

электронного представления книжного памятника» // Информационные технологии и электронные ресурсы в библиотеках: Материалы межрегион. науч.-практ. конф. (26–30 сент. 2011 г., Респ. Бурятия, г. Улан-Удэ). Новосибирск, 2011.

Баженов и др., 2011 — Баженов С.Р., Бородихин А.Ю., Шабанов А.В. Цифровая библиотека «Книжные памятники Сибири» как единый информационный ресурс фондов редких книг и рукописей Сибири // Библиотеки и информационные ресурсы в современном мире науки, культуры, образования и бизнеса: Материалы 18-й междунар. конф. (4–12 июня 2011 г., г. Суздак). М., 2011.

Шабанов, 2008 — Шабанов А.В. Факторы, влияющие на выбор технологии оцифровки русских старопечатных и рукописных книг // Библиосфера. 2008. № 4. С. 46–48.

Шабанов, 2010 — Шабанов А.В. Сравнение установок для оцифровки русских старопечатных и рукописных книг и методы обработки изображений // Библиосфера. 2010. № 2. С. 30–32.

## ЛЕТОПИСНЫЙ ТЕКСТ В СЕТЕВОЙ АРХЕОГРАФИИ

*Д. А. Добровольский*

*Национальный исследовательский университет  
«Высшая школа экономики», Москва*

Actually there are several electronic editions of russian chronicles available on the Web. Each of them has its strong and weak sides. However XML-based e-editions seem to be more convenient for everyday scientific practice, because it is much easier to adjust them to permanently changing investigators' needs.

Археография является неременной спутницей исторической науки с момента возникновения последней. На современном этапе археография сталкивается с новыми требованиями, связанными с задачами компьютеризации исторических исследований. Электронная, в том числе сетевая публикация исторического источника не может формироваться как цифровая версия публикации бумажной: такого рода издания должны отличаться от традиционных не только средствами представления материала, но и идеологией, позволяя задействовать все возможности автоматизированной обработки информации, и именно перспективы такой обработки представляются важнейшей характеристикой удачной электронной публикации. Практика показывает, что ученым нужна возможность, во-первых, осуществлять поиск по заданному шаблону, а во-вторых — сопоставлять разные версии одного и того же произведения, получая «на выходе» выровненные параллельные тексты (а если возможно, то и списки разночтений): все прочие исследовательские процедуры (поиск «общих мест» и скрытых библеизмов, текстологическая стратификация и т.п.) могут быть представлены как развитие этих двух

базовых. При этом корпус источников должен быть представлен максимально системно, чтобы обеспечить репрезентативность результатов. На сегодняшний момент в интернете существует ряд специализированных сайтов, основной или одной из основных задач которых является представление текстов такого важного источника по истории средневековой Руси, как летописи. Эти сайты создаются при непосредственном участии научного сообщества. Но помогают ли они в воплощении обозначенных выше ожиданий?

Целостную картину летописания как вида исторических источников дает проект Д.М.Котышева «Полное собрание русских летописей», созданный в Челябинском государственном университете (Режим доступа: <http://psrl.csu.ru/index.shtml>, свободный). С главной страницы сайта доступны два раздела — «перечень томов» Полного собрания русских летописей (далее ПСРЛ), где учтены все существующие на нынешний момент 43 тома данного издания, и список «летописных памятников, не вошедших в состав ПСРЛ». Название второго списка несколько дезориентирует, ибо в нем именованы, среди прочего, Лаврентьевская, Ипатьевская и Новгородская I летопись, изданные в I, II и III томах ПСРЛ, а значит речь идет не столько об источниках, оставшихся неучтенными при подготовке серии, сколько о вне-серийных публикациях летописных текстов. Впрочем, это незначительная неточность, в целом же проект дает энциклопедическое представление о развитии летописного дела от древнейших памятников и до позднейших сочинений XVII — начала XX в. Однако в качестве формы представления текстов создатели сайта избрали постраничное воспроизведение существующих печатных изданий в графическом формате djvu. В текстовом формате даны лишь справки о публикуемых памятниках, заимствованные из «Словаря книжников и книжности Древней Руси», который и без того присутствует в интернете на сайте Института русской литературы РАН (см.: <http://lib.pushkinskijdom.ru/Default.aspx?tabid=2048>). Данное решение полностью соответствует декларируемой автором цели проекта — обеспечить ученых, прежде всего провинциальных, дистанционным доступом к библиотечным ресурсам (ср.: <http://psrl.csu.ru/indexs/develop.shtml>). Но ни задачу поиска по тексту, ни тем более задачу автоматизированного сопоставления разных летописей проект Д.М.Котышева не решает.

Проблема «автоматизации филологических работ» стоит в центре внимания исследователей из Ижевска, поддерживающих сайт «Манускрипт: Древние славянские памятники» (Режим доступа: <http://manuscripts.ru/>, при условии регистрации), где наряду с прочими памятниками раннего периода славянской письменности доступен текст Повести временных лет по шести «классическим» спискам — Лаврентьевскому, Радзивиловскому, Академическому, Троицкому, Ипатьевскому и Хлебниковскому. Пользователь сайта «Манускрипт» получает доступ к запросным формам, с помощью которых можно отобразить из базы данных фрагменты текста, содержащие определенную подстроку; отобранное будет представлено на экране в максимально приближенном к рукописному оригиналу виде. Кроме того, можно знакомиться с дипломатическим воспроизведением отдельных страниц рукописи. Возмож-

ность поиска исключительно ценна для исследователей летописного текста, тем более что существующие словоуказатели [Gröber and Müller 1977–1986; Творогов, 1997] либо содержат ошибки, либо практически недоступны в России. Однако столь же существенной является задача сопоставления летописных текстов, служащего одним из ключевых приемов в восстановлении не дошедших до нас сводов, а сопоставление на сайте «Манускрипт» возможно лишь с использованием параллельного корпуса, и лишь в объеме уже отобранных фрагментов. Системное представление всех разночтений между версиями текста Повести временных лет в число задач ресурса не входит.

Аналогичную модель — вывод фрагментов текста по запросу — реализует посвященный летописям раздел сайта Некоммерческого партнерства «Рукописные памятники Древней Руси» (Режим доступа: <http://www.lrc-lib.ru/index.php%3fid=5>, свободный). Пользователь получает доступ к отсканированным (но не распознанным) страницам четырех томов ПСРЛ (I, II, III и XLIII) и четырем базам данных, содержащим лемматизированные и грамматически размеченные тексты Повести временных лет по Ипатьевскому списку, «Суздальской летописи» (т.е. текста, помещенного в Лаврентьевской летописи после Повести временных лет), киевского летописного свода рубежа XII–XIII вв. и Галицко-Волынской летописи. Возможен поиск (по каждому из четырех текстов), причем в качестве критериев могут задаваться как конкретные формы, так и грамматические характеристики искомым слов; сопоставление текстов разных летописей не предусмотрено. На сайте «Рукописные памятники...» реализовано меньше специальных инструментов, чем на сайте «Манускрипт». Однако этот недостаток компенсируется большим удобством пользователя, так как интерфейс сайта «Рукописные памятники...» интуитивно понятен, построен с использованием только стандартных технологий (для задействования всех функций «Манускрипта» требуется специальный плагин к браузеру) и допускает сохранение результатов работы в формате Microsoft Word. Поскольку же по сути два проекта практически дублируют друг друга, то соображения usability оказываются критичными, предопределяя выбор сайта для повседневного использования.

Принципиально иной способ представления летописного текста реализует проект Д. Бирнбаума «The e-PVL» (Режим доступа: <http://clover.slavic.pitt.edu/pvl/>, свободный), опирающийся на издание Повести временных лет, которое подготовил Д.Островски [Ostrowski s.a.]. Ключевое (хотя не единственное) отличие данного проекта от двух предыдущих состоит в том, что текст издаваемого источника хранится не в базе данных, а в наборе взаимосвязанных HTML (точнее XML)-файлов: один большой файл представляет реконструированный текст начала XII в. (paradosis), а без малого три сотни маленьких — текст Повести по реально дошедшим до нас спискам, разбитый на отрезки, соответствующие столбцам и строкам I тома ПСРЛ, и сопровождаемый текстом основных сводных изданий памятника (А.Ф.Бычкова, А.А.Шахматова, Д.С.Лихачева и Д.Островски). Реконструкцию и оригинальные тексты можно просматривать по отдельности, а можно — совмещенно; в этом последнем случае окно браузера разделяется на два фрейма, содержа-

ших один реконструкцию, а второй — ее источниковую основу. Поисковые механизмы на сайте отсутствуют, лемматизация и грамматическая разметка также не проводилась, хотя полностью открытый формат представления данных позволяет создать соответствующие средства самостоятельно. В итоге, проект Д. Бирнбаума, с одной стороны, содержит практически исчерпывающую информацию о том, в каком виде дошел до нас текст Повести временных лет, а с другой — дает пользователю максимально возможную свободу в оперировании с этой информацией, выгодно отличаясь в данном отношении ото всех рассмотренных выше ресурсов.

Предложенный обзор написан не с целью «отделить агнцев от козлиц». Каждый из рассмотренных сетевых ресурсов обладает своими преимуществами и имеет свои недостатки. Вместе с тем, если говорить об идеальном представлении текста с точки зрения практикующего исследователя, то придется признать — сайты, ориентированные на реляционные базы данных все-таки уступают в удобстве и функциональности сайтам, в основе которых лежит формат XML. Основным критерием в данном отношении выступает приспособляемость к конкретным задачам: если на таких сайтах, как «Манускрипт» или «Рукописные памятники...» круг решаемых вопросов полностью определяется составом и возможностями уже существующих запросных форм, то проект «The e-PVL» дает возможность реализовывать принципиально новые исследовательские программы. Представляется, что будущее сетевой археографии именно за гибкими решениями на основе таких форматов, как XML.

#### **Список литературы**

Gröber and Müller, 1977–1986 — Gröber B., Müller L. Vollständiges Wörterverzeichnis zur Nestorchronik. München, 1977–1986. Lief. 1. – 4. (Handbuch zur Nestorchronik / Hrsg. von L. Müller. Bd. 3.)

Ostrowski s.a. — Povest' Vremennykh Let / [ed. by D. Ostrowski]. Режим доступа: [hudce7.harvard.edu/ostrowski/pvl/](http://hudce7.harvard.edu/ostrowski/pvl/), свободный.

Творогов, 1997 — Творогов О.В. Лексический состав «Повести временных лет»: (словоуказатели и частотный словник): (кроме «Поучения Владимира Мономаха») // ПСРЛ. М., 1997. Т. 1. С. 581–733.

## **КОРПУС ЦЕРКОВНОСЛАВЯНСКИХ ТЕКСТОВ В СОСТАВЕ НКРЯ, ПЕРВАЯ ВЕРСИЯ: ПРОБЛЕМЫ И РЕШЕНИЯ**

***А. Е. Поляков<sup>1</sup>, Е. Р. Добрушина<sup>2</sup>, Т. Ю. Иванова-Алленова<sup>2</sup>***

*<sup>1</sup>НПБ им. К.Д. Ушинского, Москва;*

*<sup>2</sup>Православный Свято-Тихоновский Гуманитарный Университет, Москва*

The paper discusses the issues of spelling, grammatical tagging and metatext markup for a pilot version of the Church Slavonic Corpus (started in 2011) and the prospects of its development