computer Supported Processing of Early Slavic Manuscripts (Results and Prospects for Textological and Linguistic Analysis). ALLC-ACH '96. The Joint International Conference. University of Bergen, June 25-29, 1996, pp. 206-207.

последно място ще бъде създаден внушителен информационен архив от ексцерпирани старобългарски текстове (заглавие, начало и край), които ще служат като база за по-нататъшни изследвания.

В резултат на настоящия проект се дава едно принципно ново решение на въпроса за обработката на старите славянски ръкописи в електронна форма. Описанието на ръкописите с компютърни средства, съчетано с вмъкването на изображение, би могло да стане в бъдеще основа за създаване на CD ROM, съдържащ аналитично описание на старобългарски ръкописи, или самия репертоар по определени показатели. Това има значение не само като научно постижение, но и като бъдещо пособие за обучение на студенти и подобряване следдипломната квалификация на специалисти.

На следващо място информацията от обработените ръкописи лесно може да бъде пригодена за разпространение по INTERNET на специален за целта WEB SIDE. Тази перспектива без съмнение ще утвърди престижа на България в палеославистиката в световен мащаб. Важният въпрос за стандартизация на информацията в областта на старите славянски ръкописи има решаващо значение за обмен на информационни продукти с други големи Европейски научни и културни центрове.