

ОБУЧЕНИЕ КОМПЬЮТЕРНОМУ АНАЛИЗУ ИСТОРИЧЕСКИХ ИСТОЧНИКОВ (СТАТИСТИЧЕСКИХ, СТРУКТУРИРОВАННЫХ И ТЕКСТОВЫХ)

О. Л. Липницкая

Белорусский научно–исследовательский центр
электронной документации, Минск

С 1992 года на историческом факультете Белгосуниверситета введен предмет "Историческая информатика". Одна из его главных задач – научить студентов работать с историческими источниками, используя широкий круг существующих программных средств и информационных технологий.

184

В лекционном курсе по информатике студентам дается общее представление об историческом источнике, причем подчеркивается, что *«исторический источник, как носитель реальной и потенциальной, субъективной и объективной, выраженной и скрытой информации, обладает информационной неисчерпаемостью, причем это относится ко всем видам и типам исторических источников, но в первую очередь, к массовым»* [1]. Поэтому ни одна из существующих программ не может обеспечить историка всеми необходимыми средствами анализа источника. Каждая из них воспринимает информацию лишь определенного вида и интерпретирует ее своим, строго определенным образом.

В соответствии с классификацией исторических источников по принципу моделирования информации они делятся на статистические, структурированные, документы, включающие данные, текстовые или нарративные, изобразительные и картографические. Каждой из этих групп предусмотрены «свои» программные средства обработки и анализа, что является предметом обучения студентов.

Обработка статистических исторических источников. Первыми источниками, которые стали подвергаться машинной обработке, были статистические. Средствами такой обработки являются электронные таблицы (ЭТ). С их изучения и применения начинается учеба студентов.

На первых занятиях студенты знакомятся с твердой копией уже готовой таблицы, составленной историком. Перед ними ставятся задачи: ввести данные из таблицы в компьютер, используя возможности ЭТ, проанализировать данные таблицы, графически представить как исходные, так и полученные расчетные данные. При этом внимание студентов концентрируется на том или ином конкретном вопросе решаемой посредством ЭТ учебно–исследовательской задачи.

Работая с таблицей «Виды и нормы основных повинностей тяглых крестьян с волоки в имениях духовных и светских феодалов» [3] студенты учатся создавать и редактировать файл электронных таблиц, вводить данные в таблицу, модифицировать ее, форматировать, пользоваться простейшими математическими операциями (выводить суммы повинностей, размер средней повинности по повету и др.) и готовить таблицу к печати. На материалах таблицы «Структура категорий крестьянства Белоруссии XVI в. по характеру личной зависимости» они отрабатывают операции по графическому представлению данных, а по данным «Фрагмента инвентаря имения Везовец Слонимского повета Новгородского воеводства» – составляют таблицу, а затем производят такие сложные вычисления и статистические расчеты,

185

как определение ковариации, и выясняют, была ли зависимость между размером крестьянского надела и количеством «дорослых» и «недорослых» сыновей. Работая с таблицей "Потребление нефти и нефтяных остатков по Восточному району Российской империи" студенты обнаруживают и исправляют ошибки в исторических данных. Используя фрагмент учебной базы данных, созданной на основе «Пописа войска Великого княжества Литовского 1528 г.» [5], работают с файлами, хранящимися в компьютере или на диске и имеющими значительный объем информации – от 700 записей и более. Анализируя данные таблицы «Партизаны Беларуси», или фрагмента научно–исследовательской базы «Тариф подымного налогообложения Оршанского повета XVI века» (Рукопись "Тарифа", транслитерированная со старопольского языка, еще не опубликована. Автор благодарна выпускнице факультета Е. Я. Павловой и научному сотруднику Института истории НАНБ В. Е. Кудряшову за предоставленную возможность использовать их данные рукописи – сноски) учатся подводить итоги и описывать статистику.

Как показывает практика, студенты–историки лучше усваивают те приемы работы с компьютерными программами, которые были основаны на материале исторических источников, и гораздо лучше, когда

перед ними ставятся пусть небольшие, но практические задачи. Примером может служить вопросник для самостоятельной работы, составленный на основе таблиц из книг П. О. Лойко [6] и В. Ф. Голубева [7]:

1. Составить таблицу, используя данный исторический источник, взяв за основу деление Беларуси по регионам.
2. Перегруппировать таблицу в соответствии с выделенными хронологическими периодами.
3. Вычислить процентные доли всех видов земельных наделов.
4. Выбрать и представить данные для построения гистограммы.
5. Выбрать и представить данные для построения круговой диаграммы.
6. Дать интерпретацию полученных графиков.
7. Составить краткий отчет по проведенному анализу исторического источника (в текстовом процессоре WORD).

Эффективность обучения резко возрастает, когда студенты используют компьютерные технологии в своих курсовых или дипломных работах. Помогая им в этом, преподаватель провоцирует ответы на такие вопросы, как:

Каким было общее количество крестьянских хозяйств и наделов в Беларуси по трем хронологическим периодам (от XVI до XIX вв.) в целом, по регионам, по всем видам наделов?

186

Какова была доля великокняжеских владений в Оршанском повете в XVII веке?

Представить на графике структуру земельных наделов крестьян Беларуси в XVI–XVIII вв.

Или:

Есть ли зависимость между ходом военных действий на фронтах Великой Отечественной войны и статистикой вступления в партизаны?

Каким был национальный состав партизанских бригад?

Анализ и обработка структурированных исторических источников. Студенты ставят в известность о том, что многие исторические источники имеют довольно сложную структуру. Это – так называемые **структурированные источники**. Они содержат разнотипную информацию и поэтому отражаются в реляционных базах данных и обрабатываются с помощью **технологий баз данных**.

В лекциях излагаются основные подходы к проектированию реляционных баз данных, даются определения технологическим понятиям (база данных, банк данных, структура файла, система управления базами данных – СУБД), характеризуются стандартные системы управления базами данных (dBASE, FoxPro, Paradox, Access), специальные программные обеспечения для историков (KLEIO, TACT, SOCRATES, CENSSYS, WINCENS) и основные понятия и приемы работы в ACCESS.

Лабораторные занятия начинаются с изучения файла небольшой учебной базы данных, созданной на основе уже знакомого источника «Попис войска ВКЛ 1528 года». Студенты знакомятся с таблицами, их структурой, полями, записями, ключевыми полями.

Следующий этап – работа с ксерокопией исторического источника – «Пописа войска ВКЛ 1567 года» [8]. Студенты создают свою первую базу данных, на которой учатся сортировать, кодировать данные, устанавливать связи между таблицами, создавать простой запрос и запрос из двух таблиц.

Для обучения работе с большими базами данных, на которых лучше видны преимущества и возможности компьютерных технологий, студенты импортируют записи из учебной базы данных «Орша» (фрагмент научно-исследовательской БД «Тарифы подымного налогообложения Оршанского повета»). Приводя в действие функцию группировки данных, они получают результирующую таблицу с наличием максимальных, минимальных и средних значений, а также сумм и, следовательно, ответы на такие вопросы, как: титулы и звания землевладельцев Оршанского повета, количество человек, носивших тот или иной

187

титул, количество имений по сословиям, типичный представитель шляхты Оршанского повета и др.

Последующие занятия проводятся с более объемными файлами, содержащими от одной до нескольких тысяч записей, извлеченных из вышеназванных баз данных. Студенты закрепляют ранее полученные навыки и учатся выполнять новые операции, как-то: создание новой таблицы из запроса; отчетов с помощью автоотчета в столбец и ленточного, а также посредством Мастера отчетов, импорт данных в файл текстового процессора WORD, создание письменного отчета с таблицей, макрокоманд для работы с несколькими таблицами, перекрестных запросов и запросов с помощью построителей выражений, запросов на добавление, обновление и удаление, создание гиперссылок.

Например, работая с источником "Попис 1667", студент должен создать запрос с группировкой данных, чтобы затем проследить взаимосвязь между количеством имений и количеством коней, количеством имений и количеством драбов, соотношением между датой «призыва» и количеством приведенных коней, и ответить на вопрос, сколько коней поступало в войско от каждого повета и др.

Или, работая с базой данных "Тариф Оршанского повета..." студент должен выбрать в базе данных

населенный пункт (имение, фольварк,...), наиболее известный ему, составить на него "Историческую справку" в текстовом редакторе Word, создать поле гиперссылки в базе данных и связать с текстовым файлом "Справки".

Или составить отчет, в который должны войти следующие показатели: фамилия, имя, отчество землевладельца заданного повета, его должность, количество, принадлежавших ему населенных пунктов, коней и драбов. При этом вывести в отчете количество коней и драбов (суммы и доли в процентах), поставившихся маршалками, старостами, воеводами и др.

Или проследить, как землевладельцы, записанные в "Попис 1567 года", собирались в войско, сколько прибыло в каждый из месяцев года. Для этого использовать перекрестный запрос (операцию построения таблицы для вычисления итоговых значений на основе существующей таблицы или запроса). Его результатом должна стать таблица, в столбцах которой указано, сколько человек по-месячно явились на службу.

Компьютерный анализ и обработка текстовых источников. Студенты информируются о том, что часто реляционная модель ставит историков, работающих со сложными **текстовыми** источниками, перед выбором, какими свойствами источника пожертвовать – целостностью, или внутренней структурой, или

первоначальной оригинальностью текста. Дело в том, что значительная часть текстовых исторических источников является незавершенной, их исторические данные "размыты" и настолько сложны, что требуются более сложные и специфические инструментальные средства, чем те, которыми располагают коммерческие фирмы.

"Арсенал" программного обеспечения для анализа этой группы источников пока невелик. В силу известной специфичности источников для их анализа и обработки используются пакеты, разработанные историками –квантификаторами, т. е. историками, хорошо владеющими компьютерными технологиями, или программистами с глубоким историческим познанием. Примером может служить система KLEIO, разработанная в Институте истории Макса Планка (Геттинген, Германия) под руководством М. Таллера. Она соединяет в себе черты системы управления базами данных и специализированной экспертной системы, учитывающей специфику исторических источников.

Студентами-историками Белгосуниверситета изучается одна из систем анализа текстов, применяемы, в историческом исследовании, – ТАСТ, т. е. полнотекстовая база данных, созданная в Торонтском университете (Канада).

Для того, чтобы анализировать текстовый источник с помощью системы ТАСТ, его необходимо перевести в электронный вид. Поскольку операция ввода текста источника с клавиатуры ранее отработана студентами, они знакомятся с простейшими приемами сканирования и распознавания текстового источника. Материалы этой работы затем используются ими на лабораторных занятиях с системой ТАСТ.

В качестве исторических источников автором отобраны следующие: Хроника Литовская и Жмойтская // ПСРЛ, Т. 32. Хроники: Литовская, Жмойтская и Быховца. Летописи: Баркулабовская, Аверки и Панцырного. – М., 1975; Lietuvos Metrika // Knyga Nr. 8 (1499–1514), Uzrallymu knyga 8 / Parengij A. Baliulis, R. Firkovicius, D. Antanavicius. – V., 1995; Статут ВКЛ 1588 года. – Мн., 1989; William Martin Leake. Travels in northern Greece. vol 3. – Amsterdam, 1967. (Последний источник прилагается к программе ТАСТ, любезно предоставленной канадскими авторами).

Переходя к работе с ТАСТом, студенты прежде всего изучают возможности пакета: автоматическое составление словарного списка слов и словоформ, их частотность (встречаемость), методы работы с искомыми фрагментами текста (с экранами ТАСТа, их модификациями), создание категорий и персональных баз данных,

возможности поиска, выбора с помощью символов подстановки, создание запросов.

Затем студенты сами готовят отсканированные фрагменты источников для преобразования текстового файла в текстовую базу данных. В подготовку текста входят структурирование и маркировка, которые выполняются с помощью описательных языков разметки документа (Label Markup, COCOA или BYU).

Вопросы – задания, предлагаемые студентам перед началом занятий, хотя постепенно и усложняются, но в целом довольно простые: в связи с чем в Летописи упоминаются имена князя Радзивилла и новгородского князя; упоминаются ли в Статуте 1588 года сестры и братья, когда речь идет о девке на выданьи, а если «да», то в какой связи; какое наказание ожидало шляхтича, если он нарушал границу соседа (поиск вести, начиная с артикула № 21); какое наказание следовало простому человеку, ранившему или ложно обвинившему шляхтича; выявить все случаи с упоминаниями о шляхтичах, когда действие классифицировалось как убийство или ранение; создать категорию со словом «шляхтич» и «чоловек», затем в 11 разделе Статута найти статьи, в которых "чоловек простого стану" следовало наказать смертной казнью ("горлом"); во фрагменте текста Статута есть слово "мает", которое имеет значения "мает – должен"

и “мает – имеет”): найти тот случай, когда слово “мает“ употреблено в значении “имеет”.

На последних занятиях, когда студенты поняли принципы работы с источником, им предлагается самим подобрать исторический источник и попытаться произвести его анализ.

Таким образом, используя различные компьютерные технологии и применяя их к разным историческим источникам, как с рукописными, так и опубликованными, студенты учатся их обрабатывать и анализировать, начиная со знакомства с ними, работая с уже подготовленными к переводу в машиночитаемый вид и самостоятельно подготавливая к вводу в машину, отвечая на поставленные преподавателем вопросы, ставя свои задачи и находя их решение с помощью компьютера.

1. Ковальченко И. Д. Исторический источник в свете учения об информации (к постановке проблемы) // История СССР, 1982, № 3.

2. Белова Е. Б., Бородкин Л. И. и др. Историческая информатика. – М., 1996. – С. 56–62

3. Спиридонов М. Ф. Закрепощение крестьянства Белоруссии. – Мн., 1993.

4. Белова Е. Б., Бородкин Л. И. и др. Историческая информатика... С. 209.

5. Попис войска Великого княжества Литовского 1528 года. Т. 33 // РИБ. – СПб., 1915. – С. 1–324.

190

6. Лойко П. О. Прыватнаўласніцкія сяляне Беларусі. – Мн., 1991.

7. Голубев В. Ф. Сялянскае землеўладанне і землекарыстанне на Беларусі XVI–XVIII стст. – Мн., 1992.

8. Попис войска Великого княжества Литовского 1567 года. Т. 33 // РИБ. – СПб., 1915. – С. 325–1128.

191