

ТЕХНОЛОГИИ БАЗ ДАННЫХ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ

С.В. Акимов (Москва)

БАЗА ДАННЫХ КАК ИНСТРУМЕНТ ИСТОЧНИКОВЕДЧЕСКОГО АНАЛИЗА ФОТОДОКУМЕНТОВ (на примере базы данных «Фотодокументы по истории русско-японской войны 1904–1905 гг.»)

Обращение к проблеме использования и интерпретации фотодокументов в исторических исследованиях вообще и применительно к отдельным периодам и событиям исторического прошлого представляется чрезвычайно актуальным. Среди примечательных событий первых лет XX века выделяется русско-японская война 1904–1905 гг. Фотодокументы по ее истории зафиксировали быт военных лагерей, виды российских, китайских и японских городов, моменты сражений, став оригинальными первоисточниками сведений о деталях жизни всех попавших в район боевых действий.

Фотодокументы, содержащие информацию по истории русско-японской войны, занимают важное место среди других источников по данной теме. Рассмотрение событий войны в литературе до настоящего времени проводилось преимущественно на материале письменных источников. Фотодокументы использовались в ряде случаев как иллюстративный материал. Расширение и качественное изменение источниковой базы исторических исследований по данной теме путем привлечения достаточно представительного массива изобразительных источников открывает новые возможности для изучения истории русско-японской войны 1904–1905 гг.

Для раскрытия информативных возможностей фотодокументов и выяснения степени авторского вмешательства при выборе объектов съемки и в процессе подачи информации потребителю необходимо проведение тематического анализа содержания фотодокументов. Это позволит выяснить какие сюжеты были освещены в большей или меньшей степени, а какие оставались «за кадром».

Фотодокументы по истории русско-японской войны насчитывают более 80 фотоальбомов (фотодокументы которых, зачастую дублируют друг друга) и хранятся как минимум в шести архивах и библиотеках. Большинство этих фотодокументов имеют низкий уровень атрибутированности: не указан автор, дата и место съемки. В связи с этим, для однозначной идентификации фотографии, и возможности сопоставлять информацию, запечатленную на данном фото, с информацией, зафиксированной на других фото, а также в письменных источниках, необходимо подробное описание фотодокумента.

Для того, чтобы отразить максимальное количество информации, предоставляемое фотодокументом для дальнейшего анализа вне стен архива, в СУБД ACCESS была создана база данных «Фотодокументы по истории

русско-японской войны», в которой фиксируется следующая информация о фотодокументе:

- сведения о месте хранения, авторстве и названии из монтажного листа;
- аннотация - подробное описание происходящего на фотодокументе с разделением кадра на передний, средний и задний планы (если это не студийный портрет) - она необходима для получения представления о фотодокументе, особенно если в данный момент нет возможности увидеть изображение;
- примечания - дополнительные сведения, полученные из других источников относящиеся к данному снимку, а также информация о нахождении дубликатов изображения в других альбомах;
- изображение – цифровая копия фотодокумента, если есть возможность его получить.
- географический указатель - позволяет выбрать из списка страну и ближайший населенный пункт. Необходимость его введения возникает в связи с различной транскрипцией корейских, китайских и японских названий, встречающихся в источниках и литературе;
- хронологический указатель - позволяет выбрать из списка год, месяц и число зафиксированного события с учетом разницы в стилях. Также содержит хронологическую таблицу для оперативного определения даты тех или иных событий;
- тематический указатель - необходим для учета «фоновой» информации в кадре (например «Во дворце китайской императрицы в Ляояне. Полковник Попович-Липовац, В.кн. Борис и Хаим Бурбонский.»). Делится на три части: пейзаж, портрет и событие с последующим делением внутри каждой группы.
- жанровый указатель - фиксирует преобладающий жанр фотодокумента (портрет, пейзаж или событие).

БД «Фотодокументы по истории русско-японской войны» содержит 1 основную таблицу и 14 справочников и позволяет:

- максимально полно фиксировать информацию о фотодокументе
- оперативно сопоставлять описание фотодокумента как с другим фотодокументом (в альбоме или опубликованном издании), так и с информацией из письменного источника для атрибуции фотодокументов (одни и те же фотодокументы в разных изданиях или альбомах могут иметь разные заголовки, могут быть кадрированы так, что со снимка исчезает подпись автора, существующая на фото в другом альбоме) (в результате сопоставления с воспоминаниями фотографа А. Любичко удалось установить авторство и дату создания нескольких десятков фотодокументов)
- проводить анализ по автору, месту и дате съемки, тематике и жанру (по отдельности или комбинируя).

На текущем этапе исследования БД находится в стадии внесения данных (на данный момент внесены сведения о более 250 фотодокументах), также дорабатываются функции представления информации.

Таким образом, база данных «Фотодокументы по истории русско-японской войны» с одной стороны позволяет собирать и дополнять, и анализировать информацию о фотодокументах без оперативного доступа к основной массе изображений, хранящихся в архивах и библиотеках, а с другой - ставит вопрос о разработке критериев составления аннотации и классификации фотодокументов для дальнейшего использования (другими исследователями или в качестве элемента НСА архива).

А.М. Архинчев (Москва)

ОПЫТ СОЗДАНИЯ БАЗЫ ДАННЫХ ПО ИСТОРИОГРАФИИ КОРЕЙСКОЙ ВОЙНЫ 1950–1953 гг.

Пятьдесят пять лет назад закончился один из крупнейших военных конфликтов Холодной войны – Корейская война 1950-1953 годов. И спустя столь продолжительный срок процесс по мирному объединению Корейского полуострова все еще продолжается.

Интерес исследователей к этой теме связан и с открытием архивных материалов СССР после 1991 года, которые позволили подойти к проблеме происхождения Корейской войны с новой точки зрения. В результате было написано множество монографий про Корейскую войну, но представленные в них мнения настолько отличаются друг от друга, что в рамках изучения Холодной войны были сформированы несколько направлений в историографии:

1. Советская историография.
2. Реалистическое направление.
3. Ревизионистское.
4. Парадигма "национальной безопасности".
5. Постревизионистское направление.
6. "Революционно-имперская" парадигма.

Для того, чтобы выявить общие черты и различия в характеристике Корейской войны, определить основные взгляды на ее происхождение и периодизацию, свойственные указанным направлениям, в данной работе используется новый подход, который предполагает создание и количественный анализ проблемно-ориентированной базы данных по историографии Корейской войны.

Применение технологий баз данных в исторических исследованиях насчитывает уже десятки работ и проанализировано в ряде обобщающих публикаций. Однако работ, в которых этот подход применяется к анализу историографии, пока немного.

Историографическая реляционная база данных была создана с помощью программного обеспечения Microsoft Access XP. База данных состоит из десяти таблиц – одна основная (Главная) и девять вспомогательных – Тексты,

Список источников, Список литературы, Источники, Литература, Место публикации, Направления, Типы издания и Типы источников.

При анализе направлений в историографии в проблемно-ориентированной базе данных использовались математико-статистические методы количественной истории. Данные методы реализуются с помощью запросов в базе данных. Анализ работ для определения динамики развития историографии и специфики этих направлений в историографии проходил по следующим критериям:

- Число работ.
- Динамика выхода публикаций.
- Место публикации изданий.
- Источниковая база.
- Литература.

Данные критерии позволили провести необходимый анализ историографии. Так, например, удалось выяснить, что по числу работ парадигма «национальной безопасности» преобладает над остальными. Она возникает в 1980-е гг. наряду с «революционно-имперской» парадигмой, заменив реалистическое направление, а уже в 1990-е годы парадигма «национальной безопасности» занимает главенствующую позицию в исторической науке.

При анализе места публикации изданий были выбраны три страны — СССР/Россия, США и Великобритания. Во всех трех странах публиковались работы в рамках этой парадигмы. Также удалось установить, что работы этого направления, в отличие от других, опираются на определенный круг источников, которые используют многие авторы. Это, например, мемуары Н.С. Хрущева («Khrushchev Remembers») и Д. Ачесона («Present at the Creation: My Years in the State Department»), мемуары Г. Трумэна («Memoirs»), некоторые сборники документов.

Как и в случае источниковой базы, удалось выяснить, что существует определенный круг литературы, который используется во многих работах направления парадигмы «национальной безопасности», в первую очередь, монография «The Origins of the Korean War. Vol. II. The Roaring of the Cataract 1940-1950» Б. Каммингса, монография «Международный аспект Корейской войны: некоторые актуальные вопросы» Ю.В. Ванина, ряд других монографий и статей.

Удалось также сделать вывод об определенном сходстве работ «постревизионистского» направления с работами, выполненными в русле «революционно-имперской» парадигмы и парадигмы «национальной безопасности». Наблюдается также влияние советской историографии на работы направления «революционно-имперской» парадигмы.

Таким образом, результаты исследования показывают перспективность подхода, связанного с созданием и анализом базы данных при изучении историографии.

БАЗА ДАННЫХ «МОСКОВСКИЕ БЛАГОТВОРИТЕЛИ 1894–1898 гг.»

При работе над темой «Периодическая печать как источник по истории благотворительности»¹ автор столкнулся с большим количеством публикаций, характеризующих частную благотворительность. Всех их можно разложить на несколько однотипных информационных составляющих: формальные данные (*газета*, в которой опубликовано сообщение, *дата публикации, номер, страница*) и содержание публикации (*название статьи, жертвователь, характеристика пожертвования* – анонимное или подписное пожертвование, – *характеристика жертвователя, организация*, в которую было сделано пожертвование, *сумма, характер передачи денег* – пожертвование, завещание или мероприятие, *цель*, т.е. уточнение, на что именно переданы деньги: установление стипендии в учебном заведении, устройство койки в больнице и т.д., *год пожертвования, месяц пожертвования*). Наличие однотипных данных позволяет структурировать информацию прессы в виде базы данных.

В результате в программе «Access», значительно облегчающей работу с большими объёмами информации, была создана база данных «Московские благотворители». Она состоит пока из восьми таблиц, главной из которых является таблица «Благотворители», а подчинёнными – «Газеты», «Характеристика жертвователя», «Характер пожертвования», «Способ передачи сумм», «Годы пожертвований», «Месяц пожертвований». Последние состоят из двух полей: код и выражения, которые может принимать тот или иной объект. Например, таблица «Месяц пожертвований» состоит из 13 записей: 12 названий месяцев и «Месяц неизвестен».

Главная таблица «Благотворители» состоит, как уже было отмечено из 16 полей (кроме указанных смысловых составляющих представлялось целесообразным задать такие поля, как «Дополнения» и «Аналогичные сведения в других изданиях»²), в которых в формализованном виде представлена информация о пожертвованиях. Всего в таблице 7534 строки, отражающие такое же количество пожертвований. Насколько известно, это самая крупная база данных по благотворителям, существующая на сегодняшний день. Обращает внимание разброс сумм пожертвований: от 5к. до 800000р. Основная масса жертвователей – мужчины. На их долю приходится 3199 пожертвований общим объёмом 10340474р.62к.; на долю женщин приходится лишь 893 случая передачи сумм общим объёмом 5614161р.59к. Нетрудно заметить, что средний размер пожертвований в женской благотворительности значительно выше (6286р.86к.), чем в мужской (3232р.41к.). Это объясняется тем, что организованная благотворительность,

¹ Объектом исследования стали шесть ежедневных московских газет, выходивших на протяжении всего рассматриваемого периода («Московские ведомости», «Русские ведомости», «Московский листок», «Новости дня», «Русский листок», «Русское слово»). Хронологические рамки, 1894–1898 гг. объясняются особым этапом в развитии благотворительности в Москве.

² Данная графа важна для установления «пересекающейся» информации в разных изданиях, т.к. позволяет сопоставить данные о благотворителях в разных изданиях.

т.е. проходящая через разного рода организации (в частности, редакции газет) была уделом, главным образом, обеспеченных женщин. Большая часть женщин могла себя проявлять в старинной, патриархальной благотворительности (через милостыню, уход за больными и т.д.), которая не поддаётся учёту. Коллективные формы благотворительности: пожертвования организаций или групп людей, разного рода мероприятия – вносили не столь существенный вклад в дело благотворения, как инициатива частных жертвователей. Было зафиксировано лишь 1028 случаев коллективных пожертвований на общую сумму в 1446041р.70к.

Собственно пожертвования занимают большую часть от всех случаев передачи сумм: 7034, однако на их долю приходится лишь 6947761р.13к. На втором месте идут завещания, на долю которых приходится лишь 372 случая и общая сумма 10281953р.76к. нетрудно заметить, что средняя сумма по завещанию значительно превышает аналогичный показатель пожертвований: 27639р.66к. против 987р.74к. Что касается мероприятий, то на их долю приходится лишь 128 случаев и общая сумма 228446р.49к. средний размер собранной в ходе мероприятия суммы составляет лишь 1798р.79к. Приведённые данные показывают, что пожертвование в конце XIX в. оставалось наиболее распространённым и демократичным способом передачи сумм на благотворительность; завещания и мероприятия были уделом лишь наиболее обеспеченных слоёв населения.

Обращает внимание также годовая динамика пожертвований. Безусловным лидером по количеству пожертвований оказался март (1318 случаев с общей суммой 2 429 412 р. 06 к.), за ним – декабрь (987 случаев на общую сумму 996 528 р. 62 к.). Их лидерство объясняется близостью церковных праздников – Пасхи (празднуемой в марте-апреле) и Рождества (празднуемой в 25 декабря). Но это лидерство лишь по количеству пожертвований, но не по объёму собранных сумм или среднему пожертвованию. По среднему размеру пожертвования лидирует июль (7478 р. 10 к., против 1844 р. 66 к. в марте и 1041 р. 30 к. в декабре) – месяц, на который приходится лишь 271 пожертвование. Несовпадение объясняется тем, что итоговая сумма, собранная на благотворительность, зависела не столько от количества жертвователей, а желанием их пожертвовать большую или меньшую сумму.

Составленная база данных наглядно демонстрирует, что основным субъектом благотворительности был человек мелкого и среднего достатка, а не богатый предприниматель, как это явствует из историографии. Однако именно на долю богатых жертвователей приходится львиная доля всех собранных на благотворительность средств. При этом одним из основополагающих мотивов благотворительности конца XIX в. оставался религиозный.

Ю.Я. Вин, Д.Е. Кондратьев (Москва)

**СПЕЦИАЛЬНЫЕ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИЕ ПРИЛОЖЕНИЯ
СУБД «ВИЗАНТИЙСКОЕ ПРАВО»:
ПРИНЦИПЫ РЕПРЕЗЕНТАЦИИ И АНАЛИЗА ПОНЯТИЙ И ТЕРМИНОВ ***

БД «Византийское право» содержит написанные на латинском и греческом языках тексты крупнейших законодательных сводов Византии, а также их старославянских рецепций. СУБД обеспечивает ознакомление не только с содержанием текстов полнотекстовых документов, но и с их лексикой, насчитывающей почти для каждого памятника права сотни тысяч словоформ. Важнейшую научную задачу адекватной репрезентации лексики и, соответственно, понятий и терминов византийского права поддерживает ряд специальных программных приложений.

Важнейшая функция СУБД заключается в отображении информации о лексике включенных в массив данных БД источников в форме их полных Словников. Каждая внесенная в них словоформа воспроизводится в оригинальном виде с указанием частоты ее употребления в тексте источника.

Приложение «Сравнение словников», в котором реализованы опирающиеся на теорию множеств и математическую статистику алгоритмы, производит сравнение лексики изучаемых источников. Основные функции приложения «Сравнение словников» предусматривают репрезентацию отличий лексического состава сравниваемых источников путем отображения несовпадающих и общих для них словоформ, а также их суммарного объединения.

Одним из главных программных приложений СУБД является Блок понятий и терминов (БПТ). Он предназначен для систематизации и анализа аутентичных, то есть почерпнутых в текстах изучаемых источников, понятий и терминов. Результатом выполнения утилитарных информационно-поисковых процедур БПТ служит определение фрагментов текстов изучаемого источника, содержащих относящуюся к искомым понятиям и терминам лексику. По мере совершенствования БПТ предполагается фиксация логико-грамматических связей и ассоциативных отношений аутентичных понятий и терминов в высказывании. БПТ нацелен на создание технических условий для проведения изысканий в области текстологической прагматики изучаемого источника. Разработка аналитических функций БПТ направлена на построение Банка данных понятий и терминов и Тезауруса.

Не менее существенны для формирования аналитического потенциала БД Блок «Имена» (БИ) и программное приложения «Экспликация» (ПЭ). Аккумулированные в БИ лексические данные образуют коллекцию «номенов» — отдельных членов собственных имен и названий, выявленных в ходе анализа словников изучаемых источников. В процессе выполнения информационно-поисковых процедур с помощью БИ предполагается выделение фрагментов текстов изучаемого источника, содержащих лексические атрибуты искомого личного имени или названия с учетом их

* Исследование выполняется по проекту РГНФ № 08-01-00186а.

частичной или полной идентификации. По замыслу разработчиков БИ репрезентирует имена собственные и названия независимо от их транскрипции в латинском, греческом и славянском языках. ПЭ поддерживает решение задач по лексико-грамматической систематизации массива данных БД.

Наиболее наукоемкими компонентами СУБД «Византийское право» безусловно являются Модуль определения информационной близости (МОИБ) и Блок когнитивного картирования (БКК). С их помощью производится сравнение текстов изучаемых памятников византийского права и их славянских рецепций на основе методов математической статистики и когнитивного картирования, проецируемого на понятия и термины. Они включены в трехуровневую иерархическую структуру, отображающую логико-грамматические и семантические функции лексики сравниваемых фрагментов текста. На основе этих данных МОИБ выполняет автоматизированное определение весовых коэффициентов анализируемых понятий и терминов и сравнение текстов источников, а БКК поддерживает адекватную атрибуцию понятий и терминов.

Тем самым на повестку дня вынесен вопрос о проектировании Информационно-аналитического программного комплекса, призванного аккумулировать аналитические функции разрабатываемой программной системы и осуществлять машинный анализ текстов источников.

Н.Г. Горбачева (Пермь)

ИСТОРИКО-ОРИЕНТИРОВАННАЯ ПЕРСОНАЛЬНАЯ БАЗА ДАННЫХ «ЗЕМСТВО» КАК СРЕДСТВО ИЗУЧЕНИЯ СОСТАВА ГУБЕРНСКИХ ЗЕМСКИХ СОБРАНИЙ

Изучение социального состава губернских земских собраний является важной и актуальной задачей, поскольку не только государственные законы, но и общественные группы, а также конкретные личности, входящие в состав земских учреждений, влияли на определение земством целей деятельности, а также путей и средств их достижения.

Историография земства представлена достаточно большим количеством научных исследований, определенная часть которых посвящена изучению социального состава земств¹. Важно отметить, что большинство авторов ограничиваются в своих работах рассмотрением лишь сословного происхождения гласных земских собраний, либо их «классовой принадлежности»² (по куримальному распределению на выборах). Тем не

¹ См.: Азаматова Г.Б. Уфимское земство (1874–1917): Социальный состав, бюджет, деятельность в области народного образования. Уфа, Изд. «Гилем», 2005; Живодрова С.А. Пензенское земство в 1900–1914 гг. Социальный состав, бюджет и практическая деятельность. Автореф. дисс. ... канд. ист. наук. М., 1996; Двухжилова И.В. Земство Тамбовской губернии: Социальный состав и практическая деятельность (1865–1890). Автореф. дисс. ... канд. ист. наук. Пенза, 2002 и др.

² Львова .Н. Московское земство в 1865–1890 гг.: Автореф. дисс. .. канд. ист. наук. М., 1968. С. 9.

менее, важно проанализировать и другие аспекты, например, показать возрастные характеристики гласных, раскрыть образовательный уровень, увидеть профессиональный опыт и компетенции депутатов земства. Целью данной работы является изучение социального состава гласных губернских земских собраний Московской, Петербургской и Пермской губерний в 1864–1890 гг. Это значит, что для выявления социально-культурного облика депутатов необходимой основой является комплексный анализ состава гласных земских собраний. Кроме того, построение коллективных портретов земских гласных позволит обнаружить общие черты развития земского представительства в России и специфику корпуса гласных земских собраний разных губерний.

Очевидно, что лишь традиционными методами анализ совокупности социальных характеристик земских гласных затруднителен, поэтому использование компьютерных технологий, в частности баз данных, будет способствовать решению сложных исследовательских задач на более высоком доказательном уровне.

Анализ состава земских учреждений проводился на основе созданной реляционной базы данных "Zemstvo" под управлением СУБД Access. База данных включает поименованный список членов губернских земских собраний указанных губерний. База состоит из 11 таблиц (основная – "Personal" и 10 вспомогательных). Таблица "Personal" включает поименованный список гласных, описание которых структурируется по 19 полям. Созданная база данных отражает следующие характеристики гласных: ФИО, даты жизни, образование, губерния, год избрания в управу, уровень членства в управе (губернский или уездный), тип членства (председатель управы или рядовой член), уезд (от которого гласный избран в собрание), заседание губернского земского собрания (в котором принимал участие депутат), избрание в комиссии, тип участия в земском собрании (председатель, гласный, представитель ведомства и др.), год избрания гласным, сословная принадлежность, чин, членство от земства в организациях, внеземская деятельность и некоторые другие. База данных содержит полный список гласных губернских земских собраний Московской, Петербургской и Пермской губерний за 1864–1890 гг., что составляет 1080 записей.

Для проведения анализа осуществлялись запросы к базе данных, как по отдельным параметрам, так и по условиям, включающим различные комбинации характеристик гласных.

В итоге исследования были получены следующие данные:

- результаты анализа подтвердили некоторые выводы предшествующих исследователей о соотношении социальных слоев и о динамике сословного представительства в губернских земских собраниях;
- ряд полученных данных подвергают сомнению укоренившееся в земской историографии мнение о все усугубляющемся кризисе в земстве в 1870–1880-е гг., связанном с плохой посещаемостью заседаний в это время. Было показано, что регулярность посещений определенным кругом гласных была практически одинаковой. Анализ социальных

характеристик этих депутатов позволяет говорить о складывании «профессионального» слоя земцев.

- определенная часть результатов была получена впервые, например, о влиянии опыта гласного, возрастных и образовательных характеристик, а также ряда компетенций на его избрание в земство. Кроме того, сравнение социально-культурного облика гласных губернских земских собраний рассматриваемых губерний позволило выявить их общие и особенные черты.

Таким образом, использование информационных технологий, в частности базы данных, позволило найти новые аргументы для доказательства ряда существующих в исторической науке положений по истории земства, подвергнуть сомнению некоторые устоявшиеся точки зрения, связанные с социальным составом губернских земских собраний, а также привнести новое знание с доказательной базой высокого уровня.

Л.Ю. Жеребцова (Днепропетровск)

БАЗА ДАННЫХ «МЫТО» КАК ОДИН ИЗ МЕТОДОВ ИЗУЧЕНИЯ СТРУКТУРИРОВАННЫХ ИСТОЧНИКОВ

При изучении действующих на украинских землях Великого княжества Литовского в конце XV – середине XVI века пошлин, на этапе их систематизации мы составили простую таблицу, в которой фиксировали название пошлины, ее определение и документы с информацией о пошлине. Эта таблица органично подтолкнула к идее создания более сложной организации данных с учетом специфики используемых источников. Основное задание БД состоит в сохранении максимально полной информации об известных пошлинах и документах, в которых они упоминаются. Поставленные задачи обусловили использование объектно-ориентированного подхода к созданию БД, который позволил «избежать» чрезмерной концентрации на самих видах пошлин или на источниках.

Использование в исследовании документов актового характера (книг записей, книг судебных дел, книг публичных дел и книг переписей Литовской Метрики), имеющих «условный формуляр», который удалось определить методами дипломатического анализа, определило логическую модель организации данных. Такие источники имеют четкую внутреннюю структуру, что делает их наиболее удобными для отображения в формате БД.

Для организации данных была выбрана реляционная модель, которая состоит из 28 таблиц-отношений. Для установления связей между таблицами мы использовали тип связи многие ко многим, поскольку в одном документе может фигурировать несколько пошлин, а пошлина определенного вида встречается в нескольких документах. Для создания БД была выбрана СУБД MS Access, для интерфейса – Visual Basic .NET. Основу БД составляют 5 блоков, которые отражают информацию: о 42 видах пошлин – «Пошлина»; библиографическом описании опубликованных – «Библиография» и

неопубликованных источников – «Архив», «Дело»; характеризуют признаки использованных документов – «Документ». Наиболее ценным в БД «Мыто» является поиск видов пошлин по нескольким критериям одновременно. Спроектированная в БД форма «Поиск» позволяет решать ряд заданий, решение которых традиционным способом требует больше времени и усилий:

- осуществлять поиск многомерных объектов «пошлина», «источник», «документ», «территория», «лица» по нескольким, предварительно заданным, критериям одновременно. Например, можно получить данные о пошлине *«мыто»*, которая встречается в *8 книге записей Литовской Метрики* в документах вида *«лист»*, выданных во время правления короля *Сигизмунда I Старого*.
- пользователь сам может определять жесткие или гибкие условия поиска. В БД автоматически заданы жесткие условия поиска – «и», т.е., если в поле «пошлина» пользователь выбирает для поиска пошлины – мыто и капицзна, то в результате отобразятся документы, которые содержат эти 2 вида пошлин одновременно. Гибкие условия поиска в режиме «или» предусматривают получение в результате поиска всех документов, в которых встречаются названные виды пошлин.
- результаты поиска отражаются в том же окне, где задаются его критерии. Особенность состоит в том, что даже если пользователь желает получить информацию по какому-то одному критерию, например, только о количестве документов с пошлиной мыто, то БД автоматически выполнит и отразит результаты поиска и по другим критериям (которые не являются специальными объектами поиска).

Предложенная БД является одним из вариантов для изучения системных объектов, к которым относится и таможенная система, на основе массовых источников, имеющих четкую внутреннюю структуру.

И.К. Кирьянов, С.И. Корниенко, И.В. Рябухин, А.В. Сметанин (Пермь)

ИНФОРМАЦИОННАЯ СИСТЕМА «СТЕНОГРАФИЧЕСКИЕ ОТЧЕТЫ ГОСУДАРСТВЕННОЙ ДУМЫ, 1906–1917» *

Создание тематических историко-ориентированных ресурсов на основе комплексов компьютеризированных источников – актуальная задача информатизации и создания современной информационной среды исторической науки и образования. На ее решение направлена разработка и реализация проекта создания и поддержки информационной системы по истории парламентаризма в России в начале XX в., который был начат кафедрой новейшей истории России и лабораторией исторической информатики историко-политологического факультета Пермского государственного университета.

* Проект выполняется при поддержке РГНФ – грант РГНФ № 08-01-12108в., 2008-2010 гг.

Начало этой работе положило создание информационной системы «Российские парламентарии начала XX века»¹, содержащей машиночитаемые версии такого важного комплекса исторических источников по истории Государственных Дум дореволюционной России, как указатели к стенографическим отчетам. На этом этапе была спроектирована и частично реализована модель информационной системы, в электронный формат переведены указатели к стенографическим отчетам Государственной Думы I – IV созывов, и в информационную систему введены данные указателей, содержащие социокультурные параметры и показатели парламентской активности депутатов.

В настоящее время эта работа получила продолжение. Проект «Информационная система «Стенографические отчеты Государственной Думы, 1906–1917» предусматривает перевод в электронный формат и создание машиночитаемых версий другого основополагающего источника по парламентской истории России – стенографические отчеты Государственной Думы России 1906–1917 гг. на основе их ввода в информационную систему.

Как и ранее, в основу решения задачи положен источник-ориентированный подход, позволяющий осуществлять многозадачные исследования с использованием современных информационных технологий. Источник-ориентированный подход позволяет, с одной стороны максимально полно передать облик и структуру изданий, с другой, обеспечить содержательную полноту источника и удобный доступ к информации. Кроме того, обозначенный подход предполагает определенную универсальность создаваемой информационной системы с точки зрения ее научного использования.

Осуществление данной части проекта позволит создать уникальный историко-ориентированный источниковый комплекс, основанный на модели, обеспечивающей взаимосвязь между двумя группами основополагающих исторических источников по рассматриваемой проблематике и позволяющей изучать их один через другой и синхронно. Таким образом, в рамках системы впервые будет обеспечена возможность анализа указанных источников как «единого текста», что позволяет говорить о возможности существенного раскрытия их информационного потенциала.

Ввод в информационную систему полных текстов стенографических отчетов Государственной Думы I – IV созывов станет существенным шагом в создании тематического ресурса по парламентской истории России, создаст более благоприятные возможности изучения этих относительно редких и труднодоступных на сегодняшний день исторических источников.

Как и на предыдущем этапе, программирование информационной системы будет осуществляться на основе использования платформо-независимого языка. Необходимые приложения будут созданы на PL/SQL и запущены на Oracle Application Server, что обеспечит полную кросс-платформенность создаваемой системы. Благодаря такой технологической и

¹ Грант РГНФ № 03-01-12012В, 2003-2004 гг.

программной основе вся работа по дальнейшему развитию системы возможна через web-браузер с использованием сети Интернет.

В процессе реализации проекта будет осуществлена оцифровка всех доступных томов стенографических отчетов, распознанные и отредактированные тексты которых будут вводиться в информационную систему в форматах PDF и текстовом. В связи со своеобразием структуры стенографических отчетов будет модифицирована концептуальная и ER-модель системы, созданы новые интерфейсы и приложения. Будут дополнены и созданы вновь и необходимые справочники.

Р.Б. Кончаков (Тамбов)

ИНФОРМАЦИОННАЯ СИСТЕМА

"Результаты первой всеобщей переписи населения Российской империи 1897 г."*

Изданные в 112 томах результаты Переписи населения 1897 г., по мнению многих историков, являются исключительным по полноте охвата и степени репрезентативности срезом жизни Российской империи конца XIX в. Однако, возможности этого источника в полной мере не были использованы из-за трудоемкости работы с большими объемами статистического материала. Конечно, материалы результаты переписи привлекались для характеристики развития и сравнительного анализа отдельных регионов, благодаря наличию изданных томов сводных результатов с обобщающими данными по губерниям и областям Российской империи, но сравнительный анализ на уровне показателей отдельных регионов представлялся очень трудоемким. В тоже время мировая практика показала высокую эффективность создания баз и банков данных результатов национальных переписей. Многие из таких переписей доступны исследователям через Интернет, а некоторые особенно обширные ресурсы предлагают возможности on-line анализа данных. Сказанное относится в первую очередь к проекту IPUMS.

Проект по оцифровке результатов переписи 1897 года предусматривает создание доступной через Интернет информационной системы, содержащей сведения переписи по всем регионам Российской империи. Эти сведения будут дополнены картами регионов (географический интерфейс) и системой статистического анализа данных он-лайн (статистический интерфейс). Часть таблиц будет введена в базу данных, а другая часть будет доступна в виде индексированной по ключевым словам библиотеки изображений.

Важным этапом является реализуемое в настоящее время формирование базы данных. Эта работа подразделена нами на несколько направлений:

А) формирование электронного архива отсканированных томов.

Б) создание информационной модели и разработка эффективного интерфейса базы данных.

* Работа ведется при поддержке фонда РГНФ, проект № 08-01-12110в.

В) формирование группы операторов и администраторов базы данных.

На первом этапе была создана база данных отсканированных изображений страниц томов результатов переписи населения.

Ввод материалов переписи в базу данных осуществляется вручную оператором с электронного изображения листа.

В качестве базовой платформы нами была выбрана СУБД MS Access 2003, конечно, впоследствии основным хранилищем данных станет формат некоммерческой СУБД MySQL.

Для операторов был разработан специальный интерфейс, спроектированный с учетом большого опыта Лаборатории Социальной истории ТГУ им. Г.Р. Державина по созданию исторических баз данных. Оператор вводит данные лишь при помощи мыши, практически не обращаясь к клавиатуре, что позволяет добиться высокой скорости ввода и минимизировать ошибки в сжатые сроки после начала работы. Контроль ввода осуществляется администраторами базы, каждый из которых закреплен за определенной группой операторов. В то же время детальный анализ структуры таблиц переписи позволил осуществлять контроль ввода в автоматическом режиме, когда программа по завершению ввода оператором группы значений определяет соответствие введенной суммы итога группы и расчетное значение суммы этой группы.

Структура таблиц переписи, в полной мере отвечает традициям статистики конца XIX века. Чаще всего это многоуровневые группировки с итогами по каждой группе. Несомненно, удобный для человека, такой способ представления данных требует так называемой декомпозиции перед описанием его в реляционной СУБД. Мы представили каждую таблицу в виде двумерного массива данных, заменив заголовки столбцов их порядковыми номерами, а к заголовкам строк добавили специальные маркеры групп и итогов. Это дало нам возможность заменить многоуровневые группировки их метаописаниями, конкретные значения которых определяются номером колонки, номером таблицы и номером тома.

Таким образом, этот этап работы на конкретном примере формулирует более крупную проблему оцифровки статистических источников конца XIX – начала XX вв.

С.И. Корниенко, Д.А. Гагарина, Н.Г. Горбачева, С.В. Пигалева (Пермь)

**ГАЗЕТА «ПЕРМСКИЕ ГУБЕРНСКИЕ ВЕДОМОСТИ» 1838–1844 гг.:
ПРОБЛЕМЫ ИСТОЧНИКОВЕДЧЕСКОГО ИССЛЕДОВАНИЯ И СОХРАНЕНИЯ
ИСТОРИКО-КУЛЬТУРНОГО ПАМЯТНИКА ***

Газета «Пермские губернские ведомости» являлась одним из первых официальных периодических изданий в Пермской губернии и относится к важнейшим историко-культурным памятникам и историческим источникам

* Проект реализуется при поддержке РГНФ – грант РГНФ № 08-01-82101 а/У, 2008–2009 гг.

для изучения истории Пермской губернии, Урала и России в целом. Ее первый номер вышел в свет 8 (20) января 1838 г.

На сегодня сохранились неполные и немногочисленные коллекции в архивах, библиотеках и музеях, возникли проблемы с физической сохранностью этого историко-культурного памятника. В настоящее время это крайне затрудняет доступ к этому историческому источнику. Необходимость сохранения данного историко-культурного памятника, обеспечение возможности его изучения делает актуальным перевод коллекции с бумажного на современные носители. Одновременно это позволяет решать задачи применения методов научного исследования, основанных на компьютерных технологиях.

Совместная работа в этом направлении была начата лабораторией исторической и политической информатики Пермского государственного университета и Пермским краевым музеем в 2005 г.¹ В электронный формат была переведена часть коллекции за 1909–1912 гг., находящаяся в угрожающем утратой источника состоянии. Кроме того потребности обеспечения доступа и наиболее полных возможностей для изучения коллекции обусловили целесообразность создания полнотекстовой источник-ориентированной информационной системы. Созданная информационная система сразу проектировалась и разрабатывалась с учетом возможности ввода в нее в дальнейшем всей коллекции газеты, однако ориентирована была в первую очередь на издания 1909–1912 гг.

Исходя из потребностей сохранения других частей коллекции, расширения возможностей их использования в научных исследованиях и образовательно-просветительной деятельности, в настоящее время реализуется проект, направленный на создание электронных копий и ввод в информационную систему номеров газет от начала ее издания (1838 г.) до 1844 г.

Коллекция за указанные годы отличается от последующих изданий формально и содержательно, номера обладают особой структурой и своеобразием разделов, имеются многочисленные «Особые статьи» и «Прибавления», объем которых иногда достигает 60 листов. Нельзя не отметить и тот факт, что коллекция, на основе которой выполняется проект, содержит многочисленные рукописные пометки и подписи, сделанные представителями губернской администрации. Они также представляют интерес, как для понимания содержательной стороны газеты, так и в качестве самостоятельного исторического источника, и поэтому нуждаются в сохранении.

Очевидно, что особенности формы и содержания изданий требуют определенной адаптации и модернизации информационной системы. В частности, создания механизмов для отображения в ней формальных и содержательных структурных элементов номеров газеты 1838–1844 гг. Необходимо также создание в системе новых и дополнение данными

¹ «Источниковедческий анализ газеты «Пермские губернские ведомости» (XIX – начало XX вв.) на основе применения новых компьютерных технологий», грант РГНФ № 05-01-82102 а/У, 2005–2006 гг.

существующих вспомогательных справочников, модернизация форм для визуализации информации.

Кроме того своеобразие изданий предопределяют и специфику в решении задач их дигитализации, сохранения, и визуализации. Реализация проекта предполагает дальнейшую разработку средств и методов оцифровки и распознавания текста с учетом особенностей шрифтов, грамматики, существующих физических дефектов и т.д.

Таким образом, реализация проекта представляет интерес не только с точки зрения решения практических задач сохранения ценного историко-культурного памятника и обеспечения возможностей его изучения, но и дальнейшей разработки таких актуальных проблем исторической информатики как создание полнотекстовых источник-ориентированных информационных систем, информационных моделей исторической периодики XIX – начала XX вв., средств их реализации и других.

П.А. Корчагин (Пермь)

ПСИХОЛОГИЧЕСКИЙ ПОРТРЕТ ПЕТРА I. К ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ БАЗ ДАННЫХ В ИСТОРИКО-ПСИХОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

В исторической науке оценка личности Петра I всегда тесно увязывалась с оценкой его реформ. Западники давали оценки личности царя комплиментарные, подчеркивались его положительные качества (любопытность, трудолюбие). Славянофилы, напротив, оценивали Петра как тирана. Но до сих пор в науке не делалась попыток проанализировать личность Петра I в соответствии с принципами научной психологии. Существует насущная необходимость элиминировать субъективность традиционного подхода к личности императора. Возможность изучения психологии человека без самого живого человека обоснована В.А. Шкуратовым. «Чтобы связать человека с его прошлым и будущим, надо выйти на просторы большой истории. И метод наблюдения заменяется анализом результатов сделанного этим человеком»¹.

В основе настоящего исследования лежат культурно-историческая теория Л.С. Выготского и теория деятельности А.Н. Леонтьева, рассматривающие психическое развитие как результат собственной активности ребёнка в рамках социально-культурной среды. Деятельность каждого отдельного человека зависит при этом от его места в обществе, от условий, выпадающих на его долю, от того, как она складывается в неповторимых индивидуальных обстоятельствах. Помимо этого в каждой возрастной категории отмечается ведущая деятельность (которая впервые была выделена Д. Элькониним), именно в ее процессе складываются психические новообразования, имеющие значение для всей последующей жизни, из которых состоят «этажи» структуры человеческой личности. В работе использована карта

¹ Шкуратов В.А. Историческая психология. Ростов н/Д., 1994. С. 127.

возрастов Н.В. Коптевой¹ и возрастную классификацию В.Ф. Моргуна², рассматривающие периоды от младенчества до юности (от 0 до 18 лет) и зрелость (от 18 до 90 и более лет) соответственно.

Нами разработана интегрированная методика историко-психологического лонгитюда. С одной стороны, в психологии - лонгитюд - основывается на длительном наблюдении и постоянном использовании функциональных психологических проб, диагностирующих различные психические образования. Преимущества этого метода заключаются, в более высокой точности (по сравнению с методом поперечных срезов) прогноза психического развития отдельного индивида и дальнейшего хода психической эволюции вообще; в возможности определения генетических связей между фазами психического развития и установления диапазона колебаний возрастных норм и моментов перехода одной фазы в другую; в устранении такого недостатка метода поперечных срезов, как нивелировка индивидуальных социально-психологических, психологических особенностей и уравнивание все обследуемых индивидов данного возраста и данной популяции.

С другой – лонгитюд, «опрокинутый» на историю, по форме очень напоминает... летопись, которую достаточно легко формализовать с применением любой СУБД, в том числе и MSO Access. База данных «Петр I» была составлена из датированных событий личной и общественной жизни Петра (от рождения до смерти), которые в изобилие содержатся в работах М.М. Богословского «Петр I. Материалы для биографии»³, В. Бергмана «История Петра Великого»⁴ и мн. др. Все записи (в начальном варианте 894) были классифицированы по возрастным периодам и соответствующим им типам ведущей деятельности (подчиненная табл.), что позволило детально анализировать содержание любого отдельного жизненного периода.

Развитие личности Петра I удивительно точно соответствовало возрастным особенностям. К концу юности, когда заканчивается формирование характера, Петр стал вполне цельным, адекватным человеком, успешно и вовремя прошедшим все стадии становления личности. И 1695 г., начало взрослой жизни, ознаменовалось первыми самостоятельным шагом в государственном управлении – Азовским походом.

Нельзя не отметить еще один момент личностного развития Петра – это отсутствие в его жизни возрастных кризисов. Они проявляются как застой в деятельности, но у Петра таковых не отмечено. У Петра хорошо (почти идеально) прошли процессы самоопределения, самовыражения и самосовершенствования. По достижении одной цели, он ставил себе следующую, потом другую... Он был настолько поглощен своим делом, что сама возможность возникновения кризисов отпадала.

¹ Коптева Н.В. Карта возрастов. Психическое и личностное развитие от 0 до 18 лет. Пермь, 1995.

² Кулагин И.Ю., Колоцкий В.Н. Возрастная психология. М., 2001. С. 140–141.

³ Богословский М.М. Петр I. Материалы для биографии. Т. 1–3. М., 1940–1946.

⁴ Бергман В. История Петра Великого. Т. 2–6. СПб., 1833–1834.

Важно отметить, что нервное потрясение, полученное Петром во время стрелецкого мятежа, хотя и вылилось в неоправданно жестокое преследование стрельцов (заметим, исключительно московских), впоследствии не вылились в жестокость как черту характера. И хотя, в историографии распространены суждения о безжалостности Петра, но, кроме 1698 г., случаев неадекватного поведения императора отмечено не было.

В.А. Куликов (Харьков)

ПРОСТРАНСТВЕННАЯ БАЗА ДАННЫХ «Промышленность украинских губерний Российской империи в начале XX в.»

Одним из ценнейших источников по истории промышленности Российской империи являются всероссийские промышленные переписи, которых в начале XX в. было проведено три – в 1900, 1908 и 1910–1912 гг. Каждая перепись дает массовый статистический материал о состоянии промышленности Российской империи.

В докладе будут представлены результаты анализа базы данных по обрабатывающей промышленности девяти украинских губерний Российской империи в начале XX в. по материалам переписи 1900 г. База данных построена на основе публикации «Список фабрик и заводов Европейской России» (СПб., 1903). Эта публикация является перечнем фабрик и заводов, распределенных по 12 группам производств. Список фабрик и заводов составлен по материалам, собранным Отделом промышленности Министерства финансов через фабричных инспекторов. Перепись 1900 г. охватила производства обрабатывающей промышленности, не обложенной акцизом, и только заведения, подчиненные надзору фабричной инспекции. Также не были включены в обследование предприятия с годовым объемом производства до 1 тыс. рублей. То есть значительное количество мелких предприятий не было охвачено переписью. Тем не менее, материалы переписи дают возможность судить о структуре и размещении обрабатывающей промышленности.

Создание базы данных осуществлялось в СУБД MS Access™. База данных состоит из 17 полей: «Название», «ФИО владельца», «Социальный статус владельца», «Форма собственности», «год основания», «группа», «отрасль», «объем выпускаемой продукции», «число рабочих» и т. д. и более 1000 записей – отдельных предприятий девяти украинских губерний (Волинская, Екатеринославская, Киевская, Подольская, Полтавская, Таврическая, Харьковская, Херсонская, Черниговская).

Созданная база данных позволяет определить структуру обрабатывающей промышленности, установить степень концентрации предприятий и рабочих, распределение владельцев по сословиям, выяснить динамику возникновения промышленных предприятий и т. д.

«Список фабрик и заводов Европейской России» по каждому предприятию дает информацию о точном местонахождении предприятия,

что позволило визуализировать пространственное распределение предприятий.

Пространственная база данных реализована под управлением геоинформационной системы ArcGIS™. В качестве базовой пространственной информации использовались карты из Большого всемирного настольного атласа Маркса (СПб., 1910).

Предприятия, расположенные в сельской местности привязывались к центрам уезда, так как часть мелких населенных пунктов на картах атласа Маркса не обозначены; предприятия, находившиеся в губернских и уездных городских центрах обозначенных на карте привязывались к этим точкам.

Пространственная база данных дает возможность проследить географическую специфику (размещение) предприятий обрабатывающей промышленности, выявить промышленную специализацию территорий, концентрацию промышленных предприятий, влияние географического фактора на развитие промышленности.

М.Г. Кучинский (Москва)

ИНФОРМАЦИОННАЯ СИСТЕМА ИСТОРИКА: CASUS INSTRUMENTALIS

Целью исследования явилась реконструкция социальной структуры саами Кольского уезда в XVI – XVIII вв. Методикой исследования было выбрано моделирование, основанное на анализе эмпирического материала. Эмпирическим материалом является информация первичных источников. Для проведения исследования были использованы различные информационные системы.

В ходе работы была создана модель исследуемого сообщества по его состоянию на XVI – XVIII вв. Она позволила решить следующие задачи: концептуализация, построение и описание исследования, организация исследуемого материала, и, наконец, формулирование результатов исследования.

Особенность проведенного исследования состояла в том, что одновременно воссоздавались и изучались как историческая реальность конкретных людей, так и абстрагированная реальность обобщенных построений.

Все исследованные источники были условно разделены по способам использования их для данного исследования. Источник рассматривался как объект потенциальной формализации, т.е. как набор элементов информации, необходимой для исследования. Для оперирования источниками была создана БД "Источники" на основе MS Access.

Особое внимание среди источников было уделено переписям (писцовым и переписным книгам, ревизским сказкам). Они рассматривались как массивы информации. Для хранения и обработки информации переписей была создана БД "Переписи" на основе MS Access.

Ключевым пунктом построения модели является создание связей между составляющими ее массивами информации. Содержание этой процедуры

состоит в объединении информации из различных источников в систему. Это предполагает **идентификацию информации об объектах/феноменах**, содержащейся в различных массивах. Именно эта процедура является технологическим стержнем всего исследования. Объекты были представлены как комплексы элементов информации. Хотелось бы подчеркнуть, что любое ПО способно лишь помочь в идентификации. Вердикт всегда выносит исследователь.

Были выделены такие типы комплексов элементов информации как человек, природный объект, сийтг. В результате работы над идентификацией из **4502** записей о мужчинах сийтгов, зафиксированных и, естественно, дублирующихся, в переписях за период с 1678 г. по 1763 г., были выявлены **2133 мужчины**. Одним из приемов, использовавшихся при идентификации людей, был учет родственных связей. Это осуществлялось с помощью БД "Генеалогия" на основе MS Excel. Для графической презентации генеалогических линий был создан проект "Генеалогия Саами" на основе GenoPro.

В результате проведения процедуры идентификации топонимов были графически представлены **327 объектов** (т.е. более 97% упоминаемых в источниках). Также были идентифицированы фактически **все** упоминаемые в источниках **коллективы саами**.

Для идентификации топонимов была создана БД "Топонимика" на основе MS Excel. Для создания картографических материалов был создан проект "Кольский уезд и окрестности" на основе ArcView. На основе анализа были выявлены следующие институты общества саами исследуемого времени: сийтг, патрилиния, домохозяйство, вежа. Они были рассмотрены с точки зрения их стабильности/динамичности с учетом взаимосвязей между этими институтами. Интерпретации динамики этих институтов дает основу для построения гипотез об экономической специализации разных групп кольских саами в рассматриваемый период.

Применение различных инструментов и методик позволяет осуществить комплексное социально-историческое исследование. Использование информационных систем для хранения, обработки, анализа и презентации информации позволяет объемно представить исторические процессы. Конечно, недостатки ПО приводили к тому, что порой проще было осуществить какие-то операции "вручную", чем развивать приложения. Тем не менее, в технологии исследования не хватает очень существенного элемента – интеграции этих различных информационных систем.

Таким образом, данное исследование приводит к двум внешне противоположным дискуссионным выводам:

- стратегическое развитие ПО исторической информатики должно осуществляться через разработку систем с клиент/сервер архитектурой, способных интегрировать различные приложения;
- компьютер и ПО останутся лишь вспомогательным инструментом в руках исследователя.

Ю.А. Лихтер (Москва)

ВЕЩЕСТВЕННЫЙ ИСТОЧНИК И БАЗА ДАННЫХ – ПРИНЦИПЫ ВЗАИМОДЕЙСТВИЯ

1. При работе по переводу письменных источников в машиночитаемый вид была проанализирована специфика словесных исторических источников. Это привело к созданию особой дисциплины – исторической информатики, которую предлагают понимать как компьютерное источниковедение. Для вещественных источников это не сделано.

2. В вещественном источнике информация содержится в скрытом виде. Чтобы его можно было пользоваться, её необходимо преобразовать. При этом создается новая версия или модель вещественного источника.

3. Описание источника является его словесной моделью. Структура и степень формализованности этой модели зависят от цели описания, от знаний создающего её исследователя, и от познавательной модели, на которую исследователь ориентируется.

4. В музееведении и археологии созданы различные варианты описаний. Описания, предлагаемые музейщиками, ориентированы на неквалифицированного пользователя. Археологические описания ориентированы на исследователя.

5. Однако значение описаний в археологии осознается далеко не всеми исследователями. Господствующий в археологии типологический метод направлен на создание обозначений групп предметов, к которым можно отнести новые, не описывая их детально. Его теоретическую основу можно охарактеризовать как статистическую познавательную модель.

6. С появлением компьютеров усиливается интерес к описанию конкретных вещей и созданию вначале картотек, а затем Баз данных. Разработкой принципов создания картотек занимались исследователи, опиравшиеся на системную познавательную модель.

7. На основе их исследований можно построить словесную модель вещественного источника, которую можно рассматривать как целостную систему. Она складывается из внешних признаков, описывающих контекст, в который включена вещь, и внутренних, описывающих свойства вещи. Система описания внутренних признаков была разработана Ю.А. Щаповой и её учениками, в том числе автором, на семинаре «Морфология древностей». Анализ принципов, положенных в основу этой системы, позволяет утверждать, что она отвечает требованиям, предъявляемым к документальным языкам. По форме организации данных она близка к реляционной модели данных.

8. На основе реляционной модели можно создать логическую модель БД, применимую для ввода и обработки описания любого вещественного источника.

9. Нормированное описание и единая логическая структура позволяют алгоритмизировать процессы описания источника, создания физической структуры БД и ввода данных в базу.

10. Разделение описания на самостоятельные, хотя и взаимосвязанные, подсистемы, позволяет создать также алгоритм исследования информации, введенной в БД.

11. Разработанная и алгоритмизированная система описания служит программой источниковедческого изучения вещественного источника. Совместно с правильно организованной БД она может быть охарактеризована как надежный инструмент изучения древних вещей и связанных с ними разнообразных научных проблем.

В.А. Перевертень (Москва)

ВАРИАНТ ТИПОВОЙ АРХИТЕКТУРЫ ИСТОРИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ

В основе предлагаемого варианта типовой архитектуры исторических информационных систем кроме концепций построения общих для всех современных информационных систем лежит представление о так называемой **КОГНИТИВНОЙ** среде, в которой протекают информационно-познавательные процессы. Основной носитель когнитивной среды – мозг познающего индивида, который может расширять ее за счет внешних средств.

Применяя компьютерную метафору к познавательной деятельности субъекта, информационно-познавательные (когнитивные) процессы можно разделить на три вида: фиксация информации, переработка информации и передача информации. В соответствии с этим части когнитивной среды, где протекают перечисленные процессы, будем называть фиксационной, операционной и коммуникационной субсредами (средами).

До середины прошлого столетия когнитивная среда исследователя могла заметно расширяться лишь в части информационной и коммуникационной субсред. Переработка информации в основном происходила в голове человека. И поскольку основой дополнительной к мозгу человека и внешней по отношению к нему когнитивной среды был бумажный носитель, то ее естественно называть **бумажной**.

Радикальные изменения в плане выноса когнитивных процессов за пределы человека произошли с изобретением компьютера. Возник качественно новый компьютерный носитель расширения человеческой когнитивной среды, который стал позволять не только фиксировать и передавать информацию, а и производить ее переработку. Появилась **компьютерная** когнитивная среда.

Автор считает, что исторические информационные системы, как внешние когнитивные среды, должны строиться так, чтобы усиливать информационные и мыслительные возможности исследователя, быть как можно более естественным продолжением его внутренней когнитивной среды.

Фиксационная среда исторических информационных систем обязана обеспечить когнитивно обоснованные структуры для хранения всей относящейся к историческому исследованию информации.

В зависимости от характера источников и поставленных исследовательских задач, историческая информация может быть в большей или меньшей степени структурирована и представлена в текстовой, графической, аудитивной, визуальной или комплексной форме. Причем сильно структурированная информация может находиться на различных уровнях семантической: от "данных" до "знаний".

В настоящее время для представления и обработки информации в компьютерной среде существует множество информационных технологий. Сильно структурированная информация, как правило, представляется (исходя из уровня требуемой семантической) с помощью либо технологии баз данных, либо технологии знаний, а для лучшей организации слабо структурированной информации рекомендуется использовать гипертекстовые технологии, в том числе (для сложных форм представления информации) технологии гипермедиа. В общем случае очевидна необходимость сочетания возможностей всех упомянутых информационных технологий. При этом в качестве интегрирующей можно использовать технологию гипермедиа.

В **операционной среде** должны быть предусмотрены следующие основные группы программных средств: средства поддержки фиксационной среды; средства, реализующие компьютеризированные методы решения исследовательских задач; средства поддержки коммуникационной среды; управляющие средства, которые обеспечивают координацию функционирования системы; сервисные средства.

Коммуникационная среда должна обеспечить передачу информации внутри системы и комфортное взаимодействие с системой историка. В идеале средства коммуникации для исторических информационных систем следовало бы создавать исходя из модели пользователя-историка.

Среди известных принципов конструирования интерфейса человек-компьютер, основываясь на собственных наблюдениях, для организации взаимодействия историк-компьютер мы выделяем следующие: принцип максимально возможного приближения форм общения с компьютером к профессионально привычным или психологически обоснованным; принцип разумного сочетания традиционных форм представления информации с возможностями компьютера; принцип предпочтения типов и форм диалога, в которых инициатором действий является система; принцип разработки интерфейса в контакте с будущими пользователями.

Предложенная типовая архитектура исторических информационных систем устанавливает как бы родовые черты, которыми должна обладать любая система данного класса.

А.А. Раков (Москва)

**ПРОСОПОГРАФИЧЕСКАЯ БАЗА ДАННЫХ «РКЮУ»:
ТИПОЛОГИЗАЦИЯ СОЦИАЛЬНОГО ПОРТРЕТА
РАСКУЛАЧЕННЫХ КРЕСТЬЯН ЮЖНОГО УРАЛА**

В СУБД Microsoft Access нами была создана электронная просопографическая база данных «Раскулаченные крестьяне Южного Урала (1930–1934 гг.)» (БД «РКЮУ») на 1024 раскулаченные семьи (записи). Наша выборка составляет 11,8% от генеральной совокупности. Эти записи размещены в 34 полях и представляют собой данные по 1 областному и 6 местным архивам Челябинской области. Реляционная БД «РКЮУ» состоит из 8 таблиц (1 основной и 7 дополнительных).

Источниковой основой БД «РКЮУ» стали следующие группы источников:

- списки кулацких семейств, подлежащих выселению;
- списки и характеристики лиц, лишенных избирательных прав;
- анкеты и личные карточки спецпереселенцев.

Реконструировав социальный портрет раскулаченных крестьян Южного Урала по всей выборке, мы посредством запросов выделили для него несколько типологий. Одной из них стала типология по четырем возрастным группам глав раскулаченных семей, а именно: 26–35 лет, 36–45 лет, 46–55 лет, 56–65 лет.

По основным качественным показателям средние показатели всех возрастных групп совпали друг с другом: так, глава раскулаченной семьи оказался трудоспособным русским мужчиной, лишенным избирательных прав, не служившим в армии, раскулаченным по 2-ой категории в 1930 г.

По количественным показателям, которые не совпадали для представителей разных групп, распределение было следующим:

- для возрастной группы глав раскулаченных семей *26–35 лет* средний возраст главы семьи составлял 31,1 года, размер семьи – 4,7 чел., из них трудоспособных – 2,3, а нетрудоспособных – 2,4 чел.; в указанной семье присутствовали представители только одного трудоспособного поколения; имущественные характеристики были следующими: посев – 10,2 дес., количество лошадей – 3,2 гол., а коров – 3,1 гол.;
- для возрастной группы *36–45 лет* средний возраст главы семьи был равен 40,8 года, а размер семьи – 5,6 чел., из них трудоспособных – 2,9, а нетрудоспособных – 2,6 чел.; в такой семье присутствовали представители только одного трудоспособного поколения; имущественные же характеристики были следующими: посев – 10,6 дес., количество лошадей – 3,2 гол., а коров – 3,1 гол.;
- для следующей возрастной группы *46–55 лет* средний возраст главы семьи был 50,7 года при размере семьи – 5,4 чел., из которых трудоспособными были 3,7, а нетрудоспособными – 1,7 чел.; в данной семье присутствовали представители только одного трудоспособного

поколения; а имущественные характеристики были следующими: посев – 10,3 дес., количество лошадей – 3,1 гол., а коров – 3,0 гол.;

- наконец, для возрастной группы глав раскулаченных семей *56–65 лет* средний возраст главы семьи составлял 59,8 года, размер семьи – 4,7 чел., из них трудоспособных – 3,4, а нетрудоспособных – 1,3 чел.; в такой семье присутствовали представители двух трудоспособных поколений; имущественные характеристики были следующими: посев – 9,1 дес., количество лошадей – 3,3 гол., а коров – 2,9 гол.

Основные количественные показатели для разных групп попали в доверительные интервалы, рассчитанные нами для всей выборки.

Следует отметить, что чем старше становится глава семьи, тем большим становилось соотношение между количеством трудоспособных и нетрудоспособных членов в его семье. Одним из наиболее вероятных, с нашей точки зрения, объяснений здесь является то, что достаточно часто вместе с родителями жили и женатые (замужние) дети, что давало «прирост» трудоспособных членов семьи. Одним из косвенных подтверждений этого вывода является и то, что только в семьях, главы которых принадлежат к возрастной группе 56-65 лет, наиболее часто встречается наличие трудоспособного представительства двух поколений (например, трудоспособные отец и сын), а не только одного, как в других возрастных группах. Заметим также, что увеличение соотношения между трудоспособными и нетрудоспособными членами семьи по мере увеличения возраста главы семьи происходит наряду с постепенным уменьшением размера семьи, а также постепенным уменьшением размера посева, начиная с глав раскулаченных семей, принадлежащих к возрастной группе 36–45 лет.

А.А. Rogov, К.А. Rogova, К.Н. Спиридонов, М.Ю. Быстров (Петрозаводск)

ИНФОРМАЦИОННО-ПОИСКОВАЯ СИСТЕМА «ПЕТРОГЛИФЫ КАРЕЛИИ»

Наскальное искусство – это своего рода пиктографическое письмо, образующее крупнейший и наиболее важный исторический архив за последние 40 тысяч лет. В настоящее время выявлено более 35 миллионов наскальных изображений на 140 тысячах скоплениях в 120 странах мира. Петроглифы Карелии, были выбиты в 5-3 веке до н.э. на скалах восточного побережья Онежского озера и берегах Белого моря.

Информационно-поисковая система предназначена для изучения и анализа исторического наследия, обеспечивая наглядность и доступность информации о петроглифах Карелии. Данная научная разработка поддержана грантами РГНФ № 05-01-12118в, № 08-01-1211в (руководитель Н.В. Лобанова).

Информационно-поисковая система состоит из четырех блоков: базы данных петроглифов, модуля подготовки изображений для базы данных, модуля онлайн-доступа и модуля локального доступа к базе данных.

База данных является основой системы. В нее входят графитные копии, фотографии, карты, черно-белые схемы, характеристики и текстовые описания петроглифов. Выделены: группы петроглифов по местонахождению, сюжетные группы и отдельные петроглифы. Описание сюжетных групп содержит: название, кодовый номер, занимаемая площадь, карта местности, текстовая информация, фотографии и графитные копии (не менее трех для каждой). Каждый петроглиф в группе имеет собственное описание, включающее следующую информацию: кодовый номер, название, высота над уровнем моря, глубина вырезки, обрастание мхом, сохранность, угол поворота, фотографии и графитные копии каждого петроглифа (более 10 для каждого).

Модуль онлайн-доступа через Интернет реализован в виде раздела «Каталог» сайта «Петроглифы Карелии». Поиск происходит по признаку «местонахождение петроглифа» путем перемещения с помощью гиперссылок по картам. В настоящий момент на сайте находится более 500 фотографий групп петроглифов с их описаниями. Общее количество представленных петроглифов превышает 2000 фигур. Адрес сайта: <http://smalt.karelia.ru/~petroglyphs>.

Модуль подготовки изображений для базы данных состоит из подпрограмм, реализующих решение следующих задач: выделение отдельных изображений из группы, сегментация изображений, приведение изображений к единому стандарту, нанесение защитных надписей на фотографии.

Сегментация изображений производится с использованием двух методов: порогового и мультифрактальной параметризации структур. В методе мультифрактальной параметризации структур используется спектр фрактальных размерностей Реньи. Следует отметить, что эти методы не являются взаимоисключающими, а выгодно дополняют друг друга.

Модуль локального доступа к базе данных состоит из следующих подпрограмм: поиск по признакам, поиск по изображениям, классификация по признакам, классификация по схожести текстуры изображений.

Поиск по признакам изображений. Были выделены признаки, которые были проанализированы на статистическую зависимость с помощью критерия Пирсона. Пользователю предлагается выбрать значения признаков и точность поиска (количество совпадений признаков).

Поиск по изображениям предназначен для поиска изображений, похожих на данное или на его фрагмент. На вход подается исследуемое изображение, а на выходе должны появиться изображения, наиболее похожие на исходное. Для поиска используются сеть адаптивного резонанса и структурный метод поиска.

Классификация по признакам. Из базы данных пользователю предлагается выбрать петроглифы, которые требуется разбить на определенные группы, а также набор признаков. При помощи иерархического кластерного анализа (метод ближайшего соседа), все выделенные петроглифы разбиваются на группы.

Классификация по схожести текстуры графического изображения основана на гипотезе о том, что разные люди использовали разную технику

выбивания петроглифов, а значит можно провести классификацию петроглифов по степени близости (различно их текстур). В качестве характеристики текстуры поверхности петроглифа был взят *спектр фрактальных размерностей Реньи*.

Заключение. Разработанные информационные методы анализа петроглифов и полученные с их помощью результаты позволяют использовать их при анализе петроглифов других регионов России и во всем мире. Данная система рассчитана в первую очередь на исследователей наскальных изображений, а также будет интересна любому человеку, интересующемуся петроглифами. Число пользователей не ограничено.

Л.А. Серокурова (Москва)

ОПЫТ СИСТЕМАТИЗАЦИИ «ФОНДА ЛИШЕНЦЕВ» СЕВАСТОПОЛЬСКОГО РАЙОНА: ВОПРОСЫ СОЗДАНИЯ БАЗЫ ДАННЫХ

«Фонд лишенцев» Севастопольского района состоит из комплекса дел, представленных персональными делами лишенцев¹, а также делами о создании и функционировании концентрационного лагеря Георгиевский монастырь².

Сохранившиеся документы отражают разнородную информацию о лишенцах, членах их семей, об их имуществе, занятиях прошлых и настоящих и т.д. Два выделенных комплекса документов предполагают необходимость соотнесения персональных дел с данными о содержании в концлагере. Однако первостепенной задачей в систематизации «фонда лишенцев» является анализ личных дел, ввиду того, что они наиболее полно отражают структуру «лишенцев» Севастопольского района, механизм проведения политики лишения избирательных прав в регионе. В сохранившихся делах о лишении прав или восстановления в правах наиболее полно отражен период 1928–1931 гг.

Начало систематизации этого комплекса документов привело к тому, что была создана база данных³, на основе формирования «унифицированной анкеты»⁴. Этот способ для систематизации неструктурированных источников приводит к формализации разнородных данных, для последующей

¹ ГАГС (Государственный архив г. Севастополя). Ф. Р-430. Оп. 5. ДД. 1-642.

² ГАГС. Ф. Р-430. Оп. 5. Д. 643. Лл. 1–9 (об).

³ Для «лишенцев» Крыма подобный опыт анализа не применялся (См.: Неизвестные страницы политических репрессий в Крыму. Права голоса лишены // Серия реабилитированные историей. Симферополь: Таврия, 1998; Омельчук Д.В., Акулов М.Р., Вакатова Л.П., Шевцова Н.Н., Юрченко С.В. Политические репрессии в Крыму (1920–1940 гг.). Симферополь: «Таврида», 2003).

⁴ Авторы, применявшие подобный подход к составлению базы данных, полагают, что путем составления «унифицированной анкеты» можно «извлечь структурную информацию из источников личного происхождения на основе компиляции разнородных документов, относящихся к одной и той же персоналии» (См.: Тихонов В.И., Тяжельникова В.С., Юшин И.Ф. Лишение избирательных прав Москве в 1920–1930-е годы (новые архивные материалы и методы обработки). М.: Мосгорархив, 1998. С. 6).

статистической обработки. Следует отметить, что из персональных дел лишенцев Севастопольского района делать выборку нецелесообразно, ввиду того, что дела «лишенцев» могут содержать различную структурную информацию, чаще всего неполную. Поэтому выборка может привести к потере важнейших данных.

Для вышеуказанного комплекса документов была создана «унифицированная анкета», позволяющая фиксировать следующие данные о лишенцах: фамилия, имя, отчество, пол, возраст, национальность, социальная принадлежность, род занятий, статья (по которой лишили прав), основание для лишения прав, дата лишения прав, сведения о высылке, имуществе, членах семьи, месте проживания, данные о подаче ходатайства о восстановлении в правах, результате его подачи, основания для восстановления прав или отказа, время ведения дела и т.д.

Эти сведения далеко не всегда представлены в каждом деле, например национальность в большинстве случаев не указывалась, у многих нет сведений о возрасте, высылке, имуществе, иждивенцах, оснований для отказа в восстановлении прав, статьи лишения прав¹.

Заполнение базы данных подобным образом позволяет говорить о промежуточных результатах, т.к. отработана только четверть персональных дел «лишенцев» Севастопольского района (173 дела), которые позволили сделать 235 записей о лицах, лишенных избирательных прав. Однако она уже позволяет говорить о динамике реализации политики лишения избирательных прав в указанных хронологических рамках. Это также позволяет выявить спектр занятий и видов экономической активности «лишенцев», состав их семей, возрастные и имущественные категории и т.д.

В ходе проделанной работы можно выделить некоторые промежуточные результаты. Среди упомянутых 235 лиц, лишенных избирательных прав, выявлено 98 глав семей, на иждивении которых находилось около 445 человек различного возраста, на которых отразилось (прямо или косвенно) лишение прав их близких². Из 235 «лишенцев» высылке подлежали 64 человека (из них многие - высылке с семьей за пределы Крыма). 40% из 235 «лишенцев» были восстановлены в правах (в период 1927–1935 гг., т.е. до отмены института лишенчества).

Следует отметить, что пик кампании по лишению избирательных прав пришелся на 1929–1930 гг., что совпадает с периодом кампании по коллективизации – но только 20,4% «лишенцев» Севастопольского (сельского) района были определены как кулаки по объему и методам ведения хозяйства.

Что же касается оснований лишения прав, то здесь можно выделить 3 основные группы. Самая многочисленная – совершеннолетние иждивенцы

¹ В этом отношении комплекс дел «лишенцев» г. Севастополя отличается тем, что в каждом деле «городского лишенца» уже имеется анкета, составленная исполнительными органами советской власти, что существенно упрощает их обработку.

² По ст. 15 «д» совершеннолетние члены семьи лишенца подлежали лишению прав, если находились в материальной зависимости от него (См.: Инструкция ВЦИК о выборах в Советы от 4 ноября 1926 г. – ГАГС. Ф. Р–79. Оп. 1. Д. 207. Л. 12–16об).

лиц, лишенных избирательных прав – 28,8%. Следующей по численности группой «лишенцев» являются торговцы и перекупщики – 28%. Группа лиц, применявших наемный труд сверх установленной нормы, составляет 22,2%. Однако не стоит игнорировать малочисленные группы, например, служителей культа – 9,1%.

Возрастная структура лишенцев позволяет говорить о том, что были лишены прав лица преимущественно среднего (26–45 лет – 40%; 46–55 лет – 12,8%) и преклонного возраста (56–70 лет – 18%). Причем из вышеуказанных возрастных категорий «лишенцев» 65% – мужчины.

Полученные промежуточные результаты позволяют говорить о некоторой специфике Севастопольского региона (по количеству высланных), и структуре «лишенцев» указанного региона.

И.В. Скворцова (Краснодар)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИСТОРИЧЕСКОМ ИССЛЕДОВАНИИ (на примере региональной историографии Великой Отечественной войны)

Одной из задач работы было выяснить место Северного Кавказа в проблематике Великой Отечественной войны в советской (затем российской) историографии. Для этого были отобраны историографические источники (с 1960 по 2000-е гг.) о периоде 1941–1945 гг. Была проведена классификация источников по видам. Формализованный региональный материал представлен в виде проблемно-ориентированных баз данных. Для каждого вида источников была создана своя БД. Используя их, можно проанализировать научную литературу о войне в регионе, выявить её проблематику, проследить динамику, определить перспективу развития.

При анализе историографических источников о войне 1941–1945 гг. были определены устойчивые темы.

В качестве примера рассмотрим базу данных «Диссертации и авторефераты». В каждом десятилетии выявлялись наиболее значимые признаки, прослеживалась их динамика, определялись блоки признаков по частоте встречаемости в каждом десятилетии, эти блоки сравнивались между собой (также по десятилетиям).

За исследуемый период о Великой Отечественной войне было написано 3,5% диссертаций от всех по истории СССР. Северный Кавказ в войне – это 8% диссертаций от общесоюзного уровня работ, посвященных 1941–1945 гг.

Анализируя тематику диссертаций, можно отметить темы, которым отдавалось предпочтение в диссертационных работах. В целом можно выделить 3 блока признаков. Блок I составляют такие признаки, как: деятельность партийных, комсомольских и советских органов; советская экономика в период войны. В блоке II – военные действия, социальная сфера, партизанское движение, СМИ, мобилизация, историография, региональная тематика. Блок III составляют работы, исследующие проблемы оккупации, тыла, эвакуации, а также новые темы 1990-х гг.: депортация,

коллаборационизм, потери СССР, ленд-лиз, советская дипломатия, военнопленные обеих сторон, деятельность РПЦ, историческая психология.

В 1960-е гг. Великой Отечественной войне было посвящено 4,2% всех диссертаций по истории в СССР. Диссертаций в регионе – 6,8% от всех работ по войне. Именно в 1960-е гг. начинает активно разрабатываться военная тематика. Общие для региональных диссертаций черты – узкая источниковая база и научно-популярный характер многих работ. Новизна некоторых диссертационных работ 1960-х гг. – изучение зарубежной историографии, мемуаров бывших противников.

Наибольшее количество диссертаций по войне за весь советский период защищено в 1970-е гг. Почти в 1,5 раз увеличилось количество диссертаций в северокавказском регионе. К приоритетным темам можно отнести те же, что в предшествующее время, и добавить тематику СМИ. Источниковая база мало изменилась. Исследователи 1970-х гг. старались не обращаться к спорным и малоизученным темам. В региональных исследованиях темы экономики и партийного руководства являются основными.

1980-е гг. в количественном отношении похожи на 1960-е гг. В 1980-е гг. возрастает в целом количество работ по истории, но количество диссертаций по войне сокращается. Войне посвящено 3% диссертаций от всех работ по истории СССР. Однако отмечается большее разнообразие тем. Появляется новый признак – историческая психология. В регионе также сокращается число работ о войне – 4 % от всех по войне.

С 1990-х гг. тематика диссертаций по Великой Отечественной войне самая разнообразная. Диссертации по Северному Кавказу составляют 14% от всех работ по войне, они связаны с экономикой и социальной сферой. Впервые за всё рассматриваемое время защищены диссертации по оккупационному режиму в регионе. Новой можно назвать тему депортации некоторых народов Кавказа.

Сравнение советского и российского периодов даёт возможность представить историографию войны, выявить основную проблематику. В настоящее время развернулись дискуссии по многим проблемам отечественной истории, в том числе военной. Результаты этих споров должны найти отражение в диссертационных исследованиях.

А.А. Славко, Т.И. Славко

К ВОПРОСУ О СОЗДАНИИ БАЗЫ ДАННЫХ «Беспризорник 1920–1930-х годов»

Сложнейшая задача борьбы с детской беспризорностью стояла перед руководством страны с первых дней советской власти. Причины массовой беспризорности среди детей 1920–1930-х годов явились следствием политических и социальных катаклизмов, вызвавших огромные людские потери и разрушение семьи. Это империалистическая война, последствия голода и хозяйственной разрухи, «красный» и «белый» террор, лишение

избирательных прав, раскулачивание и кулацкая ссылка, введение 58 статьи Уголовного Кодекса за контрреволюционные преступления и т.д.

Одной из важнейших и малоизученных задач по истории детской беспризорности является исследование социального портрета беспризорника. Традиционные комплексы источников по данной теме отсутствуют. В связи с этим, в течение ряда лет создается база данных «Беспризорник». Ее основой послужили переписи беспризорников 1920-х годов, проведенные в Саратове, на территории Среднего Урала, Тверской губернии и др., регистрационные книги на заключенных, или расширенные списки 1930-х годов, статистические формы учета и отчетности (с середины 1930-х годов) для учреждений закрытого типа, в которые помещались дети репрессированных родителей, и др.

В настоящее время БД «Беспризорники» содержит сведения на более, чем 5 тыс. детей, оставшихся без попечения родителей.

Поля в БД сгруппированы по нескольким смысловым блокам. В первом из них зафиксированы демографические характеристики на ребенка и его родителей. В следующем приводятся сведения об условиях жизни ребенка до момента беспризорности и в период последней. Наименьшую по численности информацию (полей) составляют данные о путях выхода ребенка из состояния беспризорности.

Обобщенный социальный портрет беспризорного ребенка 1920-х годов кратко характеризуется следующим образом: преимущественно – это мальчики 12–14 лет, родители которых рабочие или крестьяне. В раннем детстве дети воспитывались в полных семьях, поэтому, несмотря на ломку психики в тяжелейших условиях беспризорности, многие из них были сориентированы на достойное будущее – получение образования и хорошей, с точки зрения ребенка, профессии, которые бы позволили обеспечить ему нормальную жизнь. Одной из важнейших форм по ликвидации беспризорности этого периода явилась организация детских домов. К концу 1920-х годов они формировались по возрастному показателю: дома дошкольные, школьного возраста и подростков. В 1930-е годы основные характеристики беспризорника меняются.

Основными учреждениями, в которые попадал беспризорный ребенок, оставались детские дома и трудовые дома, находившиеся в чрезвычайно тяжелом положении. В 1935 г., устанавливаются следующие четыре типа детских учреждений: детские дома «нормального типа», находящиеся в ведении наркомпросов союзных республик; специальные детские дома для детей, нуждающихся в длительном лечении, они подчинялись наркомздраву; специальные детские дома для детей-инвалидов, переданные в ведение наркомсобесов союзных республик; изоляторы, трудовые колонии и приемники-распределители, подчиненные Наркомвнуделу СССР. После того, как 5 июля 1937 г. Политбюро ЦК ВКП(б) принимает решение о заключении в лагеря на 5–8 лет «всех жен изменников родины членов правотроцкистской, шпионско-диверсионной организации, их дети до 15-летнего возраста попадают в категорию детей-сирот и отправляются в детские дома и закрытые интернаты наркомпросов».

Создание электронной информационной системы по беспризорникам России рассматриваемого периода и далее, позволит в дальнейшем уточнить имеющиеся сведения, обработать массовые источники, отложившиеся в региональных архивах, по единой методике. Это даст возможность сделать материалы сопоставимыми и, в конечном итоге, смоделировать единую картину состояния беспризорности как по стране в целом, так и по ее отдельным регионам.

Д.В. Стрекалов (Тамбов)

**ЭЛЕКТРОННЫЕ БАЗЫ ДАННЫХ ПО ИЗУЧЕНИЮ СОЦИАЛЬНО-
ДЕМОГРАФИЧЕСКОГО СОСТАВА НАСЕЛЕНИЯ ТАМБОВА
в конце XVIII – первой половине XIX вв.**

Урбанистические исследования являются одним из актуальных направлений современной социальной истории. Для исследования социально-демографического состава населения Тамбова в конце XVIII – первой половине XIX вв. на микроуровне были привлечены персонифицированные первичные документы: ревизии по Тамбову за 1795, 1834, 1857 (58) гг., «Списки городских обывателей губернского города Тамбова за 1831, 1832, 1833 гг.» и «Списки домохозяев Тамбова за 1850 г.», на основе которых были созданы базы данных.

БД «Ревизии г. Тамбова» составлена на основе материалов 5-й (1795 г.), 8-й (1834 г.) и 10-й (1857 г.) ревизии по г. Тамбову. БД имеет сложную многоуровневую структуру. Ее формуляр был неоднократно использован в исследовательских проектах лаборатории социальной истории Академии гуманитарного и социального образования ТГУ им. Г.Р. Державина. В данном исследовании, использована таблица условно названная «Большая семья», которая состоит из следующих полей: номер малой семьи, ФИО главы семьи и членов семьи (отдельные поля для каждого), возраст, пол, отношение к главе семьи, положение в малой семье, семейное положение, занятие (в данном случае информация о сословной принадлежности), факт смерти, грамотность, примечания. Данные позволяют рассмотреть не только статику, но и динамику населения города в конце XVIII – первой половине XIX вв.: изменение состава населения, модификации городской семьи (изменение структуры и состава). Совмещенные за ряд лет персонифицированные данные позволяют проанализировать сословную мобильность.

Однако, в ревизиях не содержатся данные о неподатных сословиях (дворянстве, купечестве 1-й и 2-й гильдий, духовенстве). Решить данную проблему позволяет информация других персонифицированных источников «Списки городских обывателей», которые содержат записи обо всех домовладельцах города и их семьях. На основе этих сведений была создана источник-ориентированная БД «Население Тамбова в 1795–1850-е гг.», которая имеет иерархическую модульную структуру. Она включает в себя 5

модулей: один главный – условное название «хозяин» (информация о главе семьи), и 4 встроенных таблицы «семья», «дом», «служба», «занятие».

Таблица «хозяин» содержит информационные поля: счетчик, ФИО, «пол», «лет» (информация о возрасте), «чин», «сословие», «откуда» (т.е. информация о том, является ли он старожилом, родившимся в городе, или вновь прибывшем с указанием, откуда прибыл обыватель), «семейный статус» (холост, женат и др.), «живет в городе» (т.е. проживает ли домовладелец в Тамбове или за ним лишь числится собственность).

В главный модуль включены поля «Примечание» (вносится дополнительная информация) и «источник» (указываются архивные данные об источнике).

Встроенная таблицы «Семья» состоит из следующих информационных полей: номер малой семьи, ФИО членов семьи, их возраст, пол (для каждого), сословие, их отношение к главе семьи, откуда происходили и «примечание». Информация данных полей позволяет определить состав и структуру городской семьи. Предварительный анализ показал, что преобладающей формой жизни для большинства домовладельцев Тамбова была малая (нуклеарная) семья, в рамках которой проживало 39%.

Таблица «Дом» содержит информационные поля, позволяющие определить местоположение имущества в городе (№ части города, квартала, дома). Данная информация позволяет выяснить и построить социально - пространственное размещение населения города. Соответствующие поля таблицы содержат информацию о предыдущих владельцах и застройщиках: ФИО, сословную принадлежность, степень родства, способ получения данного имущества (наследство, купля-продажа, получено в приданное и т.п.). Интересна информация об источнике имущества: аукционный торг, наследство и др.

Таблицы «Служба» и «Занятие» позволяют проанализировать занятия горожан. Здесь содержится информация о должности и месте службы (не всегда указывается эта информация) и ее продолжительности. В таблице «занятия», главным образом, указывается занятие тем или иным промыслом. Одним из основных видов занятий тамбовских обывателей была торговля (ок. 30 %).

Объединение информации баз данных (общее количество записей составляет более 12 тыс.) и комплексный анализ источников дают возможности для сопоставления и детализации информации о городской семье, сравнительного анализа населения города по иным социальным параметрам, а также для изучения социальной стратификации и мобильности городского населения в конце XVIII – начале XIX вв.

Д.А. Функ, К.Г. Шаховцов, М.Г. Кучинский, Е.А. Пивнева, Л.И. Миссонова (Москва)

**ПРОЕКТ «ЭТНИЧЕСКАЯ ДЕМОГРАФИЯ НАРОДОВ СЕВЕРА И СИБИРИ:
СОЗДАНИЕ КОМПЬЮТЕРНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ
ДЛЯ ФУНДАМЕНТАЛЬНЫХ ЭТНОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ» ***

Представляемый проект направлен на обеспечение решения исследовательских задач в области этнической демографии, этнической истории, социальной организации народов Севера и Сибири.

Фундаментальной задачей проекта является создание общедоступной информационной электронной системы, наполняемой данными разных по структуре источников по генеалогии и этнодемографии народов Севера и Сибири. Таким образом, будет создана возможность для проверки гипотез, возникающих в рамках исследований этнической истории, исторической демографии, социальной и родовой организации, с использованием генеалогических данных по широкому кругу народов Севера и Сибири. В качестве модельных нами выбраны массивы данных о саами, манси, южных селькупках, телеулах, шорцах, кумандинцах, челканцах и уйльта за максимально широкий временной отрезок (XVI – XX вв.).

Последующие добавления новых данных в проектируемую информационную систему позволят проанализировать как синхронную, так и диахронную массовую информацию. Такую систему можно рассматривать как вторичный источник, который, в отличие от первичного, сможет развиваться, не ограничиваясь ни хронологическими, ни территориальными рамками.

Информационная система, разрабатывается как база данных, размещаемая на сервере и снабженная web-интерфейсом, удобным для любого пользователя. Такая система позволит обращаться к специально разработанным программным модулям, осуществляющим формирование запросов к базе данных, получение от СУБД результатов их выполнения и передачу пользователю в приемлемом для него виде.

Первый этап проекта уже фактически подходит к концу и можно описать его результаты.

Участниками проекта собрано несколько коллекций массивов персональных данных по разным регионам, народам и историческим периодам. В ходе работы над проектом были проанализированы и обобщены специфические приемы обработки различных типов источников: писцовых и переписных книг, ревизских сказок, метрических и похозяйственных книг, посемейных списков. На основе этого созданы первичные таблицы базы данных, соответствующие обработанным типам источников. Разработан специфический способ унификации сходной информации, взятой из разных источников, объединения и хранения ее.

Таким образом, имеющиеся материалы касаются этнических групп, которые весьма различались своей историей, экономикой и социальной организацией. Собранные массивы охватывают все типы исторических

* Проект поддержан РФФИ, грант № 07-06-00157а

источников, необходимых для создания планируемого информационного ресурса. Общий объем собранных персональных данных превышает 10000 записей.

В результате этой работы подготовлен предварительный проект БД, включающий основные рабочие таблицы (списки лиц, семейств, домохозяйств и т.д.), вспомогательные таблицы («словари» источников, территориальных единиц, населенных пунктов, этнических групп и т.д.) и систему триггеров и ограничений целостности данных. Подготовлены предварительные описания алгоритмов декомпозиции первичных таблиц.

Начата разработка системы хранимых процедур, описывающих основные родственные отношения: созданы процедуры, осуществляющие выбор всех лиц, состоящих в непосредственном родстве с любым лицом БД (родители, дети, сиблинги) и рабочие прототипы процедур выбора всех прямых потомков данного лица по мужской линии, а также поиска 4 типов кузенных браков.

Основным результатом первого этапа работы будет разработка структуры сводных таблиц, количество которых будет соответствовать числу типов объектов учета (лица, домохозяйства, населенные пункты и т.д.). Заполнение сводных таблиц будет произведено автоматически путем выборки данных из соответствующих полей первичных таблиц.

В завершение первого этапа будет создан пробный web-интерфейс, позволяющий осуществлять запросы к созданной базе данных.

Использование информационных технологий в системе организации самостоятельной работы студентов (на примере кафедры источниковедения исторического факультета)

К.Г. Шаховцов (Москва)

ИСПОЛЬЗОВАНИЕ СЕМАНТИЧЕСКОЙ СЕТИ КАК МОДЕЛИ ПРЕДМЕТНОЙ ОБЛАСТИ ПРИ ПРОЕКТИРОВАНИИ БАЗЫ ДАННЫХ ДЛЯ ЭТНОГРАФИЧЕСКИХ ИССЛЕДОВАНИЙ

Избыток вычислительной мощности и объема дисковой памяти современных ПК позволяет проектировать реляционные базы данных для нужд гуманитарных исследований с минимальным учетом «потребностей машины», т.е. преобразуя исходные данные источников лишь в той мере, в которой это обусловлено спецификой табличного представления. При обработке массовых источников такой подход означает осознанный отказ исследователя от предварительной формализации, унификации и кодирования вводимых данных в пользу идентификации «элементарных» (непосредственных) отношений между описываемыми объектами. В результате создаваемая база данных является более адекватным и гибким представлением информации источника и позволяет строить выборки по критериям, неизвестным во время проектирования, хотя это достигается за счет некоторого усложнения структуры базы данных и потери производительности в случае простых запросов.

Данный подход реализуется в базе данных, создаваемой в рамках проекта «Этническая демография народов Севера и Сибири: создание компьютерной информационной системы для фундаментальных этнологических исследований» (грант РФФИ № 07-06-00157а, руководитель д.и.н. Д.А. Функ).

В первую очередь это относится к служебным таблицам («словарям» административных единиц, населенных пунктов, географических объектов), которые спроектированы как представления древовидных структур (как частного случая семантических сетей) в плоских прямоугольных таблицах, где запись о каждом учтенном объекте содержит указание на запись об объекте, считающемся по отношению к нему «родительским». Так, например, в записи о волости содержится указание на запись об уезде, в свою очередь ссылающуюся на запись о губернии; в записи о районе указана область и т.д. Записи в словаре населенных пунктов содержат их названия по данным каждого источника, в котором они упоминаются, с указанием на запись о современном названии.

Поскольку каждая запись делается в соответствии с данными источника, словарь административных единиц вместе со словарем населенных пунктов позволяет не только отслеживать изменения административного деления, но и формулировать, например, запросы о населении, проживавшем на территории современного района на произвольную дату в прошлом. Дополнительным преимуществом такого подхода является простота исправления ошибок, допущенных при идентификации объектов, т.к. требует изменений только в одной записи.

Принцип связи с «родительским объектом» использован и при проектировании основной таблицы «Лица», предназначенной для анализа генеалогий, брачности и т.п. В данном случае каждое лицо включено в две древовидные структуры родственных связей: по отцовской и материнской линиям. Запись о лице содержит не только указания на записи о его отце и матери, но и список (массив) всех его супругов. Таким образом, фиксация в данной таблице только элементарных и, соответственно, однозначно понимаемых родственных связей позволяет автоматически реконструировать генеалогии по данным массовых источников, например, ревизских сказок или похозяйственных книг. Фигурирующие в источниках иные степени родства («зять», «внук», «племянник» и т.д.) содержатся в записях первичных таблиц, а также в таблице, описывающей группы совместно проживающих лиц.

Для облегчения работы с представленной базой данных в настоящее время создается «конструктор родственных отношений» – набор хранимых процедур, описывающий наиболее часто встречающиеся степени родства и свойства. В частности, в качестве проверки применимости написана процедура поиска кузенных браков четырех типов.