

УДК 004:9:002

## ИНФОРМАЦИОННЫЕ СИСТЕМЫ ДЛЯ ИСТОРИКОВ: ОСНОВНЫЕ ТЕНДЕНЦИИ РАЗВИТИЯ

*Н. Г. Поврозник*

Рассматриваются основные тенденции развития историко-ориентированных информационных систем. На материале реализованных информационных систем для историков прослеживается как содержательная составляющая в их разработке, так и технологическая. В заключении рассматриваются основные функции, которые выполняют информационные системы, а также некоторые проблемы и перспективы их использования.

*Ключевые слова:* историко-ориентированная, источник-ориентированная информационные системы; компьютерные технологии в исторических исследованиях.

Историко-ориентированные информационные системы становятся все более важной составляющей работы профессионального историка, его исследовательским инструментарием. Разработка, создание и развитие такого рода систем является актуальной научной задачей. Рассмотрим основные тенденции их развития.

В основе большинства уже созданных информационных систем лежит технология баз данных, основанная на использовании СУБД (систем управления базами данных), и специализированное программное обеспечение для реализации тех или иных задач функционирования информационной системы в целом. Историко-ориентированные информационные системы в видовом отношении включают в себя и своеобразный тип источник-ориентированных систем.

Первоначально информационные системы для историков создавались как информационно-справочные комплексы, позволяющие, с одной стороны, систематизировать информацию об источниках, а с другой – организовать данные об исторических явлениях, событиях и процессах. Подобного рода системы нашли широкое применение в архивном и музейном деле. В качестве примера можно привести информационную систему «Посольские книги России конца XV–начала XVIII вв.» [1, 15–18]. Она включает базу данных об источниках, хранящихся в фондах Российского государственного архива древних актов (г. Москва) и данные об упоминающихся в делах дипломатических агентах.

Одной из тенденций развития историко-ориентированных информационных систем была трансформация существующих баз данных в более сложные по архитектуре и функциям системы. Для них разрабатывалось специализированное программное обеспечение с целью расширения возможности работы с информацией, ее анализа, визуализации информации и т.д. Классическим и одним из первых примеров подобной трансформации может служить база данных «Древнейшие тюркские рунические надписи» [2, 14–16; 3, 18–19]. Первоначально эта база данных существовала как простая полнотекстовая коллекция электронных версий рунических надписей. В дальнейшем она была технологически обновлена за счет разработки специализированного программного обеспечения для выполнения пользовательских запросов и, тем самым, приобрела качества информационной системы. В этом направлении развивались и другие проекты, например информационно-справочная система «Села и города Среднего Урала в XX веке» [4, 80–104], выполненная в Уральском государственном университете (г. Екатеринбург), или полнотекстовая информационно-поисковая система «Мыто» [5, 93–95; 6, 173–174], посвященная изучению не прямых доходов в государственную казну, созданная в Днепропетровском национальном университете.

Подобное развитие баз данных и трансформация их в информационные системы являются органичным этапом развития информационных технологий и соответствуют логике их использования в гуманитарной сфере. По мере развития информационных технологий и их активного внедрения в работу историка информационные системы для этой отрасли гуманитарного знания приобретают более сложный комплексный характер. Причем сложность заключается не только в многозадачном технологическом аспекте реализации систем, но и в тех научно-исторических задачах, которые та или иная система призвана решать.

Первоначально развитие информационных систем для историков предполагало создание локальных систем, ориентированных на группу пользователей. В этом плане исследователи получали доступ к ресурсам информационной системы через локальную институциональную сеть исследовательского учреждения. Вместе с тем с разработкой дополнительных специализированных программ обеспечивался и удаленный доступ к системе. Согласно этой логике развивалась информационная система «Журналы земских собраний как источник изучения истории местного самоуправления в России (2-я половина XIX – начало XX вв.)»<sup>1</sup>, спроектированная и реализованная в Лаборатории исторической и политической информатики историко-политологического факультета Пермского государственного университета. В основе создания информационной системы лежит база данных, что можно признать традиционным для построения моделей информационных систем. База данных объединила библиографические сведения об источниках (журналах земских собраний) с дополнительной информацией (в т.ч. о месте хранения источника). Вместе с тем при разработке информационной системы ставилась задача создания полнотекстового ресурса, который включал не только различные данные об источнике, но и сам источник, что означало многозадачный характер разрабатываемой системы.

В качестве оптимальной модели доступа была избрана концепция клиент-сервер, а СУБД послужила свободно распространяемая СУБД Fire Bird, которая позволила обеспечить работу системы в локальных и распределенных сетях. Для работы с системой были созданы необходимые приложения и интерфейсы (в том числе интерфейсы ввода-вывода данных), а также расширены пользовательские возможности, в первую очередь обеспечен доступ к системе через сеть Интернет.

Созданная информационная система имеет достаточный потенциал для развития и совершенствования за счет увеличения круга источников. Другой особенностью проектирования системы «Журналы земских собраний...» стала изначальная ее гибкость по отношению к типу репрезентируемого источника, поскольку в дальнейшем система будет развиваться и наполняться разнообразными текстовыми источниками помимо изданий земства. Поэтому разработанная логическая модель информационной системы позволяет сохранять разнотипные источники с отличающейся структурой, определение которой предоставляется пользователям при заполнении базы данных.

Размещение информационных систем в глобальной сети выводит саму систему на новый уровень, обеспечивает ей не локальный, а более широкий корпоративный масштаб пользования. Минимальным требованием для работы с такого рода системами является наличие web-браузера. Их доступность и разнообразие создает простую, не требующую дополнительного программного обеспечения основу для использования системы. Поэтому наличие web-интерфейса становится одним из необходимых признаков вновь создаваемых систем.

В своей деятельности Лаборатория исторической и политической информатики ПГУ сталкивается с необходимостью решения проблем визуализации источников, которые могут иметь не только различный тип, но и самый разнообразный физический формат издания. Так, проблема качественной репрезентации источника возникла при разработке информационной системы «Газета “Пермские губернские ведомости”»<sup>2</sup>. На первом этапе реализации проекта была оцифрована и введена в систему коллекция газет за 1909–1912 гг. Для более целостного сохранения формы и полноты содержания источника и предоставления пользователям возможностей более эффективно использовать источник было принято решение о создании различных видов отображения источника в системе в зависимости от задач пользователя. В качестве оптимального подхода было избрано отображение источника в трех форматах: 1) визуализация имиджа отдельной страницы источника в формате JPEG; 2) отображение отдельной статьи в формате PDF (с опцией распознавания «текст под изображением страницы») и 3) отредактированный текст избранной статьи. Таким образом, была реализована возможность выбора наиболее приемлемого для пользователя формата отображения источника с сохранением облика и структуры издания, а также без потери качественного содержания источника.

Безусловно, историки заинтересованы в сохранении и расширении доступа к полной коллекции этого дореволюционного периодического издания. Однако при планировании «пилотного» проекта разработчикам информационной системы необходимо было предусмотреть как оптимальные способы бережной дигитализации источника, так и определенные возможности дальнейшего развития системы. Работа над информационной системой «Газета “Пермские губернские ведомости”» – это продолжающийся проект, начало которого было связано с необходимостью безотлага-

тельного сохранения источника, изданного в 1909–1912 гг., поскольку из-за низкого уровня сохранности газет существовала угроза полной потери коллекции источника. На настоящем этапе развития системы<sup>3</sup> было принято решение о переводе в машиночитаемый формат тех коллекций источника, для которых существует угроза утраты, в частности номеров раннего периода издания источника (с момента начала его публикации в 1838 г.). До завершения очередного этапа проекта в 2009 г. планируется оцифровка и введение в систему коллекции газет за 1838–1844 гг.

Ввиду многопланового характера газеты «Пермские губернские ведомости», а также высокой востребованности этого источника исследователями создаваемая информационная система служит электронным предметно-тематическим указателем, способным обеспечить удобство поиска по источнику, осуществить запросы к информационной системе в целом в соответствии с предметно-тематическими рубриками и доступом к фрагментам текста.

Программно-технологической основой информационной системы «Газета “Пермские губернские ведомости”» стала СУБД «Oracle», работа пользователей с системой осуществляется с помощью обычного web-браузера.

Развитие созданных информационных историко-ориентированных систем, адаптация их к новым исследовательским задачам являются одними из ключевых компонентов деятельности Лаборатории исторической и политической информатики, которая является одним из научных центров, ориентированных на разработку, создание, оптимизацию и использование информационных систем для историков.

Если на примере развития информационной системы «Газета “Пермские губернские ведомости”» была показана деятельность по постепенному расширению оцифрованной коллекции одного периодического издания, то работа в области исследований по истории российского парламентаризма иллюстрирует несколько иную стратегию развития информационных систем.

Информационная система «Российские парламентарии начала XX века»<sup>4</sup> изначально разрабатывалась как исследовательский инструмент для изучения социокультурных характеристик и деятельности депутатов Государственной Думы. Источниковой основой системы стали указатели к стенографическим отчетам заседаний дореволюционного парламента, которые содержат разнообразную и, вместе с тем, четко структурированную персональную информацию о депутатах и их деятельности в Думе, а также о депутатском корпусе в целом. Важно отметить, что база данных, содержащая характеристики депутатов, имеет просопографический характер, что позволяет получать данные для создания коллективных портретов думских деятелей и изучать их. Информационная система «Российские парламентарии начала XX века» стала важным исследовательским средством для изучения депутатского корпуса Государственной Думы на основе широкого статистического материала, полученного с помощью запросов к системе.

Несмотря на то что информация указателей полностью, хотя и в распределенном виде, сохранилась в информационной системе, разработчиками было принято решение о необходимости сохранения и самого источника в оцифрованном виде, что также было реализовано в процессе работы над проектом.

Важным технологическим достоинством созданного информационного комплекса является использование при программировании необходимых приложений языка PL/SQL, а также работа системы на Oracle Application Server, что обеспечивает кроссплатформенность созданной информационной системы. Использование стандартного web-браузера пользователями системы и ее размещение в сети Интернет расширяют возможности доступа исследователей к ресурсам системы. Эти достоинства системы стали технологической основой ее дальнейшего развития.

Продолжением разработки системы, посвященной российским парламентариям дореволюционной России, стал новый проект, предусматривающий ввод в систему полных текстов стенографических отчетов Государственной Думы за 1906–1917 гг.<sup>5</sup> Таким образом, в настоящее время происходит развитие системы за счет привлечения такого источника, как стенографические отчеты заседаний Государственной Думы всех четырех созывов. Поскольку эти документы являются одним из основных источников для изучения истории дореволюционного парламента, то с переводом их в машиночитаемый формат и размещением в системе становится возможен их компьютеризированный анализ (в первую очередь контент-анализ). Появляются возможности для извлечения скрытой информации, ранее не доступной для изучения традиционными методами. Это означает, что расширение круга привлеченных источников и усложнение исследовательского инструментария за

счет роста вариативности и типов пользовательских запросов, заложенных в системе, позволяют характеризовать модернизированную информационную систему как историко-ориентированный многозадачный комплекс.

Подводя итоги, можно сказать, что информационные системы сегодня являются действенным инструментом для решения разнообразных задач, связанных с сохранением, документированием, репрезентацией и изучением историко-культурного наследия. Современные информационные системы для историков выполняют и другие функции: обеспечивают широкий доступ к историческим источникам (через локальные компьютерные сети исследовательских институтов, учреждений хранения документов и глобальную сеть Интернет); упрощают поиск различной информации по ресурсам на основе вариативных пользовательских запросов к системам (в т.ч. обеспечивают многозадачный информационный поиск); создают основу для изучения тех или иных событий, явлений и процессов (в первую очередь являются незаменимыми для компьютеризированного исследования); формируют машиночитаемые массивы источников (что позволяет анализировать их как «единый текст»).

Одной из основных тенденций развития современных историко-ориентированных систем является их трансформация из групповых по масштабности и охвату задач систем (когда система ориентирована на коллективное использование информации членами рабочей группы или коллектива отдельного научного сообщества) в корпоративные информационные системы, доступ к которым получает широкий круг пользователей.

Тенденция подобной трансформации влечет за собой как бесспорные положительные последствия в виде расширения круга пользователей, имеющих доступ к историческому и культурному наследию, что означает и расширение возможностей изучения тех или иных исторических событий, явлений и процессов, так и некоторые отрицательные моменты. К последним относится практически неконтролируемая возможность спекуляции источниками при неограниченном доступе пользователей к системе, что приводит к организации распределенного доступа к ресурсам. Подобное распределение можно представить двухступенчато: делегирование различных прав доступа администраторам и пользователям. Последние имеют возможность делать пользовательские запросы к системе, извлекать необходимую информацию. Административное звено обладает функцией контроля за действиями пользователей, располагает возможностями для трансформации и развития самой системы. Особо обратим внимание на то, что лишь сложный, распределенный характер пользовательского доступа способен обеспечить эффективную работу информационной системы.

Принимая во внимание многофункциональный и комплексный характер информационных систем, можно утверждать, что само по себе создание и разработка информационных историко-ориентированных систем является важной и актуальной научной задачей, успешное решение которой возможно лишь на основе междисциплинарного сотрудничества специалистов историков и информатиков.

### Примечания

<sup>1</sup> Проект реализован при поддержке гранта РФФИ (грант № 04-07-90283), 2004–2006 гг. Подробнее о создании информационной системы см.: *Корниенко С. И., Масленников Н. Н., Шабалина Д. В.* Журналы земских собраний: проблемы создания информационной системы // *Круг идей: алгоритмы и технологии исторической информатики: Тр. IX конф. Ассоциации «История и компьютер» / под ред. Л. И. Бородкина, В. Н. Владимирова. М.; Барнаул, 2005. С. 153–164.*

<sup>2</sup> Проект «Источниковедческий анализ газеты «Пермские губернские ведомости» (XIX–начало XX в.) на основе применения новых компьютерных технологий» поддержан грантом РГНФ (грант № 05-01-82102 а/У), 2005–2006 гг. Подробнее см.: *Волгирева Г. П., Корниенко С. И., Пигалева С. В.* Источниковедческий анализ газеты «Пермские губернские ведомости» (XIX–начало XX вв.) как основа проектирования и создания информационной системы // *Вестн. Перм. ун-та. Сер. История и Политология. 2007. Вып. 3(8). С. 79–88.*

<sup>3</sup> Проект «Газета «Пермские губернские ведомости» 1838–1844 гг.: проблемы источниковедческого исследования и сохранения историко-культурного памятника» поддержан грантом РГНФ (грант № 08-01-82101 а/У), 2008–2009 гг. Подробнее см.: *Пигалева С. В., Корниенко С. И., Гагарина Д. А., Горбачева Н. Г.* Газета «Пермские губернские ведомости» (1838–1844 гг.): традиционный источниковедческий анализ и информационные технологии // *Вестн. Перм. ун-та. Сер. История. 2008. Вып. 7(23). С. 72–76.*

<sup>4</sup> Проект реализован при поддержке гранта РГНФ (грант № 03-01-12012В). Подробнее см.: *Кирьянов И. К., Корниенко С. И.* Информационная система «Российские парламентарии начала XX века» // Вестн. Перм. ун-та. Сер. История. 2005. Вып. 5. С. 50–55.

<sup>5</sup> Проект поддержан грантом РГНФ (грант № 08-01-12108в), 2008–2010 гг. Подробнее см.: *Гагарина Д. А., Горбачева Н. Г., Кирьянов И. К. и др.* Информационная система «Стенографические отчеты Государственной Думы, 1906–1917»: задачи и возможности комплексного историко-ориентированного ресурса // Информационные ресурсы России. 2008. № 6(106).

#### Список литературы

1. *Богуславский А. А., Рогожин Н. М.* Информационно-справочная система «Посольские книги России конца XV–начала XVIII вв.»: Назначение информационно-справочной системы // Инф. бюллетень Ассоциации «История и компьютер». 1998. № 23.
2. *Васильев Д. Д.* Компьютерные базы данных памятников тюркской рунической письменности I: Эпитафийская эпиграфика Южной Сибири // Базы данных по истории Евразии в средние века. М., 1992. Вып. 1.
3. *Васильев Д. Д., Зайцев И. В., Голованов Е. И.* Гипертекстовая база данных «Древнейшие тюркские рунические надписи» // Инф. бюллетень Ассоциации «История и компьютер». 1998. № 23.
4. *Мазур Л. Н., Бродская Л. И.* Информационно-справочная система «Села и города Среднего Урала в XX веке» // Инф. бюллетень Ассоциации «История и компьютер». 2002. № 29.
5. *Жеребцова Л. Ю.* Создание базы данных «Мыто» // Инф. бюллетень Ассоциации «История и компьютер». 2004. № 32.
6. *Жеребцова Л. Ю.* Полнотекстовая информационно-поисковая система «Мыто» // Инф. бюллетень Ассоциации «История и компьютер». 2006. № 34.

*Дата поступления рукописи в редакцию: 24.10.2009*