

<sup>96</sup> Впервые опубликована в журнале «Русская мысль». № 3, 4 за 1890 г.

<sup>97</sup> См.: Гиппиус З.Н. Мережковский. Он и мы. С. 247. Речь идет о письмах, написанных в тот период, еще до свадьбы, когда они единственный раз в жизни расстались на несколько месяцев.

<sup>98</sup> См.: Мережковский Д.С. Вера. С. 284–285.

<sup>99</sup> Там же. С. 329.

<sup>100</sup> РГАЛИ. Ф. 327. Оп. 2. Д. 13а. Л. 2.

## ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

*А.В. Захаров*

### **Опыт интернет-публикаций массовых источников Разрядного приказа XVII–XVIII веков по технологии баз данных\***

С конца 1990-х годов в русскоязычном секторе сети Интернет благополучно соседствуют и взаимодополняют друг друга разнообразные формы электронных публикаций. Родственная цифровая «природа происхождения» таких изданий и методов компьютерной обработки документальных памятников дополняет перо и фолианты историка новейшими инструментами и в поиске информации по электронным хрестоматийным сборникам, и в качественном исследовании массовых исторических источников.

В узком смысле интернет-публикации – это электронные тексты, гипертексты и базы данных, локально размещенные на сервере и доступные через сеть Интернет. Важно различать их по *актуальности* – новизне сообщаемой информации, по *формату* оцифровки<sup>1</sup> – алгоритмам представления машиночитаемых данных, по *технологиям* – способам организации данных, которые в том числе применяются к конечным процедурам информационного запроса пользователя.

Среди интернет-публикаций источниковой тематики большинство составляют оцифрованные тексты, ранее вышедшие в печатном виде. Достоинства электронного дублирования очевидны в доступности, «компактности» и удобстве последующей обработки информации. Недостатки этого способа также пока чрезвычайно заметны, но относятся к числу преодолимых, например, путем разработки гибких правил интернет-издания. От пользователя иногда ускользает информация об авторской группе публикаторов на всех стадиях проекта, в электронных проектах исчезают необходимые археографические комментарии издателя-редактора, корректура часто незавершена, умалчиваются отличия версий и обновлений публикации и т. п.

Всегда актуальны электронные публикации, вводящие в оборот новые источники, опережающие или не претендующие на печатное воплощение<sup>2</sup>. Интернет-публикация предоставляет автору перманентно расширять и дополнять электронный проект, вносить исправления, устанавливать коммуникационные каналы общения с коллегами, что традиционный подход без печатного переиздания исключает.

Технология организации оцифрованных данных на различных этапах интернет-публикации (ввод, хранение данных, взаимодействие системы клиент-сервер), как правило, включает комбинацию нескольких процессов. На стадии ввода данных и преобразования исходного текста в машиночитаемый вид печатный оригинал последовательно сканируется, распознается, редактируется. При альтернативной публикации материалов,

\* Работа выполнена при финансовой поддержке РГНФ. Проект № 03-01-12035в.

не имеющих типографского тиснения, в лучшем случае возможен быстрый перевод электронного текста в html-вариант. Вероятен и полуавтоматический ввод данных с помощью встроенных функций автозамены таких текстовых процессоров, как Microsoft Word, что особенно важно в отношении к базам данных – особой формы публикации и одновременно технологии на стадиях хранения и взаимодействия системы клиент-сервер<sup>3</sup>.

Одно из наиболее распространенных определений под *базой данных* подразумевает совокупность хранимых в вычислительной системе упорядоченных данных, которые выводятся по различным запросам пользователей в форме и объеме, заданными типом систематизированной и введенной информации. Классификация баз данных динамична и сообщается бурным развитием компьютерного программного обеспечения (ПО), выбором критериев типологии. Неустоявшаяся терминология приводит к последующей критике попыток «стандартизировать» подходы проектирования баз данных<sup>4</sup>.

Известная задача баз данных так называемого фактографического вида состоит в надежном поиске информации о заданных объектах, в том числе текста исторического источника и его характеристик по значениям других атрибутов. Объектами<sup>5</sup> могут быть любые однородные предметы и свойства: имена, адреса, чины, листы документа, цвет чернил и т.п. Количество связей между объектами на несколько уровней выше числа последних и может достигать десятки миллионов. В этом смысле для компьютерной базы данных критичен не объем информации, а корректность установления связей-отношений между разнородными объектами. Вот почему качество поиска информации в базе данных отличается не только оперативностью и максимальной полнотой, но и релевантностью в отличие от «традиционного» просмотра текста по печатным указателям или ключевым словам в электронных изданиях (текстовых файлах). В зависимости от типа созданной базы данных исследователь получает возможность расширить массив обрабатываемой документации и усовершенствовать методику многомерного статистического, лексикографического или просопографического анализа массовых данных источника.

Пояснить особенности интернет-публикации источника, содержащего в том числе просопографические сведения<sup>6</sup> средствами технологии баз данных, поможет описание двух информационных полнотекстовых систем (ИПС)<sup>7</sup>, размещенных на веб-сайте ЧелГУ: «Повесточные сказки думных людей конца XVII – начала XVIII века» (ИПС ROVESTKI)<sup>8</sup> и проектируемой с 2003 года ИПС «Боярский список 1706 года» (ИПС BSPISOK1706). Компьютерная обработка боярских повесточных сказок<sup>9</sup> и боярского списка 1706 года<sup>10</sup> мотивировалась возможностью оперативного поиска информации по широкому спектру параметров и отсутствием печатных публикаций этих памятников. Исследовательские ориентиры двух проектов ИПС оказались тождественны в стратегии адекватного отображения массовых данных и уникальной полнотекстовой информации, структуры источника для реконструкции портрета традиционной правящей элиты в петровскую эпоху и эволюции приказного делопроизводства. Основной «инструмент» каждого проекта – эффективное ПО, способное обеспечить электронный ввод-вывод информации по номеру листа и местонахождению архивного документа, по фамилиям, чинам и другим объектам. Задавалось и качество ИПС редактировать хранимую текстовую информацию и наращивать количественный объем данных.

В ИПС ROVESTKI ставилась экспериментальная задача интернет-публикации источниковых данных об участии 100 думных людей в последних боярских съездах

полного состава, в ИПС BSPISOK1706 – данных о государевом дворе, пофамильном составе, службе и поместных владениях около 4800 высших чинов за 1706 год.

Разработка баз данных этих ИПС, с точки зрения техники проектирования<sup>11</sup>, представляет последовательное конструирование трех моделей: инфологической, дата-логической и физической. Инфологическое проектирование предваряется постановкой целей ИПС со стороны разработчика и потенциальных пользователей и выделения источников, имеющих идентичные формуляры и стабильную структуру документа.

Пример наиболее сложной по структуре ИПС POVESTKI позволяет рассмотреть принципы публикации памятников, терминологический аппарат и многомерные взаимосвязи данных. Основные постулаты технологии и проблематика проектирования баз данных сообщили критерии отбора памятников для интернет-публикации в ИПС – степень однотипности структуры документов и информативной насыщенности о пофамильном составе думцев на боярских «съездах». Типологически делопроизводственный комплекс боярских повесточных сказок был разделен на черновые и белые повесточные сказки, наказания и списки разрядных подьячих<sup>12</sup> из трех архивных дел 1697–1701 годов<sup>13</sup>, выявленных по печатной описи приказных столбцов Разряда (фонд 210 РГАДА).

На стадии отбора памятников проявились особенности выборочной публикации источников, отсекавшей архивные дела другой текстовой структуры, а внутри одного архивного дела фрагменты однотипной структуры, но иного содержания, например, списки стольников, составлявшихся Разрядом по различным поводам. Таким образом, точнее констатировать не столько технологические ограничения, сколько следствия фрагментарности публикации. Сочетания «источникового» и «проблемного» подхода в обработке массовых источников задавались искусственно.

Инфологическое моделирование базы данных заключается в описании предметной области, идентификации объектов и определения связей между ними. Технология баз данных основана на реляционной модели данных, предложенной в 1970 году сотрудником компании IBM математиком Э.Ф. Коддом (Codd E.F.). Вся совокупность массовых данных, сообщаемых выявленными документами, может быть отображена в нескольких программных *реляционных таблицах*<sup>14</sup>. По смыслу они аналогичны «обычной» таблице, но содержат данные нескольких взаимосвязанных характеристик одного объекта, например думных чинов. Ниже в графическом представлении видно, что одна из восьми реляционных таблиц – «Чины» описывает один объект, то есть *одну сущность* всей информации базы данных. Таблица включает два поля-столбца  $A_1$  и  $A_2$  – это *атрибуты сущности* «Чин». Слова, записанные в строках таблицы под индексами «V», называются *значениями* соответствующих *атрибутов*<sup>15</sup>.

Таблица 1

Фрагмент модели реляционной таблицы «Чины»

id	Написание чина в орфографии XVII века ( $A_1$ )	Написание чина в современной орфографии ( $A_2$ )
1	Бояря ( $V_{1-1}$ )	Боярин ( $V_{2-1}$ )
2	Околничие ( $V_{1-2}$ )	Окольничий ( $V_{2-2}$ )
3	Думные дворяня ( $V_{1-3}$ )	Думный дворянин ( $V_{2-3}$ )
4	Думная дьяки ( $V_{1-4}$ )	Думный дьяк ( $V_{2-4}$ )
n	( $V_{1-N}$ )	( $V_{2-N}$ )

Специальное поле «id» показывает *идентификатор*, который состоит из порядковых номеров, задающих код каждой уникальной строке таблицы<sup>16</sup>. Идентификатор необходим для связи атрибутов различных реляционных таблиц согласно структуре источника и алгоритмам поиска информации в базе данных. Так, атрибуты объектов «Фамилия, титул», «чин», «место жительства» связаны между собой с помощью уникальных кодов трех соответствующих реляционных таблиц, что обеспечивает поиск и вывод информации по разным запросам (см. рис. 1 и 2) к объектам других таблиц. Еще одно необходимое понятие – тип поля-столбца, определяющий вид и объем данных, которые можно ввести в атрибут таблицы (числовой, текстовый, дата и т.д.).

Атрибуты объектов базы данных определены, во-первых, по рубрикам (фамилия, чин, место жительства и др.) и фрагментам оригинального текста повесточных сказок (тексты наказов подьячим), во-вторых, по параметрам источниковой атрибуции (датировка боярских собраний, архивное местонахождение документа).

Таблица 2

*Реляционные таблицы, содержащие значения взаимосвязанных атрибутов «черновых» и «беловых повесток»*

Поля (атрибуты) основных объектов	Тип поля	Связь с таблицами
<b>ФИО, титул</b>		
Фамилия, имя, отчество, титул (современная орфография)	Строка (до 100 символов)	Чины, Места жительства
Титул, фамилия, имя, отчество, (орфография источника)		
Чин (современная орфография)	Строка (до 20 символов)	ФИО, титул
Поля (атрибуты) основных объектов	Тип поля	Связь с таблицами
<b>Чины</b>		
Чин (современная орфография)	Строка (до 20 символов)	ФИО, титул
Чин (орфография источника)		
<b>Даты боярских съездов и рассылок повесток</b>		
Дата рассылки повесток	Дата	«ФИО, титул», «Черновые повестки», «Беловые повестки»
Дата боярского съезда		
Текст наказа подьячим (в черн. повестке)	Строка (до 1500 символов)	
Текст записи о съезде (в бел. повестке)		
Номер листа черновой повестки	Целое число	
Номер листа белой повестки		
<b>Архивные фонды</b>		
Фонд, описание, дело	Строка (до 70 символов)	«Черновые повестки», «Беловые повестки», «Даты съездов, рассылок»
Номер листа	Целое число	

Поля (атрибуты) основных объектов	Тип поля	Связь с таблицами
<b>Места жительства</b>		
Место жительства (совр. орфография)	Строка (до 50 символов)	«ФИО», «Архивные фонды», «Черновые повестки», «Беловые повестки»
Место жительства (орфография источника)		
<b>Обстоятельства</b>		
Название рубрики беловых повесток	Строка (до 20 символов)	«Беловые повестки»
<b>Черновые повестки</b>		
Текст пометы	Строка (до 200 символов)	«ФИО», «Места жительства», «Даты съездов и рассылок», «Архивные фонды»
Сокращение пометы	Строка (до 4 символов)	
Археографические примечания	Строка (до 200 символов)	
Номер листа	Целое число	
<b>Беловые повестки</b>		
Текст пометы	Строка (до 200 символов)	«ФИО», «Обстоятельства», «Даты съездов и рассылок», «Архивные фонды»
Сокращение пометы	Строка (до 4 символов)	
Археографические примечания	Строка (до 200 символов)	
Номер листа	Целое число	

Для удобства возможного квантификационного анализа схожим по смыслу пометам присвоены однотипные сокращения<sup>17</sup>, расшифрованные в графической таблице ИПС «Список сокращений». Посредством поля «Археографические примечания» в базе данных введены археографические комментарии к отдельным записям текста источника.

При установлении взаимосвязи между атрибутами таблиц необходимо точно ассоциировать их, то есть установить вид взаимосвязи: «один ко многим» (1:n) или «один к одному» (1:1), или «многие ко многим» (n:n). Например, атрибут «ФИО, титул» связан с атрибутом «Место жительства» как 1:1. То есть фамилия одного человека может быть представлена в рамках одной повестки только под одним адресом рубрики «Место жительства». Атрибут «ФИО, титул» связан с атрибутом «Даты съездов и рассылок» как n:1, иначе говоря, некоторое множество фамилий думных людей включены в листок повесток одного боярского собрания).

Следующий этап – даталогическое проектирование ставит целью трансформацию инфологической модели на конкретную систему управления базой данных (СУБД). Этап включает: выбор СУБД в зависимости от специфики обработки информации и динамики базы данных<sup>18</sup>, трудоемкости, статуса СУБД (лицензионного коммерческого или свободно-распространяемого); выбор языковых программных средств с учетом возможностей сервера, на котором планируется размещение базы данных для локальной сети или Интернет<sup>19</sup>. Разумеется, качество и оптимальность этой процедуры зависит от квалификации и опыта программиста.

На программном уровне архитектура взаимосвязанных реляционных таблиц формируется набором команд, специальными программными языками<sup>20</sup>. Завершающий шаг проектирования базы данных – физическая реализация, которая включает составление и тестирование основных программ (CGI-скрипта, который обеспечивает представление базы данных в сети Интернет, и динамические запросы пользователя к базе<sup>21</sup>). Далее следует процедура заполнения базы данными, их редактирование, дизайн окон отображения, корректировка и тестирование программ. Если представить кратчайшим образом логическую модель, то спроектированные информационные системы состоят из реляционной базы данных (совокупности таблиц, содержащих данные) и поисковой системы, выполняющей запросы на языке SQL к базе данных веб-сервером.

Итак, в чем же состоят ключевые преимущества и возможности ИПС? Массивность документов разрядного делопроизводства практически исключают оперативное введение в научный оборот новых источников через печатные публикации. К тому же распыленность однотипных сведений по архивным делам и фондам затрудняет анализ и дальнейшую интерпретацию текста. Потенциально этих недостатков лишена компьютерная система. Делая поисковый запрос в ИПС, пользователь мгновенно выводит полнотекстовую информацию или полный объем данных по какому-либо атрибуту информационного массива источника. Стоит подчеркнуть еще одно преимущество технологии баз данных над характеристиками печатной публикации и даже электронного гипертекста, которое заключено в свойствах отображения, хранения и потенциальной обработки данных.

В ИПС POVESTKI предусмотрено два варианта вывода оригинального текста документа из базы данных. Первый вариант отображает полистовое содержание документа. По точной архивной ссылке (рис. 1) выводится полный текст черновой повестки – листка, извещавшего думных людей Мясницкой улицы и Лубянки, о съезде в селе Преображенском. В двух правых колонках синхронно представлены сокращения помет и археографические примечания к тексту. Второй вариант поиска – по дате боярского собрания: через указание места жительства или обстоятельства оповещения выводится лишь фрагмент документа.

Выберите фонд, описание, дело

ф. 210, оп. Московский стол. Столбцы, д. 767

Введите номер листа 320

ПОИСК

ф. 210, оп. Московский стол. Столбцы, д. 767 л. 320			
На Мясницкой			
Околничие			
Федор Савич Нарбеков	+ в вотчине своей в Московском уезде в селе Беседе	Д	
Думные дворяне			
Иван Остафьевич Власов	+ сказал самому	СС	
На Лубянке			
Околничие			
Князь Михаил Федорович Жирово-Засекин	живет на загородном своем дворе, с ведомостию хотел к нему итти человек его	Д СЧ	
Думные дьяки			
Андреи Андреев сын Виниус	сказали люди его до повестки съехал в Преображенское	П	ФИО следует читать: "Андреи Андреевич Виниус".

Рис. 1. Фрагмент окна «Фонды РГАДА» ИПС POVESTKI

В ИПС POVESTKI и BSPISOK1706 введены два варианта орфографии (современный и рубежа XVII–XVIII вв.), иллюстрированные цветом фона внутри таблиц. Текст оригинала (1 и 2 колонки слева в таблице на рис. 1) с сохранением аутентичной орфографии размещен на светлых (желтых) оттенках табличного фона. Современный текст (сокращения, примечания, выходные данные документа) передан оттенками темного (красного) фона в окне «Фонды РГАДА».

Пофамильные сведения об участии думцев в боярских съездах выводятся в окно (браузер программы) с перечнем дат боярских съездов и разосланных повесток, упоминающих искомую фамилию (рис. 2).

**Фамилия:** Зотов Никита Моисеевич  
**Чин:** думный дьяк  
**Место жительства:** Пречистенка

**Краткие данные черновых повесточных сказок**

Дата рассылки	Дата съезда	Сокр. черновых
1700-02-09	0001-01-01	П
1700-02-09	0001-01-01	П
1700-02-04	1700-02-05	П
1698-04-04	1698-04-04	СС
1700-02-03	1700-02-05	С
1700-01-31	1700-01-31	П
1699-10-13	1699-10-14	Д

**Краткие данные беловых повесточных сказок**

Дата рассылки	Дата съезда	Сокр. беловых
1700-06-13	1700-06-15	СС
1698-04-26	1698-04-26	+

[Перейти на главную страницу](#)

Рис. 2. Фрагмент окна «Информация по заданной фамилии» ИПС POVESTKI<sup>22</sup>

Процесс проектирования базы данных не должен сводиться к самодостаточному выражению требований реляционной алгебры и устранению избыточности построенных программ. Качество источниковедческой публикации определяет разработка археографического аппарата, оптимальное соотношение восприятия потенциала компьютерного ПО и многомерной ткани источника, скрывающей историческую информацию.

Полное отображение (при выводе в браузер пользователя) текста массовых источников Разрядного приказа рубежа XVII–XVIII веков предлагается передавать по следующим принципам и допускам в базе данных:

1. Структура текста и настройка рубрик (реляционных таблиц).

1.1. Порядок следования рубрик текста и фамилий персоналий внутри рубрик строго задается структурой архивного оригинала. Какие-либо исключения оговариваются в специальном поле «Примечания».

1.2. Написание значений рубрик (напр., названий чинов, улиц), имен и фамилий передается в устоявшейся для конца XVII века орфографии. Все текстуальные казусы, в том числе и ошибки писцов, рассматриваются в примечаниях к тексту.

1.3. Допускается возможность настройки отдельных рубрик при выводе их значений всегда во множественном числе, независимо от фактического количества значений, помещенных под названием рубрики. Например, под чином «Думные дворяня», в

электронном виде документа записан один человек, следовательно, в тексте архивного оригинала название рубрики читается: «Думный дворянин».

Также названия рубрик с потенциально малочисленными записями могут передаваться всегда в единственном числе, независимо от количества значений, помещенных под названием рубрики.

Во всех случаях ошибки писцов в окончаниях слов во множественном и единственном числе фиксируются в примечаниях.

1.4. Значения, не соответствующие атрибутам основных объектов базы данных (как правило, это избыточные данные, напр., имена лиц, чины, исключаемые тематикой базы) записываются в поле «Примечания».

1.5. В поле «Пометы» кроме собственно помет документа допустимо записывать те фрагменты текста, которые не получили учета в наборе атрибутов таблиц.

1.6. Текстовые неструктурированные по атрибутам фрагменты допустимо отображать над таблицей данных. Примечания к этим фрагментам заключены в двойные квадратные скобки.

## 2. Настройка отображения орфографии и пунктуации источника.

2.1. Текст источника передается буквами современного русского алфавита. Не используются буква «й» (введена в 1735 г.) – пишется «и», «ё» (окончательно утверждена в 1956 г.) – пишется «е». Отображение в конце слов буквы «ъ» оговаривается особо для конкретной базы данных. Устаревшие буквы заменяются аналогичными современными: «ижица» на «и», «ять» на «е», «кси» на «кс», «фита» на «ф». Буквы «зело» и «земля» передаются как «з».

2.2. Буквы, нераскрытые в тексте источника в словах под титулом, а также выносные буквы, вносятся в строку текста согласно традиционному чтению слов в конце XVII века.

2.3. Все цифры, переданные в оригинальном тексте буквами кириллической азбуки, в базе данных заменены арабскими цифрами.

2.4. Текстовые древние символы не раскрываются и обозначены в графической таблице «Принятые сокращения». Надстрочные знаки, обозначающие сокращения, краткость звука и ударения, не воспроизводятся.

2.5. Ошибки писца, вставки в текст, зачеркнутые, затертые слова и прочие аутентичные исправления оговариваются в примечаниях.

2.6. Разделение слов и предложений, пунктуация в текстовых фрагментах базы произведены по современным нормам.

Эти правила, первоначально составленные для ИПС POVESTKI, оказались востребованы и в ИПС BSPISOK1706 (рис. 3).

Говоря о недостатках применения технологии баз данных при публикации в ИПС массовых источников приказного делопроизводства необходимо дифференцировать «объективные» ограничения и потенциально устранимые. Первые вытекают из оптимально настроенных программных связей, к ним относится выборочность структурирования объемных фрагментов текста, не содержащих атрибутов, заданных базой. Это и привело к созданию правил представления полнотекстовых данных. Подобное ограничение компенсировано устойчивостью программ и высокой производительностью базы данных. Ограничения «внепрограммного» свойства вполне устранимы, они видятся в неполном представлении массива данных. В ИПС POVESTKI пока представлено около 35 % документов от числа выявленных в архивных фондах, в ИПС BSPISOK1706 – 100 % текста.

Ф.И.О.	Помета	Кол-во дворов	Примечания
РГАДА, ф. 210, оп. 6а Московский стол. Книги, д. 176 л. 1			
Бояря			
[Основной список]			
Князь Петръ Ивановичъ Прозоровской			
Князь Михаило Алегуковичъ Черкасской			
Князь Петръ Ивановичъ Хованской	в Астрахани		
Князь Михаило Григорьевичъ Ромодановской	на службе		
Князь Юрьи Семеновичъ Урусовъ	в Нарве		
Князь Михаило Яковлевичъ Черкасской	в Тоболску		
Князь Борисъ Ивановичъ Прозоровской			

Рис. 3. Фрагмент окна «Фонды РГАДА» ИПС BSPISOK1706

Благодаря первой экспериментальной ИПС была создана основа для дальнейшего изучения массового источника о судьбе традиционных институтов власти и правящей элиты России в петровскую эпоху. Программы ввода-вывода и база данных ИПС BSPISOK1706 станет основой для полнотекстовой публикации боярских списков из делопроизводства Разрядного приказа.

В рассмотренных ИПС реализованы настройки базы данных на оперативное пополнение новыми документами. Они предназначены для широкого доступа к информации исторического памятника в Интернете и нацелены на практическое применение в исследованиях по генеалогии, истории учреждений XVII – начала XVIII века, при создании биографических справочников, в преподавании источниковедения и исторической информатики. По инициативе многих исследователей и при поддержке научных фондов и организаций становится реальным доступ широкого круга пользователей сети Интернет к некоммерческим базам данных, созданным в большинстве случаев пока по массиву известных источников. В настоящее время для профессиональных исследователей ввести в научный оборот массовые источники, в том числе приказного происхождения, видимо, затруднительно без специальных усилий в освоении технологии баз данных.

### Примечания

- <sup>1</sup> Можно выделить интернет-публикации, во-первых, в виде текстовых файлов, с расширением .txt, .doc, .htm и др., передающих текстовые знаки и разметку текста набором команд соответствующего программного алгоритма или языка. Например, наиболее распространенный в Интернете гипертекстовый формат публикаций пишется на языке HTML (HyperText Markup Language). Во-вторых, не менее популярны графические и комбинированные форматы, узнаваемые по расширению .tif, .jpg, .pdf и др. Такие файлы содержат закодированные данные и в том числе текст, который для пользователя передается в графическом виде, целиком копирующем изображение оригинала. Например, pdf-файлы, читаемые программой «Acrobat Reader», внешне напоминают «обычную» ксерокопию.
- <sup>2</sup> В большинстве случаев авторское право защищает процедура государственной регистрации электронных изданий, баз и банков данных, проводимая ФГУП «Информрегистр». См.: [www.inforeg.ru](http://www.inforeg.ru)
- <sup>3</sup> С точки зрения технологии интернет-публикация реализуется с помощью взаимодействия специальной программы (веб-сервера), запущенной на компьютере-сервере, где хранятся данные и другой программы (веб-клиента) компьютера пользователя, который запрашивает необходимые документы из хранилища компьютера-сервера.

- <sup>4</sup> См.: Гарскова И.М. К дискуссии о двух «подходах» // Информ. бюл. ассоциации «История и компьютер». № 19. М., 1996.
- <sup>5</sup> В теории реляционных баз данных используется термин «сущности» и производные понятия: «тип сущности» – род или свойства объектов (напр.: город; цвет), «экземпляр сущности» – конкретные названия «сущностей» (напр.: Москва, Казань; черный, синий). Ср.: Кириллов В.В. Основы проектирования реляционных баз данных. СПб., 1994.
- <sup>6</sup> Просопографическими можно считать массовые данные, характеризующие множество параметров социальной группы, обработка (статистика) которых направлена на создание «коллективной биографии». В ИПС ROVESTKI это данные о месте проживания, чинах, участии в боярских съездах думных людей в 1697–1701 гг.
- <sup>7</sup> Поскольку набор программных средств, позволяющий делать поисковые запросы к базе данных, выводит информацию только по заранее заданным параметрам, обе ИПС необходимо рассматривать как полнотекстовые в информационном, а не программном отношении. В последнем смысле базы данных могут быть настроены на поиск и статистический учет отдельных лексем по всему текстовому массиву базы.
- <sup>8</sup> Начальная стадия проектирования ИПС ROVESTKI под названием «Полнотекстовая база данных и поисковая система «Повесточные сказки думных людей конца XVII – начала XVIII века» поддержана Российским гуманитарным научным фондом. Проект № 00-01-12016в. См.: [www.csu.ru/ourprogram/index.html](http://www.csu.ru/ourprogram/index.html).
- <sup>9</sup> Повесточные сказки – комплекс текущей массовой документации, ведшейся Разрядным приказом для созыва боярских собраний. Он включает черновые формуляры повесток думцам, беловые повестки, пофамильные списки подьячих-курьеров и указы приказным людям; приказную переписку и объяснительные сказки «бояр», пропустивших думное собрание. См.: Захаров А.В. Изучение и компьютерная обработка нового источника о пофамильном составе боярских съездов при Петре I // Вестн. Челяб. ун-та. Сер. 1. История. 2001. № 1. С. 26–29.
- <sup>10</sup> О содержании и информационных возможностях источника см.: Захаров А.В. Уникальный источник начала XVIII века Боярский список 1706 года // Вестн. Челяб. ун-та. Сер. 1. История. 1998. № 1. С. 78–81.
- <sup>11</sup> См.: Гарскова И.М. Базы и банки данных в исторических исследованиях. М., 1994.
- <sup>12</sup> *Черновая повестка* – это небольшого формата листок со списком думных людей одной московской улицы. Подьячие сообщали думным людям о времени и месте боярского «съезда», а пометами напротив фамилии каждого фиксировали факт его оповещения («сказано самому», «сказал, что сего апреля 1-го поедет на Воронеж и о полате быть некогда», «сказ сыну ево Володимеру», «по скаске стряпчего Ивашки, что отпущен на неделю в подмосковную вотчину» и т. п.). *Список подьячих*, рассылавших приглашения, не имел однотипной структуры с повестками, но он также включен в ИПС в полном текстовом формате, содержащем своеобразный наказ курьерам о дате боярского съезда. *Беловые повестки* фиксировали перечень приглашенных на боярский съезд; фамилии думных записывались в рубриках по обстоятельствам оповещения или пребывания на «съезде» («были», «не были», «больны», «в деревнях» и т. п.) и внутри них по чинам.
- <sup>13</sup> Набор текущей документации по учету службы высших чинов государева двора окончательно сложился в конце XVII в. Составление повесток, судя по устоявшейся внутренней структуре записей, практиковалось задолго до петровского царствования. Известно сочинение Джильса Флетчера, английского издания 1591 года с упоминанием о рассылке «повестки писцом Думы Дорофеем Бушевым, который получает приказ из Разряда, или от верхнего констабля, чтоб пригласить их (думных бояр. – А.З.) к назначенному времени». См.: Дж. Флетчер О государстве Русском // Россия XVI века. Воспоминания иностранцев. Смоленск, 2003. С. 55. Прекращение в 1701 году созыва боярских съездов полного состава (то есть всех думцев, находившихся в Москве) не перечеркнуло опыт составления повесток. Во всяком случае, в фонде Разряда встречаются более поздние документы, аналогичные по назначению и структуре повесточным сказкам. Это, например, список «ближних людей», созывавшихся для «славления» на Потешный двор в 1710 году.
- <sup>14</sup> Здесь и далее курсивом приводится соответствующая математическая терминология реляционной теории баз данных.
- <sup>15</sup> Непременное действие инфологического проектирования – определение алгоритмов связи атрибутов: *арности* связей, то есть количества атрибутов объектов и *мощности отношения* (выражается количеством данных в каждом поле).

- <sup>16</sup> Строки таблицы – это *реляционные записи*, которые формируют *кортежи*. Например, кортеж 1, в свою очередь, состоит из двух пар отношения «атрибут-значение»  $A_1:V_{1-1}$  и  $A_2:V_{2-1}$ . Значения  $V_1$  поля «Написание чина в орфографии XVII века» применяется при отображении текста источника (см. рис. 1 и 3), значения  $V_2$  – при выводе из базы «сводных» данных (см. рис. 2).
- <sup>17</sup> Сокращения необходимы в случае поиска «сводных» данных по одной персоналии из всего массива документов, так как полное отображение часто встречаемой пометы «тоже» («тож») может быть бессмысленным вне контекста всего источника.
- <sup>18</sup> В этом отношении СУБД должна быть многофункциональной, многопользовательской, функционировать на множестве моделей компьютеров, поддерживать оптимальное пополнение базы новыми данными.
- <sup>19</sup> В ИПС POVESTKI даталогическое проектирование, физическая разработка и частичная отладка программ выполнены С. Н. Белинским.
- <sup>20</sup> В ИПС POVESTKI и BSPISOK1706 используются языки Perl и PHP. В качестве СУБД выбрана созданная в Калифорнийском университете г. Беркли PostgreSQL в 1996 г. на основе разработки проф. М. Стоунбрейкера (M. Stonebraker) в 1975–1980 гг. СУБД Ingres.
- <sup>21</sup> См.: Павлов А. CGI-программирование: Учеб. курс. СПб., 2000. С. 305–307.
- <sup>22</sup> Принятые сокращения соответствуют пометам: «П» – «в походе», «поехал»; «СС» – «сказано самому»; «С» – «сказано»; «Д» – «в деревне».

## НАУЧНАЯ РЕФЛЕКСИЯ

*И.Э. Любчанский*

### Керамический комплекс и формально-статистические методы его обработки\*

Керамический комплекс является наиболее массовым материалом в системе материальной культуры древних обществ и цивилизаций. Имея свои специфические особенности морфологического, технологического и идеологического характера, он выступает благодатным материалом для исследования отдельных сторон жизнедеятельности отдельных сообществ. В «чистом» виде керамический комплекс представляет «грудку» обожженной глины. Если ее разобрать и упорядочить, ввести в определенные схемы, то эти глиняные черепки начинают приоткрывать завесу тайн не только технологического, но и исторического порядка. Предлагаемая работа посвящена истории изучения методов и методик исследования керамических комплексов в отечественной и зарубежной научной литературе, динамике изменения подходов к его исследованию.

Первая попытка классификации керамической коллекции в русской археологической науке была предпринята А.Е. Теплоуховым в 1884 г. на материалах зауральских стоянок<sup>1</sup>. Керамика была разделена следующим образом: А – черепки, орнаментированные чеканом; Б – разные черепки. Далее черепки, орнаментированные чеканом, делились на десять типов орнаментов и по количеству зубцов на чекане на оттиски 3, 5, 6, 7, 8, 9 зубцов.

Еще одну попытку классификации также на зауральских материалах провели О.Е. Клер и К.И. Фадеев. Их метод основан на анализе способа нанесения орнамента. Все узоры были разделены на 15 типов: 1) линейный; 2) напоминающий веревочный; 3) оттиски в виде змеек; 4) чешуйчатый; 5) крестовой; 6) чекан многозубчатый; 7) чекан изогнутый; 8) треугольный; 9) копытовидный; 10) лунчатый; 11) как бы из насечек; 12) ногтевой; 13) глубоких круглых вдавлений; 14) ямочный; 15) кружковый. Кроме то-

\* Работа выполнена при поддержке гранта РГНФ, проект 02-01-00504а/Т.