

# ТЕХНОЛОГИИ БАЗ ДАННЫХ И ИНФОРМАЦИОННЫЕ СИСТЕМЫ

---

---

*С.А. Баканов, А.А. Фокин (Челябинск)*

## **ОПЫТ СОЗДАНИЯ И ПРОВЕРКИ РЕПРЕЗЕНТАТИВНОСТИ ЭЛЕКТРОННОЙ БАЗЫ ДАННЫХ ДИССЕРТАЦИЙ "СОВРЕМЕННАЯ УРАЛЬСКАЯ ИСТОРИОГРАФИЯ"**

База данных "Современная уральская историография" имеет своей целью исследование основных тенденций развития современной отечественной историографии, затрагивающей вопросы истории Урала, на основании изучения тематики диссертационных исследований, защищенных по историческим наукам в 1992–2004 гг. Источником информации для наполнения базы данных послужил электронный каталог Российской государственной библиотеки, размещенный на ее официальном сайте [www.rsl.ru](http://www.rsl.ru) и открытый для свободного пользования. Информационный поиск осуществлялся по каталогу диссертаций, к которому было сформулировано несколько запросов, ограниченных сроком защиты диссертации, принадлежностью ее к специальностям исторической науки и географическими названиями, связанными с Уралом.

База данных, созданная с помощью специализированного программного пакета SPSS, содержит 480 записей о кандидатских и докторских диссертациях по истории Урала. По своей структуре она состоит из одной таблицы, наполнение которой велось по 10 атрибутам: Автор, Название, Год защиты, Искомая научная степень, Код специальности, Пол автора, Научное направление, Предмет, Хронологические и Территориальные рамки исследования. Помимо первых шести атрибутов, информация по которым имеется в исходных данных, остальные были сконструированы авторами базы для решения аналитических задач. По каждому из атрибутов эмпирическим путем были выявлены устойчивые группы значений, к которым можно отнести каждую работу. Так, например, для атрибута "Научное направление" мы выявили следующие группы: военная история, история государственных учреждений, историческая биография, история науки, история образования, историография, источниковедение, история культуры, история межнациональных отношений, этнография, политическая, социальная и экономическая история. Вне нашего внимания остались археологические исследования, так как их территориальные рамки часто трудно локализовать уральским регионом. Для атрибута "Хронологические рамки" устойчивыми группами выступили исторические периоды: до 1861 г.; 1861–1917 гг.; 1917–21 гг.; 1920–30-е гг.; 1941–45 гг.; 1945–85 гг.; 1985–90-е гг.; XX век в целом. При заполнении атрибута "Территориальные рамки" мы сознательно отказались от агрегирования данных, чтобы представить разнообразие авторских подходов к понятиям

"Урал", "Приуралье", "Зауралье" и т.п. Наибольшей проблемой при создании настоящей базы данных стало отнесение тематики конкретной диссертации к устойчивым группам каждого атрибута, так как на этом этапе работы возникла вероятность субъективной оценки тематики.

Работа над данным проектом проводилась по аналогии с уже апробированной одним из авторов настоящей статьи базой данных "Постсоветская историография", которая также строилась на основе каталога РГБ, но имела меньший хронологический охват, ограничиваясь только диссертациями по новейшей истории<sup>1</sup>. Таким образом, обе базы можно рассматривать как независимые выборки из генеральной совокупности диссертаций, защищенных в 1990-е годы по историческим наукам. Совпадение по хронологическому периоду и году защиты составило 100%, несмотря на то, что запросы к каталогу РГБ в обоих случаях были разными. В целях проверки репрезентативности каталога РГБ мы воспользовались опубликованными списками работ, защищенных в двух диссертационных советах Урала, действующих с 1990 г. при Институте истории и археологии Уральского отделения РАН и с 1991 г. при Челябинском государственном университете. В каталог РГБ не попало от 10 до 15% работ, защищенных в этих советах. Если экстраполировать такую пропорцию на остальные советы, в которых обычно защищаются работы по уральской тематике, то выборка диссертаций, попавшая в нашу базу данных, включает более 85% генеральной совокупности, а учитывая то, что советы, по которым проводилась проверка репрезентативности, в среднем проводили защиты чаще других уральских, то и более 90%.

*И.М. Баравуля (Минск)*

#### **СОЗДАНИЕ БД "НАСЕЛЕНИЕ БЕЛАРУСИ В 1897 г." (ПО МАТЕРИАЛАМ ПЕРВОЙ ВСЕРОССИЙСКОЙ ПЕРЕПИСИ)**

Материалы переписи населения 1897 г. с момента их опубликования всегда привлекали внимание исследователей, но традиционно использовались лишь некоторые ее таблицы, а также отдельные статистические данные. До настоящего времени нет достаточно полного представления этого источника в электронной форме, что не дает возможности рассмотреть социально-демографические процессы, в частности, в Беларуси, на рубеже XIX – XX вв. Сложность заключается в объеме массива информации, необходимости большой предварительной работы по приведению данных в формат, соответствующий выполнению исследовательской задачи – созданию соответствующей характеру источника базы данных, дающей возможность использования специальной методики обработки данных, например, многомерного статистического анализа.

---

<sup>1</sup> См. подробнее: Баканов С.А. Историография отечественной истории новейшего времени в 1990-е гг.: опыт статистического изучения тематики диссертационных исследований // Круг идей: алгоритмы и технологии исторической информатики. М.: Барнаул, 2005. С. 85 –100.

В работе была поставлена цель перевода интересующих нас материалов Первой всероссийской переписи населения 1897 г. в электронный формат и создания реляционной базы данных<sup>1</sup>. Для этого была выбрана СУБД Microsoft Access. При создании базы данных учитывались особенности исследуемого источника, связанные как с изданием 100-летней давности, так и с возможностями структурирования материала.

В базу данных была полностью перенесена информация по следующим губерниям: Минской, Могилевской, Гродненской, Витебской и Виленской<sup>2</sup>. Информация по каждой из губерний вошла в 25 таблиц БД. Они содержат сведения о населении (наличном и постоянном, коренном и приезде из других стран) в целом и по возрастным группам, географии проживания и семейному положению, по сословиям и месту рождения, по образовательному уровню, вероисповеданию и родному языку, по хозяйственной деятельности и физическим недостаткам. При этом дается различное сочетание этих групп показателей.

Принципы формирования описываемой базы данных предусматривают точное воспроизведение содержания источника, возможность агрегирования данных с определенной корректировкой, их поиск и анализ в зависимости от поставленной задачи изучения переписи как таковой.

Для анализа БД можно использовать такие методы, как факторный анализ, кластерный анализ, многомерная классификация объектов на основе теории нечетких множеств и др. Применение многомерного статистического анализа позволяет обнаружить зависимости между рассматриваемыми показателями, выявить более и менее значимые признаки по степени их влияния на интересующие нас показатели, определить характер и структуру взаимосвязей между компонентами исследуемых многомерных явлений.

*Ю.Я. Вин (Москва)*

### **БАЗА ДАННЫХ "ВИЗАНТИЙСКОЕ ПРАВО": ПРЕЗЕНТАЦИЯ НОВОЙ ВЕРСИИ**

База данных "Византийское право" – научный проект, осуществляемый с целью систематизации изучаемых правовых памятников. Она представляет собой универсальную систему хранения и представления текстов смешанного диакритического греческого, латинского, а после проведения модернизации СУБД – также старославянского алфавитов. Оригинальная СУБД обеспечивает выполнение процедур формирования массива данных, разного вида их поиска, системной поддержки средствами вспомогательного аппарата. Система обладает рядом специальных функций, обуславливающих научно-исследовательский потенциал БД. Введенные в нее тексты организованы в

<sup>1</sup> Население Беларуси в конце XIX века: база данных, составленная по материалам Первой всероссийской переписи населения 1897 г. / Разработчик И.М. Баравуля. Мн.: БГУ, 2006.

<sup>2</sup> Источники: Первая всеобщая перепись населения Российской империи 1897 г. СПб., 1900–1904. Вып. IV. Виленская губерния. Тетради 1–3; СПб., 1899–1903. Вып. V. Витебская губерния. Тетради 1–3; СПб., 1904. Вып. XI. Гродненская губерния; СПб., 1904. Вып. XXII. Минская губерния; СПб., 1903. Вып. XXIII. Могилевская губерния.

виде электронной библиотеки. Главное место занимают Юстинианово законодательство и "Василики", включая схолии к ним. План развития БД предусматривает систематизацию и представление ее материалов в четырех разделах:

- "Гражданское право",
- "Каноническое право",
- "Юриспруденция и юстиция",
- "Славянские рецепции".

Основной функцией БД является предоставление пользователю полных текстов правовых памятников и отображение с помощью так называемых аннотаций их структуры и атрибуции с указанием системных связей отдельных частей и разделов законодательных сборников. Информационно-поисковая система аннотаций позволяет находить искомые разделы, заголовки, связи, авторов и датировки прототипов и прочие дополнительные сведения там, где они указаны в комментариях к текстам.

Для анализа лексики правовых памятников прежде всего могут быть использованы так называемые словники, которые формируются при вводе текстов в БД и содержат сведения о частоте употребления каждой словоформы, а также программное приложение "Сравнение словников". Основным инструментом анализа содержания текстов служит обновленная в ходе модернизации СУБД информационно-поисковая система, предназначенная для поиска лексики. Эта процедура реализуется с использованием шрифтов трех алфавитов, допуская одновременный поиск слов разных естественных языков. Фактически возможен поиск как отдельных слов, так и целостных словосочетаний. Его режимы предусматривают альтернативное выполнение поиска части или целого слова, с учетом регистра и диакритики, а также всего списка ряда слов. В сравнении с прежней, обновленная система информационного поиска заметно увеличила скорость выполнения операций. Будучи построена на табличной форме репрезентации результатов, поисковая система позволяет сохранять результаты предшествующего поиска и при необходимости возвращаться к ним. Тем самым обеспечивается поиск супплетивных форм лексики и проведение простейших концептуальных изысканий.

К особо перспективным аналитическим средствам БД относятся специальные виды информационного поиска. В их числе – проектируемое программное приложение "Имя", которое нацелено на осуществление трехуровневого поиска собственных имен и названий, а также проходящий пробную эксплуатацию Блок понятий и терминов (БПТ). Он репрезентирует иерархическую структуру данных аутентичных, то есть почерпнутых в текстах византийских правовых памятников, понятий и терминов, начиная с так называемых ключевых слов и грамматических производных форм. Они служат посылкой выявления в текстах изучаемых источников лексем, иначе говоря – инвариантов понятийно-терминологического значения, и – через посредство аутентичных словосочетаний – обладающих собственной логикой синтагм. Приложения "Имя" и БПТ являются средством анализа и сопоставления понятийно-терминологического аппарата источников, написанных на названных выше естественных языках. Первая очередь программного приложения

"Экспликация" служит упорядочиванию грамматического строя латинской, греческой и старославянской лексики словарей отдельных памятников византийского права. Все описанные разработки призваны реализовать компьютерный анализ данных.

*Г.П. Волгирева, С.И. Корниенко,  
Д.Г. Решетников, С.В. Пигалева (Пермь)*

**ИСТОЧНИКОВЕДЧЕСКИЙ АНАЛИЗ –  
ОСНОВА СОЗДАНИЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ  
"ГАЗЕТА "ПЕРМСКИЕ ГУБЕРНСКИЕ ВЕДОМОСТИ" (XIX – НАЧАЛО XX вв.)" <sup>1</sup>**

Газета "Пермские губернские ведомости" (1841–1917 гг.) представляет собой источник разнообразных сведений по истории среднего Урала – его экономики, культуры, этнографии и фольклора. Один из наиболее полных комплектов ПГВ в России и наиболее полный на Урале представлен в фондах Пермского областного краеведческого музея (ПОКМ).

С точки зрения сохранности целый ряд годовых комплектов "Пермских губернских ведомостей" находится в катастрофическом состоянии – особенно это относится к подшивкам за 1909–1913 гг. Кардинальное решение проблемы сохранения этого источника лежит в русле создания машиночитаемой версии "Ведомостей" на электронных носителях. Именно это позволит сохранить уникальный исторический памятник для потомков, продлить его жизненный цикл.

Для решения указанных задач в проекте, осуществляемом сотрудниками ПГУ и ПОКМ, предусмотрено создание информационной системы, содержащей электронные версии газеты "Пермские губернские ведомости" за 1909–1913 гг. и реализующей функции электронного путеводителя к данной коллекции. Информационная система предполагает перевод газеты в электронный формат и создание источник-ориентированной полнотекстовой базы данных.

Особенности информационного объекта предъявляют особые требования к концептуальной и логической моделям информационной системы уже на этапе их проектирования. В частности, уже на этапе проектирования необходима первичная источниковедческая, научно-исследовательская работа, связанная с внешней и внутренней критикой источника.

Так, на первом этапе (2005 г.) реализации проекта первоочередное внимание было уделено первичному источниковедческому анализу указанной части номеров газеты и выработке методов нормативной систематизации газетных материалов. Выявлены состояние коллекции, ее полнота и сохранность, формат изданий, природа и история источника, состояние его изученности, осуществлены формально-логический, структурный, выборочный проблемно-тематический анализ номеров и др.

---

<sup>1</sup> Проект реализуется при финансовой поддержке РГНФ, № 05-01-82102а/У.

Периодические издания по своему содержанию отличаются полифонизмом и представляют собой комплексный источник. Поэтому систематизация всего информационного потенциала газеты на первоначальном этапе потребовала разработки и применения метода выборочного поэкземплярного описания каждого отдельного номера "Пермских губернских ведомостей" за 1909–1913 гг. С помощью этого метода определены основные типы объектов, их атрибуты, типы связей между ними. На основе полученных результатов, спроектирована и реализована следующая ER-модель информационной системы.

Сущность PUBLICATION – номер (выпуск) состоит из страниц (сущность PAGE). Статьи (ARTICLE) через сущность пересечения POSITION размещаются на страницах номера. Структура номера реализована рекурсивной сущностью SECTION, что обеспечивает неограниченную иерархическую структуру. Каждая статья относится к определенному разделу структуры и характеризуется типом публикации (сущность TYPE OF ARTICLE). Персоналии (сущность PEOPLE) связываются со статьей через сущность пересечения PERSON REL, при этом тип отношения к статье определяется сущностью TYPE\_OF\_REL. Кроме этого отношение персоналии к номеру в целом определяется сущностью EDITOR.

По представленной ER-модели создана табличная модель, которая реализована в СУБД "Oracle". Приложения для работы с базой данных реализованы в виде пакетов на языке PL/SQL. В качестве интерфейса для работы с системой используется стандартный браузер. Страницы системы формируются динамически в процессе работы процедур пакетов PL/SQL.

Таким образом, созданная на основе результатов первичного источниковедческого анализа база данных, как основа информационной системы, позволяет достаточно эффективно реализовать источник-ориентированный подход, максимально полно передать облик и структуру издания, обеспечить содержательную полноту источника и удобный доступ к информации.

*А.Ю. Володин (Москва)*

### **РОССИЙСКИЕ ФАБРИЧНЫЕ ИНСПЕКТОРЫ: ОПЫТ ПРОСОПОГРАФИЧЕСКОЙ БАЗЫ ДАННЫХ <sup>1</sup>**

Институт фабричной инспекции в России был проводником и блюстителем законности на фабриках и заводах в годы дореволюционной индустриализации. Важную роль в её успехах сыграли сами фабричные инспекторы. Для создания коллективного портрета чинов этого важного государственного учреждения была создана просопографическая база данных.

БД "Российские фабричные инспекторы" является проблемно-ориентированной реляционной базой данных, которая должна включать по возможности максимальное число сведений, почерпнутых из различных источников, о каждом фабричном инспекторе на каждый год службы. Из ис-

---

<sup>1</sup> Работа осуществляется в рамках научно-исследовательского проекта РГНФ 05-01-01388а.

точников следует назвать следующие: Свод отчетов фабричных инспекторов за [1901–1914 гг.] СПб., 1903–1915; Список личного состава министерства финансов. СПб., 1894; Личный состав фабричной инспекции. СПб., 1889, 1897. Для наполнения БД привлекаются также адрес-календари разных лет и отдельные издания по истории инспектората, например, книга А.А. Микулина "Очерк из истории применения закона 3 июня 1886 г. о найме рабочих на фабриках и заводах Владимирской губернии" (Владимир, 1893). Прежде всего, интересны сохранившиеся данные о самом инспекторе (должность, звание, чин, фамилия и инициалы, точная дата вступления в должность) и об участке его ведения (территориальный состав, количество промышленных заведений, число рабочих, количество паровых котлов, место пребывания инспектора). На настоящий момент собраны фактически все минимальные сведения о каждом из числа фабричных инспекторов, отраженных в списках личного состава.

Обратимся для примера к данным за 1905 г., которые можно считать показательными как по полноте, так и по представительности. Количество полей в БД соответствует максимальному количеству данных, приведенных в списке личного состава фабричной инспекции в "Своде отчетов фабричных инспекторов за 1905 год" (СПб., 1908. С. 1–31), а именно: округ, губерния, участок, состав участка, количество промышленных заведений, число рабочих, количество паровых котлов, место пребывания фабричного инспектора, звание, чин, фамилия и инициалы, точная дата вступления в должность или её вакантности.

Надзор фабричной инспекции в Российской Империи в 1905 г. распространялся на 231 участок в 64 губерниях, объединенных в 6 округов. Всего по фабричной инспекции Департамента торговли и мануфактур в 1905 г. служило 226 участковых инспекторов (всего в инспекторате в 1905 г. служило 268 инспекторов, включая старших и окружных, которые не несли обязательств надзора над отдельным участком), из которых 8 находились в течение отчетного года на военной службе. Главными представителями фабричной инспекции во время революционных событий 1905 г. стали, конечно, окружные фабричные инспектора – И.А. Федоров, И.И. Гончаров, Н.А. Тихомиров, А.А. Микулин, Р.Н. Рума, А.Г. Гнедич. В это время служили по фабричной инспекции и такие известные деятели, как А.К. Клепиков, И.М. Козьминых-Ланин, В.П. Литвинов-Фалинский. Деятельное участие в общественной жизни принимали старшие фабричные инспекторы и участковые фабричные инспекторы столичных губерний.

Фабричные инспектора 1905 г. были преимущественно инженерами-техниками (186 человек), инженерами-механиками (32) и техниками (28). Среди фабричных инспекторов были и врачи (5), и горные инженеры (3), и военные инженеры (2). В состав инспектората также входили доктор медицины, кандидат инженерно-технических наук, кандидат физико-математического факультета, кандидат естественных наук, кандидат математических наук, инженер-химик, корабельный инженер, механик-строитель, выпускник Михайловской артиллерийской академии и ученик инженера-техника.

Чины фабричных инспекторов отличались достаточным разнообразием: в инспекторате служило 2 действительных статских советника, 24 статских советника, 27 коллежских советников, 36 надворных советников, 71 коллежский асессор, 66 титулярных советников, 31 коллежский секретарь, 4 губернских секретаря и 1 коллежский регистратор.

Опыт использования БД, построенной по материалам списка личного состава фабричной инспекции "Свода отчетов фабричных инспекторов за 1905 год", позволяет выяснить профессиональный состав инспекторов, и оценить в целом их положение в обществе, а также узнать, в каком регионе работал тот или иной инспектор. Такие сведения, представленные в динамике, представляют не только научный, но и справочный интерес.

Особенно важным для нашего исследования аспектом развития личного состава фабричных инспекторов в динамике являются ротация и увольнения. Ротация кадров в фабричной инспекции – сложная для изучения проблема, которая характеризует, с одной стороны, возможности личного роста в государственном институте, а, с другой стороны, показывает насколько "осёдлыми" были инспектора. Оказалось, что, как правило, рост в чинах был не связан с изменением места (губернии, не участка) надзора. При этом существенные перемещения (например, из Саратова в Ригу), официально обосновывались желанием укрепить фабричную инспекцию в губернии. Однако есть мнение, не подтвержденное документально, что практика переводов была связана ещё и с тем, что некоторые инспекторы попадали под подозрение в корыстных связях с промышленниками, а с 1903 г. такие переводы происходили и в случае размолвок с губернаторами.

По результатам запросов к БД возможно также выяснить, внесли ли существенные коррективы в личный состав инспектората изменения подчинения инспекторов, вводившиеся с 1904 г., и события революции 1905–1907 гг. Оказалось, что и административные изменения, и революционные события незначительно повлияли на личный состав инспекции. Ежегодно из инспектората, в том числе и переводами на другую должность, уходило 10–15 человек. Таким образом, основной состав инспекции оставался постоянным.

*Н.Г. Горбачева (Пермь)*

**БАЗА ДАННЫХ КАК ОСНОВА ИСТОРИОГРАФИЧЕСКИ-ОРИЕНТИРОВАННОГО  
КОМПЛЕКСА ДЛЯ АНАЛИЗА ДЕЯТЕЛЬНОСТИ  
МЕЖДУНАРОДНОЙ АССОЦИАЦИИ HISTORY & COMPUTING**

Международная ассоциация "History & Computing" (АНС), созданная в 1986 году, и ее национальные организации накопили огромный опыт в области информатизации и компьютеризации исторической науки и образования, сыграли большую роль в возникновении, становлении и развитии такой новой и важной области исторического знания, как историческая информатика.

На современном этапе информатизации науки вообще и исторической науки в частности, историческая информатика подошла к очередной "точке



бифуркации", когда со всей остротой встал вопрос о магистральных направлениях и перспективах ее развития<sup>1</sup>. В этих условиях исторический путь, пройденный АНС и ее национальными ветвями, накопленный ими опыт приобретают особое значение, изучение их становится весьма актуальным<sup>2</sup>.

Такая "саморефлексия" предполагает, наряду с традиционными, использование и современных методов организации, обработки и анализа информации, основанных на использовании компьютерных технологий, в том числе методов, основанных на применении баз данных.

Так, для анализа основных направлений деятельности и состояния научных исследований АНС предполагается использовать создаваемую в лаборатории исторической и политической информатики историко-политологического факультета Пермского государственного университета базу данных, основу которой составляет библиография АНС. При ее проектировании и создании учитывается опыт национальных ветвей АНС по созданию и использованию библиографических баз данных, например, библиографической базы данных ассоциации "История и Компьютер"<sup>3</sup>.

Библиографическую основу базы данных составляют описания публикаций журнала АНС "History & Computing", других изданий АНС и об АНС. Для описания публикаций предусмотрены, прежде всего, стандартные библиографические реквизиты и атрибуты: "автор", "заглавие", "подзаголовок", "год издания", "вид издания", "тип издания" и другие.

Для расширения возможностей использования базы в аналитических целях предусмотрены такие специальные атрибуты-рубрики, как "тематическая", "хронологическая", "географическая", "компьютерная программа", "компьютерная технология", "организация" и другие.

База данных является основой информационного *историографически ориентированного комплекса*, который включает тексты публикаций АНС и об АНС в электронном формате. Библиографические записи в базе данных и объекты специализированной электронной библиотеки связаны гиперсвязями.

База данных реализована на основе реляционной модели под управлением СУБД Access. Она имеет две связанные таблицы. Основная таблица состоит из полей, соответствующих указанным атрибутам. Через поле "гиперссылка" реализуются упомянутые гиперсвязи. Вспомогательная таблица содержит переводы библиографических описаний на русский язык.

Запросы к базе данных позволяют получать количественные и качественные параметры, характеризующие состав, основные направления деятельности, состояние научных исследований в АНС. Их анализ дает возможность

---

<sup>1</sup> Бородкин А.И. Историческая информатика в точке бифуркации: движение к Historical Information Science // Крут идей: алгоритмы и технологии исторической информатики. М.; Барнаул: Изд-во Алт. ун-та, 2005. С. 14.

<sup>2</sup> См.: Boonstra O., Breure L. and Doorn P. Past, Present and Future of Historical Information Science. Amsterdam, NIWI-KNAW, 2004. 130 p.

<sup>3</sup> См. Библиография АИК (1992-1999) // Информационный бюллетень Ассоциации "История и компьютер". №25. С. 20-73; Гарскова И.М., Иванов С.А. Библиографический анализ исторической информатики как научного направления // Информационный бюллетень Ассоциации "История и компьютер". №30. С. 254-256.

изучать публикаторскую активность, структуру изданий по типам и видам, языкам, направлениям исследований, сравнивать тематику исследований по странам и регионам, сопоставлять распределение числа публикаций по национальным отделениям АНС, а также получать динамику публикаций по периодам и освещаемым проблемам. Кроме того, созданная база данных позволяет более эффективно решать задачи организации информации для применения других методов компьютеризированного анализа (например, контент-анализа).

**М.А. Дорофеев (Москва)**

**НОВЫЙ ЭЛЕКТРОННЫЙ РЕСУРС ПО РАБОЧЕЙ ИСТОРИИ:  
БАЗА ДАННЫХ "ТРУДОВЫЕ КОНФЛИКТЫ В МЕТАЛЛООБРАБАТЫВАЮЩЕЙ  
ОТРАСЛИ ПРОМЫШЛЕННОСТИ ЦПР в 1895–1901 гг." <sup>1</sup>**

Исследование конфликтов между рабочими и предпринимателями на предприятиях дореволюционной России получило в последние 10–15 лет новую источниковую базу – многотомное издание "Рабочее движение в России. 1895 – февраль 1917 г. Хроника". В 1992–2005 гг. вышли в свет 9 томов, содержащих описания тысяч стачек и других трудовых конфликтов в российской промышленности 1895–1903 гг. В подготовке выпусков "Хроники" были использованы не только архивные фонды, но и материалы прессы. Это нелегальные, а позднее – легальные газеты и журналы конца XIX – начала XX веков, содержащие информацию о рабочем движении. Как отмечают составители "Хроники", "исходной для сбора сведений о большинстве форм борьбы явилась разработанная методика сбора сведений о стачке как одной из основных форм борьбы пролетариата" <sup>2</sup>.

На основе материалов "Хроники" нами создана база данных по трудовым конфликтам на металлообрабатывающих фабриках Центрально-Промышленного района в 1895–1901 гг. Веским основанием для того, чтобы использовать для создания базы данных сведения о стачках рабочих металлообрабатывающей отрасли, был тот факт, что именно эта отрасль была ведущей в ходе дореволюционной индустриализации России, именно в ней возникали наиболее острые трудовые конфликты конца XIX – начала XX вв.

Создание БД осуществлялось нами на основе СУБД MS Access. Все данные о трудовых конфликтах в металлообрабатывающей отрасли промышленности ЦПР с 1895 по 1901 гг., содержащиеся в "Хронике рабочего движения", находятся в одном файле "Hronica.mbd". БД состоит из двух таблиц (отношений): "Главная", созданная для ввода данных из статей "Хроники", и "Реквизиты", предназначенная для размещения ссылок на первичный источник. БД содержит более 20 информационных полей и более 100 записей по числу трудовых конфликтов за рассматриваемый период в указанном регионе. Собранная в БД информация позволяет проводить анализ по самым раз-

<sup>1</sup> Работа ведется в рамках проекта при поддержке гранта РГНФ №06-01-00220а.

<sup>2</sup> Рабочее движение в России. 1895 – февраль 1917 г. Хроника. Вып. I. М., 1992. С. 12.

личным аспектам трудовых конфликтов, выявляя динамику средней численности участников трудовых конфликтов, продолжительность конфликтов, их результаты и т.д.

В докладе проводится сравнение трудовых конфликтов в двух ведущих отраслях дореволюционной российской промышленности – текстильной и металлообрабатывающей (анализ стачечного движения рабочих-текстильщиков ЦПР был проведен на кафедре исторической информатики МГУ ранее). В результате проведенного анализа выявлен ряд отраслевых особенностей в форме проведения и организации забастовок, в характере требований рабочих, реакции властей и администрации предприятий и т.д. Разработанная нами БД "Трудовые конфликты в металлообрабатывающей отрасли промышленности ЦПР в 1895–1901 гг." и результаты анализа содержащейся в ней информации будут представлены на сайте кафедры исторической информатики в середине 2006 г.

*Л.Ю. Жеребцова (Днепропетровск)*

#### **ПОЛНОТЕКСТОВАЯ ИНФОРМАЦИОННО-ПОИСКОВАЯ СИСТЕМА "МЫТО"**

Представленная в данной работе полнотекстовая информационно-поисковая система (ИПС) является усовершенствованным вариантом созданной ранее базы данных "Мыто", представленной на IX конференции АИК. Поэтому автор считает более целесообразным остановиться на историко-ведческих особенностях ИПС "Мыто".

История организации и функционирования таможенной системы на украинских землях Великого княжества Литовского в конце XIV – второй половине XVI в. (до 1569 г.) отражена в комплексах архивных и опубликованных источников. Поскольку информация о видах пошлин, способах их взимания, эволюции в течение изучаемого периода содержится в разных видах документов ("листы", "привилеи", "вольности", "аренды", "уставы", "квитанции" и т.д.), возникает необходимость построения классификации источников по нескольким принципам и критериям. Целью создания полнотекстовой проблемно-ориентированной ИПС "Мыто" было отображение разнообразия действовавших пошлин на основе архивных и опубликованных комплексов источников.

ИПС "Мыто" обладает обновленным интерфейсом, функционирует на двух языках (украинском и русском), имеет расширенные возможности поиска, а также удобные переходы между таблицами ИПС. С помощью блоков ИПС "Мыто" возможно:

- Проследить общее количество документов (архивных и/или опубликованных), виды документов ("листы", "вольности", "аренды" и т.д.), в которых содержится информация о конкретной выбранной пошлине и непосредственно перейти к ним из таблицы "Мыто".
- Получить библиографическое описание использованных архивных (РГАДА. Ф. 386. Литовская Метрика, книги судебно-административных органов Украины в фондах ЦГИАУ, г. Киев и т.д.) и опубликованных ис-

точников (книги Литовской Метрики, Архив Юго-Западной России, публикации персональных коллекций и т.д.), общее количество подчиненных документов, взятых из данного источника, а также виды документов, которые встречаются в данном источнике.

- С помощью блока "Библиография" определить количество и виды пошлин, которые встречаются в конкретном источнике. Это, в свою очередь, позволит выяснить приблизительное время появления разных пошлин.
- Выявить основные атрибуты средневекового документа и провести его дипломатический анализ на основе полного текста документа.
- Осуществить выборку документов из таблицы "Документ" по таким критериям, как документы одного вида; документы, выданные в одном месте (Вильнюс, Берестье, Торунь, Люблин и т.д.); документы, выданные в период правления определенного правителя (Александр, Сигизмунд Старый, Сигизмунд Август и т.д.).
- С помощью вспомогательных таблиц "Меры и веса" и "Монеты" получить определение средневековых метрологических и нумизматических терминов.
- Получить некоторую информацию о торговых путях, которыми должны были следовать купцы, складских городах, арендаторах и откупщиках государственных пошлин, регионах, где находились государственные и частные таможи, лицах, которые обладали правом беспошлинной торговли, а также других аспектах таможенной системы.

Созданная ИПС "Мыто" является дополнительным инструментом исследования такого многогранного явления, как таможенная система, а также вторичным историческим источником. В перспективе планируется дополнение ИПС картографическим материалом для отображения сети государственных и частных таможен и их филиалов.

*О.И. Кирьянова (Пермь)*

### **"СОВЕТСКИЕ ЖЕНЩИНЫ НА СТРАНИЦАХ ЖУРНАЛОВ": ВОЗМОЖНОСТИ И ТЕХНОЛОГИЯ КОМПЬЮТЕРИЗИРОВАННОГО АНАЛИЗА**

В современную эпоху рекламы и политехнологий проблема репрезентации реальности (и современной, и исторической) является как никогда актуальной. В свете этого интересной проблемой представляется репрезентации советской реальности в СМИ. Нами выбран лишь один из аспектов: "Советская женщина: журнальные образы и реальность. Середина 1940-х – середина 1950-х гг.". Цель работы – сопоставление советской "женской" реальности и ее журнальной репрезентации. Задачи: определить социально-экономический статус советских женщин; проанализировать государственную политику в отношении к женщине; выявить наиболее и наименее типичные образы, в которых представала женщина на страницах журналов, а также их эволюцию; проверить, "работает" ли на страницах журналов уже

утвердившаяся в науке применительно к советскому периоду концепция "контракта работающей матери"; выявить механизмы конструирования женских образов и ролей. Для исследования выбраны ориентированные на женщин журналы "Работница" и "Крестьянка", а также "Огонек" за 1944–1956 гг.

В журналах присутствуют и визуальные, и вербальные образы. Но поскольку именно иллюстрации обращают на себя внимание в первую очередь и обладают особым эффектом воздействия, целесообразно сделать акцент именно на них. Учитывая большой объем источников (т.е. журнальных изображений: фото, рисунков, плакатов и т.д.), для получения адекватных исследовательским задачам результатов недостаточно использовать лишь традиционные методы. Выходом представляется создание реляционной базы данных на основе СУБД Access.

Создание БД на основе изобразительных источников и применение этой технологии и основанных на ней методов анализа не получило еще достаточного распространения. Проблемы здесь очевидны: визуальные образы являются собой совершенно особый тип данных и с трудом поддаются классификации. Однако, во-первых, возможности классификации зависят от того, насколько продуманы атрибуты БД. Во-вторых, уже тот факт, что одни изображения поддаются типизации более или менее легко, а другие – с большими трудностями, говорит о многом.

Для описания графических источников в рамках нашего исследования выделено около 30 атрибутов, на основе которых построены реляционные таблицы. Прежде всего, это атрибуты, которые содержат формальные сведения об иллюстрациях (название, год издания и номер журнала, тип иллюстрации и т.д.) и изображениях (цветовая гамма, пол и количество изображенных людей и т.д.).

Женщине не обязательно быть на изображении, чтобы мы получили представление о ее журнальном образе. К тому же, наша цель – не только типизация и описание женских образов, но и выявление сопутствующих им обстоятельств, признаков. Поэтому важнейшими атрибутами являются "объект изображения\_люди" и "объект изображения\_не люди".

Интересные взаимосвязи с основными образами позволяет установить введение поля "символика", где фиксируется наличие на иллюстрации каких-либо советских, государственных символов.

Поскольку изображения оказывают воздействие на "зрителей" через впечатление, которое производят, необходимы соответствующие атрибуты: "яркость изображения" и "эмоциональный посыл". Еще один атрибут, который следует назвать – это "ситуация показа". Отдельный столбец "мужчины" отводится под типы мужчин, которые встречаются на изображениях.

Третий атрибут имеет непосредственное отношение к женским образам. Наиболее важные среди них: "возраст", "типаж", "профессия/сфера деятельности", "женственность" и др. Для каждого признака предлагается фиксированный набор значений, но не исключается возможность добавления новых. Для некоторых признаков весьма показательным является значение "не выражен".

Создание описанной БД позволит применить метод компьютеризированного анализа изображений на основе информационных запросов к ней.

Это даст возможность, во-первых, составить "коллективный портрет" образов советских женщин. Во-вторых, через сопоставление результатов запросов по различным атрибутам выявить более глубокие связи и зависимости, проникнуть в механизмы создания образов.

*С.И. Корниенко, Д.А. Гагарина (Пермь)*

### **ИНФОРМАЦИОННАЯ СИСТЕМА "СЕТЬ НАУЧНЫХ ПОНЯТИЙ": ОПЫТ ПОСТРОЕНИЯ, ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ**

Проблема построения определенных научных понятий и их систематизации является одной из актуальных проблем для науки вообще и гуманитарного научного знания, в частности. Для ее решения на основе современных информационно-коммуникационных технологий разработана и частично реализована информационная система "Сеть научных понятий", ориентированная на применение в научных исследованиях и образовательной деятельности, в том числе и в области исторической науки.

На начальном этапе разработки подобных систем встает вопрос о модели представления понятийных знаний.

В основу модели рассматриваемой системы положена семантическая родовидовая сеть научных понятий, представляющая собой совокупность узлов (т.е. понятий) и отношений между ними. Для каждого узла определяется модель его представления и одно или несколько видов отношений между ними. Для задания отношений между понятиями используются строгие родовидовые определения. Такая модель построения системы позволяет представить научные понятия в виде единой иерархии. Каждое понятие в ней может быть как специфическим понятием некоторой предметной области, так и достаточно общим понятием.

Разработка и реализация конкретной модели, естественно, должны осуществляться в соответствии с областями научных знаний и целями дальнейшего использования системы.

Для реализации модели семантической сети научных понятий на программном и физическом уровне выбрана СУБД Microsoft Access. Каждое понятие – узел сети – задается в базе данных несколькими полями, ключевым является код. Система кодирования задает отношения между понятиями – структуру семантической сети. Для добавления нового понятия в сеть пользователю необходимо выбрать из имеющихся понятий определяющее, при этом кодирование происходит автоматически.

В настоящее время в системе разработаны и функционируют модули, связанные с образовательными целями: освоением знаний и тестовым контролем. В системе имеются разделы для использования в различных предметных областях, в том числе в области истории, политологии, а также на междисциплинарном уровне. Например, имеется блок, содержащий понятия системы государственного управления Российской Федерации. В нем содержатся такие понятия, как Президент, Государственная Дума, Совет Федерации, Кон-

ституционный Суд, Правительство и т.д. Для задания отношений используются определения, взятые из Конституции РФ.

Важнейшей частью разработанной информационной системы является модуль тестирования. Поскольку иерархия понятий представляет собой частично упорядоченное множество с отношением родовидовой связи, специальным образом закодированное, такое частично упорядоченное множество понятий может использоваться для автоматической генерации тестовых заданий типа обобщенного парного сравнения. Причем для большего разнообразия каждое понятие может представляться в виде нескольких моделей (например, различные определения понятия).

Система может использоваться и для решения научно-исследовательских задач: построения сетей понятий на основе системного подхода в различных областях наук для последующего создания, например, поисковых словарей, тезаурусов и др.; построения понятийных моделей различных авторских концепций явлений и процессов с целью их последующего сравнительного анализа. Для реализации этих задач необходимо, наряду с иерархической моделью представления понятийных знаний, использовать и иные, более сложные модели. Например, сетевую, допускающую циклы (правда, в таком случае множество понятий уже не будет частично упорядоченным, и тестирование знаний будет затруднено). При построении таких сетей понятий можно будет использовать не только отношение подчинения, но и другие виды отношений между понятиями, например, равнозначность, перекрещивание и др.

Ряд компонентов информационной системы "Сеть научных понятий" разработан полностью. Они применяются в реальном учебном процессе. Система имеет интуитивно-понятный графический интерфейс, не требующий от пользователя специальных знаний и навыков. Она может использоваться как через Интернет (следовательно, может применяться в дистанционном образовании), так и локально. В то же время продолжается работа по дальнейшей разработке системы и наращиванию возможностей ее использования в научной и образовательной областях.

*С.И. Корниенко, Н.Н. Масленников (Пермь)*

#### **ПРОЕКТ ИНФОРМАЦИОННОЙ СИСТЕМЫ "ЖУРНАЛЫ ЗЕМСКИХ СОБРАНИЙ..." В КОНТЕКСТЕ СОВРЕМЕННЫХ ПРОБЛЕМ ИСТОРИЧЕСКОЙ ИНФОРМАТИКИ <sup>1</sup>**

Успешная реализация возможностей, открывающихся перед исторической наукой в результате развития процесса информатизации, в первую очередь зависит от того, насколько быстро и эффективно решается задача создания машиночитаемых исторических научно-образовательных ресурсов. Теоретические и прикладные аспекты создания машиночитаемых исторических источников рассматриваются в качестве основных задач новой специ-

---

<sup>1</sup> выполняется при поддержке РФФИ. Грант № 04-07-90283

альной исторической дисциплины, исторической информатики, как у нас в стране, так и за рубежом.

Информационная система "Журналы земских собраний как источник изучения истории местного самоуправления в России (II половина XIX – начало XX века)" создается как источник-ориентированный комплекс для всестороннего, последовательного, системного изучения проблем земского самоуправления как исторического феномена российской истории.

При решении задач проекта получили свое дальнейшее развитие концептуальные подходы, программно-алгоритмические решения, связанные с созданием информационных систем на основе исторических источников, представляющих собой квазиструктурированные данные.

Так, база данных "Журналы земских собраний" предназначена для хранения различных земских изданий. При наличии общей формальной структуры, каждое издание имеет свои особенности. Поэтому следует говорить о том, что издания представляют квазиструктурированные данные. В этом случае структура данных известна частично. Известны основные описания базовых сущностей или классов системы. Вместе с тем имеется необходимость хранения и обработки заранее неопределенной дополнительной информации. Указанные обстоятельства требуют нестандартных решений уже при создании базы данных. Это нашло свое выражение на уровне проектирования модели базы данных. Разработанная для базы данных модель предусматривает возможность ввода не только новых данных, как в обычных реляционных базах данных, но и ввод, и редактирование метаданных (схемы БД). Работающий с БД исследователь вводит новую информацию и при необходимости дополняет некоторые сущности новыми реквизитами. Связь между новыми реквизитами и сущностями осуществляется с помощью ассоциативных отношений.

Для нахождения информации в такой не совсем традиционной базе данных была разработана специальная технология поиска, которая обеспечивает поиск не только по стандартным реквизитам, но и по дополнительным атрибутам.

Содержание любого раздела издания можно вывести на экран в отдельном окне в виде PDF- или RTF-документа. В информационной системе "Журналы земских собраний" реализована и возможность доступа к информации, хранящейся в базе данных из локальной или распределенной сети используя веб-интерфейс. В связи с тем, что HTTP – серверы не имеют механизма для обращения к СУБД, была использована платформонезависимая технология CGI (Common Gateway Interface). Эта технология была выбрана в связи с тем, что доступ к данным предполагается осуществлять, используя несколько веб-серверов, работающих под различными операционными системами и на различных платформах.

Для реализации данной технологии было разработано CGI-приложение, которое на основе данных, полученных от СУБД, формирует HTML-страницу, которую возвращает веб-серверу, а тот направляет её клиенту (веб-браузеру). В конечном итоге пользователь имеет возможность получить полный текст любого раздела в виде PDF-документа.



В результате создания информационно-поисковых компонентов системы в значительной мере решены задачи ее реализации как многозадачного комплекса, ориентированного на применение различных компьютерных методов извлечения, обработки и анализа информации источников.

*Т.С. Михеенок (Минск)*

### **ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ БАЗ ДАННЫХ ДЛЯ ОБЕСПЕЧЕНИЯ ДОСТУПА К ДОКУМЕНТАМ НАЦИОНАЛЬНОГО ИСТОРИЧЕСКОГО АРХИВА БЕЛАРУСИ**

Архивы как хранители огромных массивов информации представляют большие возможности для внедрения всевозможных информационных технологий, первыми из которых стали технологии баз данных. С начала 1990-х Национальный исторический архив Беларуси (НИАБ) активно участвует в процессе автоматизации. Одной из первых прикладных программ, внедренных в практику работы НИАБ, была система "Опись типовая". Эта система представляла собой глобальный проект с большим количеством функций. Ее разработка, в связи со сложностью, заняла несколько лет. На стадии постановки задачи наблюдалось отсутствие "общего языка" между разработчиками и будущими пользователями. Внедрение в работу архива новых, ранее не использовавшихся технологий, требовало обучения сотрудников, поэтому ввод информации в базы данных производился очень медленно. Все эти причины приводили к тому, что к моменту начала полноценной эксплуатации такой ИС выявлялось множество недостатков, у пользователей появлялись новые требования, а программное обеспечение уже устаревало.

Другие базы данных архив создавал в стандартных приложениях, таких, как Microsoft Access или Excel. Несмотря на то, что эти форматы легко конвертируются в любые другие стандартные форматы, данные в этих базах совершенно не унифицированы и содержат многочисленные повторения одной и той же информации (прежде всего, географических названий).

Таким образом, имеющиеся в НИАБ базы данных обладают многими недостатками, такими, как избыточность и возможная противоречивость данных; невозможность поддержки целостности данных; невозможность совместного использования данных. Отсутствие стандартов не позволяет интегрировать существующие БД. Очевидно, что без серьезной реорганизации существующих данных их эффективное использование невозможно.

Одновременно отметим, что основными особенностями архивных проектов является долговременность и трудоемкость наполнения БД информацией, а также длительная перспектива использования таких систем. Поэтому автоматизация архивов должна базироваться на самых современных компьютерных технологиях и использовать самые современные инструментальные средства, обеспечивающие создание качественных и надежных программных продуктов в самые короткие сроки.

Применение современных технологий автоматизированного проектирования (CASE-инструментов) позволит избежать непонимания между пользователями и разработчиками, реализовать весь комплекс требований, найти

компромисс между часто противоречивыми пожеланиями будущих пользователей, а также ускорить создание программного кода.

Анализ современных информационных технологий позволяет поставить перед архивом и долгосрочные, перспективные цели – это создание на базе своих документов хранилищ данных и применение средств оперативной аналитической обработки данных; а также интеграция баз данных в среду WWW с использованием стандартизированных архивных описаний, адаптированных к международным стандартам.

В настоящее время весь накопленный НИАБ массив данных используется только для самого примитивного атрибутивного поиска, который не использует всего потенциала имеющихся данных. Сегодняшний традиционный и электронный НСА позволяет получить в результате поиска только перечень документов, в которых встречаются искомые тематика, фамилия или населенный пункт. В хранилище данных можно будет получить ответы на более сложные запросы, а также на запросы произвольной формы. Технология аналитической обработки данных является одним из наилучших способов выявления значимых тенденций и закономерностей в огромном объеме данных. Эта технология позволяет отыскать внутри хранилища такую информацию, которая не может быть обнаружена ни с помощью запросов, ни посредством создания отчетов.

Например, из простейшего учета метрических книг можно извлечь информацию по истории развития конфессий, данные об открытии и закрытии приходов, их переименовании и т.д. Агрегирование данных из нескольких различных источников поможет моделировать исторические процессы, такие, как развитие хозяйств, миграция населения и многие другие.

Очевидно, что получение такого рода информации послужит развитию многих других прикладных исторических дисциплин, таких, как источниковедение, историческая демография, историческая география, генеалогия и др.

*А.И. Назарова (Чебоксары)*

#### **СОЗДАНИЕ БАЗЫ ДАННЫХ "ГОРОДСКИЕ СОВЕТЫ ЧУВАШСКОЙ АССР В 1946–1964 гг."**

Городскими представительными органами власти республиканского подчинения в Чувашской АССР в указанный период являлись советы Алатыря, Канаша и Чебоксар. Всего для анализа состава среднего звена органов местного самоуправления нами были взяты биографии 3236 человек, являвшихся членами указанных органов.

Создавая базу по персональному составу городских советов депутатов трудящихся, необходимо было установить полный список тех, кто являлся депутатом с 1946 по 1964 гг. Базовым источником, в котором содержались основные данные о народных избранниках (ФИО, год рождения, пол, партийная и профессиональная принадлежность), стала местная периодическая печать. Вспомогательными источниками, пополнившими базу данных сведе-

ниями о депутатах, стали агитационные листки и списки уже избранных представителей народа.

На первоначальном этапе работы, которая происходила в архиве, биографические данные переносились на специальные бланки для последующего ввода информации в память компьютера.

Для построения базы данных по составу горсоветов Чувашской АССР была выбрана программа Microsoft Access. База данных содержит более или менее подробные сведения о каждом депутате горсоветов; эти сведения могут быть положены в основу статистических расчетов, отражающих закономерности эволюции этих органов государственной власти на протяжении всего рассматриваемого периода. База данных состоит из двух таблиц. В первой содержатся неизменные данные о депутатах: ФИО, пол, год и место рождения, национальность и т.д.; всего 9 полей. Во второй таблице содержатся те сведения о народных избранниках, которые могли меняться с течением времени: созыв, округ, город, партийность, образование, награды, профессия и т.д.; всего 17 полей. Поле "код депутата" является ключевым.

В результате анализа базы данных можно получить списки депутатов со всеми их характеристиками по любому признаку, например, списки избранных, представляющие различные отрасли народного хозяйства или профессии. Первичные данные могут быть использованы для уточнения фактов биографии людей, широко известных в республике в тот или иной период. Наконец, можно проанализировать количественный состав депутатского корпуса по значениям всех признаков, которые содержатся в анкетах и записаны в память компьютера. Кроме того, база данных позволяет с помощью так называемого панельного анализа (анализа по спискам) сопоставлять изменения во времени, происходившие с группой одних и тех же людей. Так как часть депутатов избиралась несколько раз, то, используя данные об участии депутатов в работе советов разных созывов и в разных городах, можно выявить эту группу депутатов, идентифицируя их по фамилии, инициалам и дате рождения.

*А.А. Раков (Москва)*

#### **ОСОБЕННОСТИ СОЗДАНИЯ БАЗЫ ДАННЫХ ПО МАТЕРИАЛАМ АРХИВНЫХ ИСТОЧНИКОВ С РАЗНЫМИ ФОРМУЛЯРАМИ (НА ПРИМЕРЕ БД ПО РАСКУЛАЧЕННЫМ ЮЖНОГО УРАЛА)**

Задача данного – "пилотного" – этапа исследования, конечной целью которого является реконструкция социального портрета раскулаченных Южного Урала в период 1930–1934 гг., определяется как составление просопографической БД по материалам местных и региональных архивов, а также региональных ИЦ УВД Южного Урала. Для создания БД была выбрана традиционная для подобного типа исследований СУБД Microsoft Access.

Несмотря на признание современной историографией необходимости изучения микроистории и обилие появившихся в последнее время работ по воссозданию коллективного портрета различных категорий репрессирован-

ных Советской властью<sup>1</sup>, реконструкция социального портрета раскулаченных на материалах Южного Урала еще не становилась предметом серьезного исследования.

Основными источниками настоящего исследования являются личные дела раскулаченных, анкеты к выселению или списки кулацких семей, внесенные общегражданским собранием, имеющие хорошо разработанный формуляр, который позволяет перевести их сведения в электронный формат источник-ориентированной базы данных, и списки лишенных избирательных прав, из которых в первую очередь и выбирались "кулаки". При этом следует отметить, что у списков "лишенцев" не только нет разработанного формуляра, но и вообще какой-либо постоянной информационной структуры.

Вследствие того, что работа проводилась с различными источниками, данные о конкретной семье могли быть неполными. Это утверждение иллюстрируется сопоставлением различных структур используемых источников информации. Изначально в базу данных по Троицкому району, позже дополненную данными по другим территориям Южного Урала, были внесены параметры, соответствующие следующей информационной структуре документа, при этом имущественное положение представлено достаточно подробно, а также указана дата и причина лишения избирательных прав<sup>2</sup>:

№№ п/п	
Ф.И.О. главы и членов семьи	
Возраст	
Посев	27/28 с.-х. г. 28/29 с.-х. г.
Сенокос	27/28 с.-х. г. 28/29 с.-х. г.
Имущественное положение	Быков В т.ч. рабочих надв. пост. на сумму
С/х налог	27/28 с.-х. г. 28/29 с.-х. г.
Суд. Когда и за что	
Сл-ба в армии	
Хар-ка (осн. к выс.)	

Для документов на лишенных избирательных прав информативную часть в самом общем случае можно представить в виде формуляра<sup>3</sup>:

<sup>1</sup> Смирнова Т.М. Социальный портрет "бывших" в Советской России 1917–1920 годов. (По материалам регистрации "лиц бывшего буржуазного и чиновного состояния" осенью 1919 г. в Москве и Петрограде) // Социальная история. М.: РОССПЭН, 2000. С. 87-126; Социальный портрет лишенца (на материалах Урала): Сб. документов. Екатеринбург: УрГУ, 1996; Тихонов В.И., Тяжельникова В.С., Юшин И.Ф. Лишение избирательных прав в Москве в 1920–1930 годы. М.: Мосгосархив, 1998.

<sup>2</sup> Архивный отдел Администрации Троицкого района Челябинской области. Ф. 356. Оп. 1. Д. 10. Л. 25.

<sup>3</sup> Архивный отдел Администрации Октябрьского района Челябинской области. Ф. 9. Оп. 1. Д. 7. Л. 7.

Ф.И.О. главы семьи  
Возраст  
Кол-во чл. сем.  
Посев 27/28 с.-х. г.  
28/29 с.-х. г.  
Коров Дойных  
Подрастающих  
Лошадей Взрослых  
Подрастающих  
МРС  
Овец  
Когда и за что лишен избир. прав  
Характеристика (основание к выселению)

При этом количество лошадей и коров часто сдваивалось, а в исследуемых единицах хранения иногда встречаются следующие примечания: "в графе "всего скота", кроме указанного скота, записаны и свиньи", поэтому определить количество свиней, например, представляется возможным лишь в отдельных случаях. Надо сказать, что о недочетах в оформлении документов на раскулаченных и "лишенцев" говорится и в архивных материалах <sup>1</sup>.

По указанным причинам из множества параметров базы данных были выделены для анализа основные, так или иначе присутствующие в большинстве исследованных документов, а именно: пол главы семьи, национальность главы семьи, возраст главы семьи, размер семьи, количество трудоспособных и нетрудоспособных, причина и год лишения избирательных прав, характеристика, служба в армии, имущественное положение, которые и будут составляющими "социального портрета".

Таким образом, основной трудностью при создании БД из структурно различных источников является разработка их формуляров, а также сопоставление последних, их слияние и выделение наиболее общих параметров для анализа и последующего синтеза.

*Ю.А. Святец (Днепропетровск)*

**ИНФОРМАЦИОННАЯ СИСТЕМА "ВЫБОРОЧНЫЕ ПЕРЕПИСИ  
КРЕСТЬЯНСКИХ ХОЗЯЙСТВ УКРАИНЫ. 1922-1926":  
ТЕХНОЛОГИЧЕСКИЙ АСПЕКТ**

Историческое исследование можно определить как итерационный процесс формулирования и проверки гипотез на основе исторических данных. В связи с этим организация данных имеет немаловажное значение для многократного обращения к ним с различными запросами. Введение исторического источника в научный оборот в условиях компьютерных технологий при-

---

<sup>1</sup> Объединенный государственный архив Челябинской области. Ф. 170. Оп. 1. Д. 652. ЛЛ. 68-69.

обретает новое звучание, поскольку сами данные становятся более доступными для обработки. Для хранения и обработки информации о соответствующей предметной области создаются информационные системы. Основой ИС является информационный ресурс – информация, характеризующаяся определенной ценностью и полезностью в производительной и познавательной деятельности человека.

В нашем исследовании создается ИПС ретроспективного поиска.

Технология проектирования ИПС состоит из шести этапов:

1. Исследование и анализ существующих информационных систем с целью выделения необходимого и достаточного объема информации для решения исследовательских задач.
2. Выбор методов удовлетворения информационных запросов пользователей, определение контуров ИС, режимов ее функционирования.
3. Определение состава и структуры базы данных ИС.
4. Определение технической базы функционирования ИС.
5. Проектирование технологических процессов обработки данных, обеспечивающих получение необходимой информации в соответствующих режимах (регламентном, запросном, диалоговом).
6. Решение вопросов организационного обеспечения функционирования ИС.

Компьютерную информационную базу составляют данные опубликованных материалов выборочных переписей крестьянских хозяйств Украины за 1922–1926 гг.

*И.Г. Силина, А.С. Щетинина (Барнаул)*

#### **КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В ИЗУЧЕНИИ ВЫНУЖДЕННЫХ МИГРАЦИЙ НА ЮГЕ ЗАПАДНОЙ СИБИРИ В 1917–1930 гг. <sup>1</sup>**

Существующие компьютерные базы данных (БД) по истории различных категорий вынужденных переселенцев отражают отдельные фрагменты процесса формирования этого вида миграции и его влияние на жизнь общества. Не хватает комплексного взгляда на проблему соотношения таких разновидностей вынужденных миграций, как принудительных и добровольно-вынужденных. В то же время по источникам достаточно сложно разделить категории беженцев и переселенных в принудительном порядке. До начала массовых репрессий вынужденные миграции, на наш взгляд, следует понимать более широко, они затрагивают различные категории в той или иной степени стимулированных на передвижения масс людей.

Среди известных БД, в той или иной мере затрагивающих вопросы вынужденных миграций на юге Западной Сибири, абсолютное большинство составляют базы данных по репрессированным и спецпоселенцам. Нам неизвестны БД, содержащие сведения о беженцах, хотя тема получила определенное освещение в отечественной историографии. Данные о беженцах в

---

<sup>1</sup> Исследование поддержано РГНФ, грант № 05-01-01128а.

Западную Сибирь содержатся в документах разных архивохранилищ и в силу этого обстоятельства используются фрагментарно. Создание базы данных по беженцам первой мировой войны необходимо как для комплексного исследования вынужденных миграций, так и для того, чтобы составить более подробное представление об отдельной категории вынужденных переселенцев.

Комплексное рассмотрение вынужденных миграций заключается в определении таких параметров, как общая численность мигрантов, причины и условия миграции, направление движения по вынужденному переселению, территория расселения и проживания мигрантов. Необходимо также указать сроки и время происходивших вынужденных перемещений населения. При этом необходимы сведения по определенным видам и категориям беженцев.

Для составления базы данных по беженцам первой мировой войны наиболее полные сведения можно получить из анкетных листов всероссийской сельскохозяйственной переписи населения 1917 г., однако она не дает информации по городам. Для того, чтобы проследить динамику в положении беженцев, можно дополнить базу данных информацией из волостных списков за 1920–1921 гг., включающих в себя схожий набор сведений, хотя и менее подробный, чем в переписных листах 1917 г.

Создаваемая база данных даст возможность проанализировать состояние беженцев с разных сторон. Практически во все поволостные списки, а также в анкетные листы переписи включена информация о возрасте, что позволяет составить представление о половозрастном составе беженцев. Сведения о месте проживания в период эвакуации могут быть использованы при определении территорий расселения беженцев в пределах Алтайской губернии, а также для выявления этого расселения. К этому можно добавить информацию о национальности или подданстве беженцев, проследив, таким образом, зависела ли территория размещения прибывающих на Алтай беженцев от их национальной принадлежности.

Информация о членах семьи беженца (общее количество членов семьи при главе, возраст каждого, обозначение родственных связей) позволяет выявить изменения демографической ситуации среди беженцев: уровень рождаемости и смертности, количество заключенных браков. На основании указанного в разных источниках места проживания можно говорить о возможности внутренней миграции беженцев на момент пребывания в эвакуации.

В переписных листах 1917 г. записан статус вынужденных переселенцев – беженец или гражданский пленный. При изучении списков беженцев за 1920–21 гг. нами выявлено несколько специальных списков для бывших гражданских пленных и интернированных. Информация о статусе вынужденных переселенцев также может быть включена в базу данных, и таким образом можно будет узнать, изменялся ли этот статус за период с 1917 г. по 1920–21 гг. Такие сведения позволят также выяснить, были ли существенные различия между прибывавшими беженцами и гражданскими пленными в особенностях их жизни на Алтае.

Таким образом, создание базы данных по беженцам первой мировой войны, проживавшим на Алтае, имеет большое значение для изучения одной из сторон вынужденных миграций на юге Западной Сибири в первой трети XX в. Использование компьютерных технологий позволит создать ком-

плексную источниковую базу по беженцам с возможностями разносторонней обработки материала.

*И.В. Скворцова (Краснодар)*

**СЕВЕРО-КАВКАЗСКИЙ РЕГИОН В БАЗЕ ДАННЫХ  
"ЭТНИЧЕСКИЕ КОНФЛИКТЫ В СССР. 1918–1945 гг."**

Процесс этноидентичности у основной массы народов Северо-Кавказского региона приходится на начало XX в. После Октябрьской революции 1917 г. в регионе резко увеличивается число этноконфликтов. В сети Интернет в рамках проекта "Электронные публикации в сфере культуры и гуманитарных знаний" была размещена текстовая база данных "Этнические конфликты в СССР. 1917–1945 гг." (Сост.: Е. Зубкова, С. Константинов, Д. Люкшин, Т. Филиппова), состоящая из 2500 информационных карточек с развернутыми формализованными данными и архивным описанием (время, место, субъекты, предмет, форма выражения, способы разрешения конфликта, информатор, его адресат, форма представления информации). Основу этой базы данных составляют материалы центральных архивов (ГАРФ, РГАСПИ и др.).

Всю информацию, относящуюся к Северному Кавказу, мы перенесли в базу данных СУБД Access. БД состоит из двух таблиц: "Конфликт" (с указанными признаками, к которым был добавлен признак "место хранения источника") и "Описание конфликта" (выдержки из документов с требованиями конфликтующих сторон, причинами недовольства и т.д.). Кроме этого, к уже существующей информации были добавлены сведения из некоторых фондов (в основном материалы региональных северокавказских архивов), а также опубликованных сборников документов, не вошедших в исходную текстовую базу данных.

Хронологические рамки рассматриваемых в регионе этнических конфликтов: 1920–1941 гг., т.е. от установления Советской власти до окончания эксперимента с административно-территориальными преобразованиями, завершения национально-государственного строительства. В течение указанного промежутка на Северном Кавказе зафиксировано 176 конфликтов, большая часть которых приходится на 20-е гг. Рассмотрим причины этноконфликтов и их основных участников.

Анализ таблицы "Конфликт" позволил сформулировать несколько причин возникновения конфликтов между отдельными народами или этносами и властью. Причиной большей части из них была земля, вернее, ее отсутствие у некоторых народов, стремившихся получить ее за счет соседей, и территориальные разногласия. Национальный вопрос переплетался с земельным, что приводило к территориальным спорам между этносами. Напомним, что систему хозяйства определяли география и природно-климатические условия.

Другая значимая причина этноконфликтов – русификация в национальных районах (немецких, греческих, эстонских, украинских). Еще одной при-



чиной этноконфликтов был кадровый вопрос, особенно при выборах в Советы и др. местные органы власти. Следует отметить, что в горных районах Чечни и Дагестана этнические конфликты перерастали в политические.

Названные причины этноконфликтов тесно связаны с национально-территориальными преобразованиями. Анализ признака "место конфликта" позволяет выделить 3 блока административно-территориальных единиц: Горская АССР с входящими в ее состав округами и Дагестанская АССР; Чечня, Дагестан, Кабардино-Балкария; Кабардино-Балкария, Северная Осетия, Ингушетия.

Анализ признака "субъекты конфликтов" выделяет 38 национальностей, а также казаков, местную и центральную власть. Количественный подсчет выступлений дает следующую картину: 48% – это доля русских, доля местной и центральной власти как субъектов конфликта составляет 20% и 22% соответственно, далее в порядке убывания – чеченцы, карачаевцы, кабардинцы, ингуши, черкесы, балкарцы, доля остальных этносов незначительна.

Таким образом, основными причинами этноконфликтов были земельные претензии между отдельными этносами, кадровая политика, выступления против власти. В центре этноконфликтов оказались русские как этнос и власть, местная и центральная.

*О.В. Скобелкин (Воронеж), Ю.Ю. Юмашева (Москва)*

#### **ПРОБЛЕМЫ СОЗДАНИЯ ПРОСОПОГРАФИЧЕСКОЙ БАЗЫ ДАННЫХ "ИНОЗЕМЦЫ НА РУССКОЙ ВОЕННОЙ СЛУЖБЕ В ПЕРВОЙ ТРЕТИ XVII в."**

К началу XVII в. служилые иноземцы стали неотъемлемой частью вооруженных сил России, а их количество в русском войске исчислялось несколькими тысячами. С учетом этого основную задачу исследования в самом общем виде можно сформулировать как изучение "коллективной биографии" иноземцев на русской службе в первой трети XVII в. Это, в свою очередь, определило исследовательский жанр, известный под обобщающим названием "просопография".

Исследование охватывает период с 1601 по 1629 гг.

Источники, содержащие информацию об иностранцах на русской службе, отложились в двух десятках фондов Российского государственного архива древних актов (РГАДА).

Все документы написаны скорописью XVII в., и в ходе работы необходимым этапом является их перевод в машиночитаемую форму. Отработка методики такого перевода и создание коллекции текстов о служивых иноземцах в формате текстового редактора начались еще во второй половине 90-х гг., и настоящее исследование частично базируется на полученных тогда результатах.

Общее количество выявленных единиц хранения, содержащих интересующую нас информацию, – около 300. Их объем варьируется от одного листа до десятков и даже сотен листов, и все они различны по своей информативности и достоверности.

Среди привлекаемых источников имеется порядка полутора десятков видов документов личного происхождения и приказного делопроизводства. Каждый из видов источников имеет свой, достаточно устойчивый, формуляр.

Избранная методика (создания и анализа просопографической базы данных) исследования должна помочь решить следующие задачи:

1. Сформировать полный список всех учтенных иноземцев.
2. Предоставить возможность создать полный послужной список конкретного лица, в котором в хронологическом порядке была бы расположена информация об основных этапах его жизни и службы в России.
3. Сформировать списки иноземцев по различным характеристикам: время приезда в Россию; категория "иноземчества"; принадлежность к той или иной войсковой группировке (роте); участие в конкретных боевых действиях; характер и размеры жалования; уезд испомещения; размеры поместных и денежных окладов; переход в православие; наличие и состав семьи; разовые поручения, не связанные с военной службой; факты и причины обращения к властям с челобитной.
4. Оценить тесноту взаимосвязи между различными характеристиками. Специфика привлеченных источников потребовала выработки нестандартных подходов и алгоритмов к решению типологических задач подготовки и переноса информации источника в структуру базы данных с учетом того, что в выбранных для исследования документах содержатся все 4 возможных варианта соотношения "источник – изучаемый объект".

Разработанный в ходе проектирования базы данных алгоритм работы позволяет:

- полностью сохранить информацию всех источников в базе данных с целью ввода в постоянный научный оборот выявленного массива документов;
- создать такую структуру базы данных, которая позволила бы зафиксировать и точно локализовать все разночтения, которые имеются в источниках, проводить "прозрачную" верификацию достоверности и гибкий отбор представленной информации;
- применить весь комплекс исследовательских методик для формирования коллективной биографии изучаемых персонажей.

В ходе исследования стало очевидным, что предлагаемый подход требует изменений устоявшихся взглядов на значение вспомогательных таблиц и, как следствие, изменения последовательности работы с базой данных и увеличения времени, которое должно быть затрачено на ее разработку. Однако ввод в постоянный научный оборот источников с возможностью их многократного повторного использования без ущерба для оригинала, равно как и повышение верифицируемости результатов научных изысканий оправдывает эти трудозатраты.

А.А. Славко, Т.И. Славко

## МЕТОДИКА ФОРМИРОВАНИЯ БАЗЫ ДАННЫХ "РОССИЙСКИЕ БЕСПРИЗОРНИКИ 1900–1929 ГОДОВ"

В настоящее время формируются региональные базы данных в рамках Государственного банка данных о детях, оставшихся без попечения родителей. Согласно Федеральному закону от 4 апреля 2001 г., эти информационные ресурсы предназначены для сбора, обработки, накопления, хранения и поиска информации о таких детях, в том числе с целью передачи гражданам, желающим принять их в свои семьи на воспитание. Ценность такого банка данных как исторического источника существенно расширяется при создании соответствующих баз данных за более ранние исторические периоды.

В данной работе излагаются некоторые итоги создания и использования базы данных "Российские беспризорники 1900–1929 годов".

Единого источника, содержащего для рассматриваемого периода однородные сведения о беспризорниках на персональном уровне, нами выявлено не было.

До революции в России отсутствовала централизованная система сбора и обработки информации о беспризорных детях. Только по Москве в рамках Елизаветинского благотворительного общества существовало 227 местных комитетов, которые собирали сведения о детях, нуждающихся в прирешении. Подобная система учета вызывает сложности в выявлении информации о беспризорных детях, сосредоточенной в различных архивных фондах.

После Революции 1917 г. сбор сведений о детях, находящихся без присмотра и обеспечения, был возложен на губернские и уездные исполкомы. Однако последние не справлялись с огромным потоком беспризорных детей. Что касается учета беспризорников, совершивших противоправные действия и содержащихся в исправительно-трудовых учреждениях закрытого типа, то здесь наблюдается ситуация еще более сложная, чем с детдомовцами. Статистические формы учета и отчетности для учреждений закрытого типа вводятся только с середины 1930-х годов. До этого периода отсутствует единая методика предоставления сведений и наблюдаются разночтения. Неточность данных по малолетним правонарушителям, находящимся в специальных учреждениях закрытого типа, была обусловлена в определенной степени совместным содержанием детей и взрослых преступников, постоянным движением контингента правонарушителей. В целом же можно констатировать, что единая система учета беспризорников в России в рассматриваемый период отсутствовала. Это, в свою очередь, подчеркивает актуальность формирования единого информационного массива данных.

В настоящее время удалось выявить следующие источники и сформировать на их основе ряд баз данных. Это материалы переписи беспризорных детей Саратова 1924 г. и Свердловской области 1926 г., автобиографии беспризорных детей – воспитанников одного из Московских детских домов, регистрационные карточки детей, попавших в детские приемники в Москве и Твери, регистрационные книги воспитанников Трудовых домов для несовершеннолетних преступников.

Докомпьютерное структурирование данных связано с разработкой единой системы измерения признаков, использованием методики контент-анализа, привлечением математических методов.

Создание информационной системы по беспризорникам в досоветской и советской России первых десятилетий Советской власти позволит в дальнейшем уточнить имеющиеся сведения, обработать массовые источники, отложившиеся в региональных архивах, по единой методике. Это дает возможность сделать материалы сопоставимыми и, в конечном итоге, смоделировать единую картину состояния беспризорности как по стране в целом, так и по ее отдельным регионам. Пока же можно говорить об отдельных, хотя и важнейших, тенденциях и закономерностях.

*И.В. Титова (Чебоксары)*

**БАЗА ДАННЫХ ПО ПРОБЛЕМЕ "СОЦИАЛЬНЫЙ ПОРТРЕТ ДЕПУТАТА  
ВЕРХОВНОГО СОВЕТА" (НА МАТЕРИАЛАХ ЧУВАШСКОЙ РЕСПУБЛИКИ):  
ПРЕДВАРИТЕЛЬНЫЕ ИТОГИ**

В век компьютеризации обычной практикой стало использование информационных технологий в исторических исследованиях. Новые технические возможности позволяют разрабатывать проблемы, связанные с анализом массовых явлений и процессов. Одной из таких проблем является изучение возрастного, образовательного, национального аспектов социального облика депутатов Советов разных уровней. Нами предпринимается попытка создания базы данных по депутатскому корпусу Верховного Совета Чувашской АССР (1938–1994 г.). В ходе работы данного органа власти было сформировано двенадцать созывов, причем численность народных избранников увеличивалась от созыва к созыву (I созыв – 96 депутатов, X созыв – 200 депутатов) и составила в общей сложности 1738 человек. База данных создается в СУБД MS Access.

Источниковой базой исследования являются личные карточки депутатов Верховного Совета Чувашской АССР, которые не опубликованы, так как составлялись исключительно для служебного пользования. Использовались также агитационные листовки, выпускавшиеся в период предвыборных кампаний, статистические отчеты Мандатной комиссии каждого созыва. Эти материалы находятся на постоянном хранении в Государственном историческом архиве Чувашской Республики.

На данный момент обработаны данные по I–V созывам Верховного Совета Чувашской АССР, то есть по 524 депутатам. В базу данных были включены признаки: пол, возраст, национальность, образование, партийность, социальное положение, место работы, должность, являлся ли ранее депутатом Верховного Совета Чувашской АССР, правительственные награды, избирательный округ.

Предварительный анализ БД показал, что в пяти первых составах Верховного Совета мужчины составляли 61–73% от числа народных избранников, причем их доля достигает максимума в I созыве; самому молодому депу-

тату был 21 год (I созыв), максимальный возраст народного избранника составил 68 лет (II созыв), а средний – 40 лет. Во всех рассматриваемых созывах преобладали представители титульной нации (чуваши) – 52–70% и русские – 26–40%. Наблюдается тенденция постепенного роста доли чувашей и уменьшения доли русских среди депутатов.

Относительно уровня образования депутатов Верховного Совета следует отметить, что он возрастал с каждым новым созывом: если в первом созыве начальное образование имели 44,8%, а высшее – 16,7%, то в V созыве, соответственно, 38,0% и 42%. Большинство депутатов имели полное среднее образование.

Подавляющее большинство народных избранников (54–75%) являлось членами КПСС. Число депутатов, не состоявших в рядах партии, оставалось в пределах 24–31%.

Результаты обработки данных свидетельствуют, что большинство народных представителей, за исключением I созыва, относилось к государственным и партийным служащим (51,4–73%); достигнув максимума во II созыве (73%), их доля постепенно сокращалась. Следующей преобладающей категорией населения в составе высшего представительного органа власти были работники сельского хозяйства – их численность составляла от 16,7% (II созыв) до 42,7% (I созыв). Третье место заняли работники промышленности, они составляли 10 – 30% депутатского корпуса.

Характеризуя состав Верховного Совета Чувашской АССР первых пяти созывов, следует отметить, что к V созыву среднестатистическим депутатом являлся чуваш в возрасте 40 лет, с высшим образованием, член КПСС, работник органов государственной власти.

Таким образом, работа с базой данных значительно повышает эффективность обработки массовых источников, содержащих биографические сведения о депутатах, позволяет выявить закономерности формирования и изменения депутатского корпуса Верховного Совета Чувашской АССР, определить основные черты социального портрета народного представителя.