

ИСТОРИОГРАФИЯ ИСТОРИЧЕСКОЙ ИНФОРМАТИКИ

HISTORIOGRAPHY OF HISTORICAL INFORMATION SCIENCE

СТАНОВЛЕНИЕ КОМПЬЮТЕРНОЙ АРХЕОГРАФИИ ИСТОЧНИКОВ ДЕЛОПРОИЗВОДСТВА XVII–XVIII ВВ.

HISTORY OF DIGITAL ARCHEOGRAPHY OF XVII–XVIII CC.
RECORD MANAGEMENT SOURCES

Захаров Андрей Викторович

Кандидат исторических наук, доцент кафедры истории России института гуманитарного образования Челябинского государственного университета.

E-mail: elural@yandex.ru

Andrey V. Zakharov

Анализ сущности и предмета компьютерной археологии позволяет выяснить ее статус как междисциплинарного направления. Основная цель компьютерной археологии состоит в обширной практике публикаций в сети Интернет и создании компьютерных поисковых инструментов и электронного научно-справочного аппарата. Подготовка научной электронной публикации архивных документов должна выдерживать несколько основных стадий, среди которых: обоснованный выбор технологии электронной публикации и публикуемых документов, поиск оптимального варианта транслитерации текста, обращение исследователя к оригиналу исторического документа для корректуры электронного текста и описания документов. Изучены особенности четырех сетевых проектов издания документов делопроизводства XVII–XVIII вв. для выяснения типичных проблем компьютерной археологии в настоящее время.

Ключевые слова: компьютерная археология, историческая информатика, базы данных, исторические источники, критическое издание архивных документов, информационно-поисковые системы, архивные фонды.

The analysis of the essence and the subject of computer archeology allows the author to determine its status as an interdisciplinary knowledge. The main purpose of computer archeology is extensive practice publications in the Internet and creating computer search tools and electronic finding aids. The prepare scientific electronic edition of archival documents should be a several basic phase. These are reasonable choice of the technology for electronic publication and published documents, the best options transliteration of the text, the researcher work with the originals of historical documents for correction of electronic text and a description of documents. The features of the four network projects publications archival documents XVII–XVIII centuries are studied for find out the typical problems of computer archeology now.

Key words: Computer Edition, Archeology, Historical Information Science, database, historical sources, critical edition of archival documents, information retrieval system, archival fonds.

Упоминания о компьютерной археологии на страницах периодических изданий, сборников конференций с каждым годом появляются чаще, что нельзя сказать о выяснении ее предметной области и науковедческого статуса. Эти вопросы пока не получили конвенционального оформления в исторической информатике и российской историографии. Терминология и научная институционализация компьютерной археологии коррелируются с состоянием «классической» археологии — научной дисциплины по подготовке и описанию публикаций исторических документов. Известность получил тезис о широком понимании «триады» археологической деятельности: полевой — по выявлению документальных памятников, камеральной — по описанию, эдиционной — по публикации документов. Узкое понимание археологии относится к публикаторской работе и знаниям об эдичии (издании) исторических источников. Почти вековая российская дискуссия о предмете археологии обусловлена искусственным обособлением археологических задач, научных институтов и терминологии. Некоторые историки считали сам термин «археология» архаичным, от которого еще в XIX в. отказались в западноевропейских странах. В начале 1960-х гг. С. Н. Валк предлагал переименование собственно предметной области по публикации исторических источников и самой науки археологии в «общепринятое точное название» — «издание документов»¹. За три последних десятилетия появлялись новые вариации названий: «эйдотика», «ретроэдиторика», «публикациеведение», ни один из которых не рассматривался научным сообществом как необходимая альтернатива. В пользу сохранения термина «археология» достаточно сказать о его благозвучности и распространении. Устоявшиеся в России и многих восточноевропейских странах археологическая сфера знаний, процедуры и принципы документальных публикаций фиксируют статус археологии как специальной научной дисциплины и публикаторской деятельности.

В специальной литературе также давно обсуждаются полисемичность самого слова «археология» и многих археологических терминов. Предлагается радикальное обновление терминосистемы археологии², но даже рационально объясненные новые термины иногда воспринимаются негативно как чужеродные. Однако общеизвестно, что длительное искусственное сохранение терминосистемы и методологического аппарата науки губельны принижением альтернативных трактовок, способных ответить на новые вызовы времени. Консервация терминологического аппарата в когнитивном понимании выражает период так называемого нормального развития науки до смены научной пара-

дигмы³. С утверждением новой парадигмы логично ассоциировать компьютерную технологическую революцию и методологический переворот в гуманитарных науках, которые не обошли стороной археологию и источниковедение.

В контексте расширения междисциплинарных связей внутри современного Digital Humanities объектно-предметные границы новых научных направлений и субдисциплин по отношению к традиционным («классическим») дисциплинам приобретают собственные очертания. Примером может быть сопоставление предмета изучения общего («классического») источниковедения и компьютерного источниковедения как нового междисциплинарного направления⁴. Общим объектом «классического» источниковедения и нового направления является исторический источник. Междисциплинарность трактуется не суммированием методов и аналитических процедур различных наук, а выходом представителей этих дисциплин на уровень теории познания, выявлением многоаспектных связей общего объекта изучения⁵. Становятся очевидными и закономерности разграничения предметного поля дисциплин. Методологический фундамент и понятийно-терминологический аппарат многих гуманитарных наук оказываются если не общими, то весьма близкими.

Иначе обстоит дело с формулировкой предмета изучения междисциплинарного направления. Установлено, что междисциплинарность не отменяет специальной предметной области новых научных направлений⁶. Предмет последних расщепляется от прежнего понимания предмета изучения «классической» дисциплины. Поскольку новое направление имеет новую постановку предмета и задач, оно аккумулирует новые методы и методики, пришедшие из других наук. Например, лингвистические приемы анализа текста акцептируются компьютерным источниковедением. Генерируется и специфическая терминология, кажущаяся с позиции «классической» дисциплины избыточной или ненаучной. У «классического» источниковеда нет необходимости понимать отдельные термины из лексикона специалиста по компьютерному источниковедению, но обратная ситуация понимания общего источниковедения является аксиомой. Что касается разделения предметной области компьютерного источниковедения как нового направления, то оно практически состоялось.

Насколько возможно провести аналогичную дифференциацию предмета изучения специалистов-археологов, применяющих компьютерные информационные технологии для публикации исторических документов, от предмета «классической» археологии? Этому вопросу и специальному изучению практики компьютерных публикаций

архивных документов делопроизводства просвещена данная статья.

Правомерность проблемы конструирования предметного поля компьютерной археографии подтверждают публикации исторических источников по принципиально новым технологиям и постепенное появление новых значений терминов, заимствованных из исторической информатики, компьютерной лингвистики, текстологии. Проблема соседствует с ощутимым признанием электронных публикаций исторических документов лишь как новой формы публикаций в «классической» научной дисциплине. Развитие методик и публикаций исторических документов вне аналоговых, типографских форм говорит все-таки больше чем о приращении задач и обогащении предмета «классической» дисциплины. Констатировать лишь простое раздвижение «классических» границ археографии невозможно вследствие нового специфического предмета, а не только в связи с новыми компьютерными технологиями. В практике электронных публикаций такая постановка выражается имплицитно-когнитивными установками авторов компьютерных проектов по отношению к типографским публикациям источников. Разработчики непременно отмечают преимущества компьютерных технологий, но предметное поле реализуемых проектов относительно традиционного археографического, как правило, четко не обозначается.

Появление термина «компьютерная археография», одновременное распространение синонима «электронная археография»⁷ подтверждают скорее усилия авторских коллективов в придании особого статуса электронным публикациям исторических документов. Впервые термин «компьютерная археография» появился в российской печати конца 1990-х гг. и за полтора десятилетия использовался в нескольких значениях: от определения практики электронных публикаций до обозначения нового направления в исторической информатике⁸.

Термин «компьютерная археография» семантически точнее названия «электронная археография»⁹. Слово «электронный» в большей степени относится к свойствам электронных вычислительных машин и средств, поэтому не случайно в лексике закрепились словосочетания «электронные документы», «электронные ресурсы». По аналогичному принципу приняты научным сообществом названия компьютерных технологий, методов и научных дисциплин (компьютерной лингвистики, компьютерной лексикографии и т. д.).

О «компьютерной археографии» упоминается в настоящее время преимущественно в контексте институциональных обобщений и дискуссий. Авторы компьютерно-археографических по сути проектов, как правило, не идентифицируют свою

работу с новым направлением, но стараются обозначить междисциплинарный характер своих изысканий. Такая ситуация кажется парадоксальной. Дело в том, что авторы электронных публикаций источников редко ставят специальную задачу определить эпистемологический статус собственной работы, при том, что эдиционная составляющая редко вызывает сомнения. Эта деятельность, как будет показано ниже на нескольких примерах, имеет характер практической работы, которая способствует введению в оборот новых данных и появлению новых знаний. Технологии электронной презентации документов разрабатываются прикладными компьютерными науками, а общие принципы описания, публикации исторического документа и состав основных элементов¹⁰ научно-справочного аппарата разработаны общей археографией.

Учитывая постулаты науковедческих дисциплин, сформулирую следующее определение. *Компьютерная археография* — это междисциплинарная исследовательская область, направленная на создание, изучение научных электронных публикаций исторических документов по принципам археографии, ведущая разработку компьютерных методик или программ для управления данными исторических документов или их описаний. Кажущееся противоречие в трактовке научной институализации компьютерной археографии между теоретико-аналитической и практической публикационной работой должно быть снято. Компьютерная археография может включать несколько видов деятельности: собственно теоретико-аналитическую по критическому изучению компьютерных методик публикаций и описаний документов; компьютерную эдиционную работу по созданию публикаций; компьютерную камеральную — по созданию и внедрению компьютерных методик и программ описания исторических документов. Например, компьютерная эдиционная работа ставит целью электронную публикацию новых исторических источников с применением эффективных поисковых инструментов, что невозможно без исследования структуры, атрибуции, систематизации публикуемых документов, критического применения компьютерных технологий.

Критика практического опыта и имплицитно выраженных установок специалистов-археографов, использующих компьютерные технологии, позволяет установить, что *предмет компьютерной археографии* состоит в подготовке электронных публикаций и описаний исторических документов с обеспечением многовариантного поиска источниковой информации. Методы, программы обработки и анализа текста или визуальной информации разрабатывают информатика или такие междисциплинарные направления, как компью-

терное источниковедение, поэтому они не входят в предметную область компьютерной археографии.

В чем состоят основные условия адекватной электронной публикации исторического документа? Они обеспечиваются несколькими стадиями эдиционной работы. Первая стадия — подготовительная, состоящая из обоснований выбора формы электронной публикации и выбора исторических документов для публикации. Как правило, выбираются архивные документы различных стадий изготовления: беловые экземпляры, копии, отпуски, черновики и другие оригиналы. Фактически археограф ставит задачу адекватно передать образ исторического документа. Воспроизводимый в печатном или электронном виде документ всегда представляет собой образ¹¹ или полного текста, или фрагмента документа, или формализованных (сокращенных, структурированных археографом) текстовых данных — регестов. Презентация текста памятников в виде регестов особенно актуальна для массовых источников делопроизводства. Любая археографическая передача образа документа возможна при условии работы публикатора с оригиналом или фотокопией, что позволяет в том числе воспроизвести точные выходные данные, пагинацию документа. Эти действия должны указывать на возможность отыскания документа и проверки его воспроизведения. Качественному проектированию электронной публикации способствует изучение истории и особенностей предшествующих публикаций (при их наличии), что отражается в исследовательских материалах, обзорных статьях.

Основу второй стадии эдиционной работы составляет транслитерация — передача знаками современного алфавита знаковой системы публикуемого документа, иногда неточно называемая компьютерным набором. Продумываются способы транслитерации: ручной или полуавтоматический (на основе алгоритмов замены повторяющихся слов или символов). Чрезвычайно важен и оптимальный выбор кодировки шрифта — программного набора символов алфавита и знаков письма.

Третья стадия состоит из корректуры — исправления ошибок электронного текста. В качестве альтернативы проводится сверка электронного текста предшествующих публикаций. Во всех случаях исправления должны производиться по оригиналу документа или его качественной копии. Непосредственная работа с историческим документом необходима и потому, что адекватное воспроизведение текста (или графики, визуального источника) невозможно без анализа степени сохранности документа, палеографических характеристик, установления пагинации.

Последующие действия могут быть синхронны обозначенным стадиям, но в целом соотноси-

мы с логикой общих археографических принципов: составления и редактирования вводного комментария, научно-справочного аппарата и его частей — археографической легенды, примечаний и поисковых инструментов. Таким образом, для адекватной электронной публикации исторического документа должны сохраняться основные принципы, выработанные «классической» археографией, а обращение к оригиналу исторического документа справедливо считать незаменимой составляющей компьютерной эдиционной работы.

Предмет компьютерной археографии предполагает более технологичное сопровождение издания, чем при «бумажной» или другой аналоговой¹² публикации исторических документов, что нацеливает исследователя предоставить пользователю разветвленные средства поиска текстовых или визуальных данных. Многовариантный поиск источниковой информации присущ далеко не всем электронным публикациям, а обеспечивается, во-первых, внутренними поисковыми инструментами программ, структурирующих документы, во-вторых, поиском сведений по метаданным источника (выходным данным хранения, заглавию, подвиду или авторству документа).

Задачи компьютерных эдиционных проектов для поиска или описания исторических памятников тождественны стратегии практической археографии по трансформации документа в исторический источник. Однако понимание «превращения» архивного документа в исторический источник, исполняемое якобы только археографом, обоснованно критикуется¹³. Еще А. А. Лаппо-Данилевский, развивая идею Э. Бернгейма, обосновал научно-эмпирическое понимание источника как «всякого реального объекта, который изучается не ради его самого, а для того, чтобы через ближайшее его посредство получить знание о другом объекте»¹⁴. Действительно, археография не имеет монополии «превращения» документа в источник. Обращение историка, например, к архивному документу в контексте конкретно-исторического исследования уже «рождает» из документа исторический источник. Такое понимание следует экстраполировать и на компьютерную археографию, общий вектор которой направлен на один из способов введения исторических документов в исторический источник — виртуальную трансформацию.

Как научная дисциплина и практическая деятельность археография всегда стремилась реализовать надежные способы введения в научный оборот документальных памятников. На современном этапе, что было ясно с середины 1990-х гг.¹⁵, востребован радикальный прорыв именно в сфере эдиционной. Обширная публикаторская практика с созданием компьютерного научно-справочно-

го аппарата в электронных публикациях — очевидное призвание компьютерной археологии.

Из этой первейшей цели вытекают практические задачи технологического свойства: адаптация технологий публикации текстовых и визуальных данных; создание методик адекватной миграции данных с аналоговых, в том числе с бумажных, печатных публикаций на электронно-магнитные носители информации для презентации аутентичной информации исторического документа; создание эффективного многомерного поискового аппарата; обеспечение доступности и сохранности произведенного электронного продукта. Разумеется, каждая из задач сложна и имеет различные варианты решения, нуждается в научной критике.

В российском сегменте сети Интернет и на оптических носителях информации доминируют электронные републикации «бумажных» изданий исторических источников. Они имеют различные предназначение и уровень подготовки, на что в литературе уже обращалось внимание¹⁶. Так, хороший уровень публикаций задан кафедрой исторической информатики МГУ, открывшей в 1999 г. электронную библиотеку источников для учебно-научных целей¹⁷. Другими научно-информационными центрами создано множество чрезвычайно важных открытых для свободного пользования библиотек и отдельных комплексов оцифрованных источников. Все они имеют колоссальное значение для организации сетевого доступа к редким или малодоступным книжным публикациям, ценны при поиске и обработке источниковой информации. Вполне оправдано, что некоторые электронные републикации потенциально могут стать отправной базой для проектов компьютерной археологии. Как отмечалось выше, гранью таких проектов является соответствующая постановка предмета и задач, обращение к оригиналу публикуемого документа, специально созданный поисковый и научно-справочный аппарат публикации.

Обращусь к четырем примерам электронной публикации делопроизводственных источников XVII–XVIII вв., которые не получили должного освещения в российской периодике. Будучи различными по эдиционной методике и по технологии разработки данные ресурсы, несомненно, должны рассматриваться в контексте компьютерной археологии, но их сравнение продуктивно именно с точки зрения методико-технологической, и не вполне корректно по другим параметрам¹⁸. Первые три проекта являются самостоятельными публикациями, т. е. не существовали в аналоговой версии; четвертый имеет републикационную основу. Различными были стратегические задачи, из которых вытекают определения объемов транслитерированного и факсимильного материала, степень

завершения работ. Остановлюсь на особенностях этих публикаций для выявления типичных проблем компьютерной археологии делопроизводственных источников в настоящее время.

Первое издание «Повесточных сказок думных людей XVII–XVIII веков»¹⁹ инициировано в 1999 г. автором данной статьи с целью полнотекстовой публикации и введения в научный оборот исторических источников о созыве думных людей на боярские съезды в Кремль и подмосковное село Преображенское в 1697–1701 гг., более известными как Дума. Повесточными сказками публикатором назван памятник делопроизводства, известный в практике Разрядного приказа с конца XVI в. как совокупность нескольких подвидов документов: повесток и сводных списков «бояр» в черновых и беловых вариантах, росписей подьячих, боярских сказок. Проект был заявлен как археологический для самостоятельной (автономной) публикации всего памятника в сети Интернет и поддержан РГНФ в 2000–2001 гг. как первое в Рунете полнотекстовое издание массового исторического источника по технологии объектно-реляционной базы данных. Данная технология была выбрана совместно с IT-лидером проекта С. Н. Белинским по принципам доступности программного обеспечения (ПО), функциональности поиска, быстрого набора и коррекции текстовых данных. В течение двухлетнего срока планировалось на основе открытого ПО создать программу управления базой данных, разместить ее на веб-сервере Челябинского государственного университета, а далее регулярно пополнять массив текста. Параметры поиска данных были задуманы для исследовательских целей историков и генеалогов по нескольким атрибутам базы данных: выходным данным и подвидам архивных документов, датировке боярских съездов, фамилиям думных людей, чинам, адресам их жительства. В течение месяца были созданы две программы, одна из которых использовалась для набора и корректуры текста в базе данных на локальном компьютере под операционной системой Linux семейства Unix. Другая программа-скрипт, написанная на языке Perl под СУБД PostgreSQL, размещалась на веб-сервере для агрегирования полнотекстовых источниковых данных в режиме «клиент-сервер» при помощи вышеназванных поисковых запросов.

Значительное время было затрачено на копирование публикуемых документов. Сначала фотоспособом дублировались пленочные копии архивных микрофильмов, позже удалось сделать более качественные монохромные скан-изображения. После пятилетнего перерыва проекта база данных пополнялась новыми текстами, отражающими участие 102 думных людей в позднейших боярских собраниях полного состава. Финальный объем

электронного текста (523 л.) был сверен с архивными оригиналами документов. Поисковые запросы и построчные записи были настроены для возможной трансформации электронной публикации в печатную версию. Перенастройка осуществлялась синхронно с изменениями первоначального варианта археографических правил ввода и отображения данных в браузере. Редактирование правил должно было максимально учесть существующие археографические принципы публикации архивных документов и ввести новые археографические инструкции. В чем же состояли основные технологические подходы? Технология базы данных оказывается оптимальной в отношении трудозатрат и приобретаемых функций при наличии в базе многочисленных стабильных фрагментов текста — значений в полях. Например, были фиксированы записи в полях «места жительства», «имена персоналий», «чины», поэтому гибко отражать возможные редкие изменения этих значений (словоформ, ошибок писцов) без дополнительных манипуляций невозможно. Для сохранения контекста записей при агрегации данных по персоналиям было введено поле «примечания», давшее возможность корректировки значений в каждой построчной записи текста документов. Выяснить верный вариант прочтения формы числа слов, обозначающих рубрики и чины, в данной работе возможно только благодаря «Правилам» электронной публикации, что, конечно, является паллиативом. К тому же синхронное изменение правил потребовало коррекции ряда уже внесенных данных, что увеличило трудозатраты, предсказуемые в пилотном проекте. Проблема передачи словоформ в названии рубрик документов могла быть решена с помощью иной взаимосвязи реляционных таблиц. Построчные примечания использовались только для указания вариантов имен персоналий.

Итогом проекта стала информационная система, действующая с 2000 г., печатная публикация просопографических таблиц, html-версия текстов документов на CD-диске, который тиражировался для монографии²⁰. Выявленные документы памятника были опубликованы полностью, а способы решения возникших эдиционных проблем предложены в другой интернет-разработке²¹. Критика решений²² звучала скорее в адрес технологии реляционных баз данных, которая не позволяет оптимально вводить глубокое структурирование данных на уровне лексем, проводить статистические подсчеты. С точки зрения компьютерной археографии слабая структуризация действительно снижает поисковые возможности публикации. Однако задачи лексического разбора и статистики должны решаться при помощи других технологий в контексте компьютерного источниковедения и компьютерной

лингвистики. Промежуточные итоги проекта и тематического исследования опубликованных источников докладывались на Второй международной летней школе по истории России в 2005 г.²³ К этому времени популярность среди средств интернет-публикаций завоевала новая технология гипертекстовой разметки XML, которая для публикаций некоторых источников делопроизводства (служебных списков, кадастров) будет более затратной по времени и ресурсам, но для обработки нарративных источников с синхронным размещением графических изображений удачно используется.

Продуктивным для российской компьютерной археографии будет изучение опыта французских исследователей Национальной школы хартий в проекте «Средневековые акты аббатства Сен-Дени»²⁴. Работа была инициирована в 2006 г. научным руководителем Olivier Guyotjeannin и IT-лидером Florence Clavaud. Интерес привлекает источниковая основа, в качестве которой на первом этапе были выбраны не многочисленные оригиналы и копии актов, а два не изданных ранее каталога кратких перечней актов: средневековый двухтомник XIII в. и многотомный «генеральный» картулярий 1680–1728 гг. В этих документах делопроизводства записи отдельных актов по сути являются регестами, выполненными монахами три века назад. Полнотекстовое издание каталогов понадобилось для последующей реконструкции корпуса 3300 актов аббатства Сен-Дени, расплывшихся в настоящее время по различным архивохранилищам Франции. В начале проекта были выработаны общие принципы электронного издания, которые отличались от принятых принципов дипломатических изданий. Учитывалось постепенное присоединение оригинальных актов, их списков. Вполне закономерным было намерение авторов создать основу реконструкции корпуса актов за несколько лет (фактически понадобилось около пяти лет), а далее постепенно исправлять ошибки, добавлять акты, что осознавалось как многолетняя работа. Два рукописных каталога-картулярия были оцифрованы, а полноцветные графические изображения синхронно связаны с электронными записями актов в рукописях.

Процедуры подготовки электронной публикации включали оцифровку рукописей, транслитерацию, редактирование в MS Word, далее ручную проводилась разметка текста (языки XML, EAD 2002 DTD). Особенностью кодирования стало описание не единичных символов и знаков, а больших сегментов текста и слов, которые могли быть полезны при поисковых запросах. Отлаженная навигация по документам дает возможность исследователям быстро и точно получить информацию по любому из «регестов» актов по следующим за

просам: краткому перечню актов, сопоставлению электронных текстов и изображений двух рукописей, указателю имен персоналий, номеру записи, датировке акта, типу акта и документа, словам. Для привязки текстов к изображениям через гиперссылки создавались файлы METS. Выбор форматов оцифровки и языков кодирования как ответственный выбор, определяющий всю перспективу проекта, был задан возможностью интегрирования с системами французских национальных архивов и использования в сети Интернет.

Французские исследователи особо подчеркнули безальтернативность обращения к оригиналам рукописей или их цифровым копиям при транслитерации и кодировании текста. Кроме плановых процедур, приходилось готовить файлы данных преобразовать из стандарта TEI P4 в P5. Необходимость преобразований текста при потенциальной смене стандартов явно не дает преимущества этой технологии. Интерес представляет итог расчета трудозатрат, что недостает российским разработкам. Разметка текста заняла 2,5 года, или в пересчете на полный рабочий день занятий 16 человеко-месяцев, причем наем студентов для транслитерации текстов сэкономил время. Труд финансировался грантом французского министерства культуры при условии сотрудничества с частной компанией. Организационными трудностями были оперативный поиск специалистов, «имеющих все научно-технические навыки», невозможность проведения работ непрерывно. Промежуточные итоги подводились через 6 лет²⁵, когда завершилось электронное критическое издание трех томов каталога начала XVIII в. с инструментами навигации и поиска сведений о древних актах.

Третий изучаемый проект «Олонецкая воеводская изба» задействовал силы двух научных центров: исследовательской лаборатории локальной и микроистории Карелии, кафедр Петрозаводского госуниверситета и Научного архива Санкт-Петербургского института истории (СПбИИ) РАН. Научный руководитель коллектива²⁶ И. А. Чернякова выдвинула план создания «исчерпывающе полного, исторически сложившегося документального комплекса»²⁷. В разработке 2005–2007 гг., поддержанной грантом РГНФ, пока удалось опубликовать источники в основном второй половины XVII в., около 1000 листов-сставов (276 архивных дел) из 12,5 тыс. листов фонда 98 «Олонецкая воеводская изба» научного архива СПб ИИ РАН. Фонд охватывает документы делопроизводства (челобитные в составе «дел», отписки, памяти, наказы, «выписи», «обыски» и др.) и акты источники. До сетевого размещения коллекции были составлены электронные описи, произведены цифровая фотосъемка, обработка фотокопий, транслитерация, коррек-

тура электронного текста. Интернет-публикация коллекции подготовлена по технологии генерации файлового архива в редакторе «Apple iWeb», который предназначен для облегченной подготовки веб-сайтов и работает в операционной системе Mac OS. Транслитерированные тексты сопровождают цифровые фотокопии, необходимые для «потребностей специалистов-филологов»²⁸ и задач компьютерной археографии. Изображения и тексты документов представлены в графическом виде, поэтому словарный поиск исключен. Навигация ведется по выходным данным и заглавиям документов внутри дел. Синхронно первому этапу проекта составлялись указатели, а текст и фотокопии обрабатывались для последующей публикации в формате pdf. Трудно согласиться с тем, что традиционные указатели имен и географических названий могут заменить поиск данных средствами программы Adobe Acrobat, поскольку орфография и написание имен собственных в XVII в. не устоялись, а поиск по переносам слов, ошибочным написаниям и словоформам оказывается невозможен. Проект создавался в научных и информационно-образовательных целях, не адаптировался к одному исследованию. Авторами ставилась более глобальная задача введения в оборот архивных документов по локальной истории. Электронную коллекцию документов воеводской Олонецкой избы в настоящее время необходимо рассматривать как промежуточную публикацию, полный объем которой мог бы действительно способствовать широкому спектру научных работ.

Относительно роли и значения графических изображений документов, параллельно сопровождающих символьное воспроизведение текста, существуют различные мнения. С точки зрения архивистов наиболее полной публикацией можно считать сочетание графического (приоритетного) и символического воспроизведения²⁹. Но в археографическом контексте символическое воспроизведение текста документа всегда имело ведущую, а не второстепенную роль.

Четвертый проект должен рассматриваться в контексте компьютерной археографии и архивоведения. Он выполнялся коллективом лаборатории исторического краеведения Барнаульского государственного педагогического университета (ныне АлтГПА)³⁰. Работа посвящена электронной публикации «архива» А. В. Беэра³¹ (1696–1751) — выдающегося инженера и организатора горного дела, ставшего первым начальником Кольвано-Воскресенских заводов на Алтае. Более двух тысяч документов уникальной архивной коллекции отражают первые годы выплавки серебра на заводах Акинфия Демидова и деятельность комиссии бригадира Беэра в 1743–1748 гг. Задачи разработки состояли

в ознакомлении научного сообщества и студентов с документами комиссии, архив которой послужил началом комплектования Государственного архива Алтайского края. Описан 391 архивный документ, 18 из них изданы целиком. В основном это документы делопроизводства: рапорты, доклады, письма, контракты и др. Многочисленные разновидности и различия структуры документов, небольшой срок проектирования предопределили объем работы и выбранную методику публикации. Технология объектно-реляционной базы данных (скрипт-программа, написанная на языке PHP) близка первому проекту, но способ презентации полнотекстовых документов значительно упрощен. Поэтому публикация могла следовать «Правилам издания исторических документов» 1990 г. Приведена авторская библиография по изучению «Архива Бера», но опущена часть элементов легенды: сведения о предшествующих публикациях источников и основания их датировки. Электронные тексты не структурированы в реляционных таблицах, выводятся в браузер документами формата MS Word, текстовый поиск возможен только внутри единичных документов средствами текстового редактора. Это говорит о приоритете в базе данных архивного научно-справочного аппарата. Поиск сведений построен по тематике, видам документов, словам из заглавий, номеру единицы хранения и хронологии документов. Сведения таблиц описания архивных документов включают 13 позиций, в том числе современные заглавия документов и датировку. Публикация базы данных стала одним из результатов конкретно-исторического изучения истории Западной Сибири и непосредственной работы авторов с архивными документами.

Анализ четырех археографических веб-сайтов позволил установить, что средний срок подготовки полнотекстовой интернет-публикации объемом до 10 учетно-издательских листов с созданием уникального программно-поискового кода, первичной (впервые осуществляемой) транслитерацией, корректурой текстов при частичной занятости команды из 2–3 участников составляет не менее трех лет. Поэтому для рационального использования времени эдиционные и технологические операции выполняются параллельно или большим составом научного коллектива³².

Последние годы вопрос о необходимости создания единых правил электронной публикации активно не дискутируется. Соглашусь с мнением о бесперспективности стремлений выработать единые археографические правила электронной публикации исторических документов в силу индивидуальной специфики ресурсов и применяемых технологий. Тем не менее актуален ориентир на универсализацию основных эдиционных прин-

ципов, большинство из которых выработаны общей археографией.

Подчеркну, что выбор методик и технологий интернет-публикаций документов делопроизводства XVII–XVIII вв. зависел от целей, опыта, ресурсных возможностей публикаторов. Различными были уровень постановки и решения эдиционных задач, эдиционная мотивация и стратегические авторские установки³³. Как правило, реализуется одна или несколько стратегий развития научных проектов компьютерной археографии: 1) оперативная компьютерная публикация прежде неизданных источников; 2) создание информационной системы для реконструкции комплекса или аккумуляции полнотекстовых документов; 3) предварительная подготовка более масштабного издания или банка данных; 4) публикация с синхронным созданием многовариантных поисковых инструментов; 5) ознакомительная или «демонстрационная» версии.

Сегмент высокотехнологичных завершенных российских проектов компьютерной археографии среди электронных публикаций исторических источников в настоящее время, видимо, не превышает долей процента. В чем причины такой малой доли, как ее расширить? Консервативность и недоверие потенциальных публикаторов к автономным публикациям уже отмечались специалистами³⁴. Многозадачность и трудоемкость компьютерных проектов также очевидна. Одна из весомых причин состоит в том, что долгое время не было сформулировано представлений о предмете, задачах и минимальных требованиях компьютерной археографии. Для достижения приемлемого уровня компьютерной эдиционной работы требуется ряд необходимых условий: обоснованный и объясняемый публикатором выбор документов публикации; корреляция задач технологических (программных) и археографических; обращение к архивному оригиналу или качественной копии публикуемого документа; создание поискового и научно-справочного аппарата публикации. Разумеется, добиться хорошего результата можно разработчикам с определенным уровнем археографических и компьютерных знаний.

Следующая причина узкого сегмента компьютерных археографических публикаций документов делопроизводства кроется в сложности организации таких работ. Успех, как можно убедиться на конкретных примерах, пока возможен только в сотрудничестве историков и специалистов по компьютерным технологиям. Все вышеназванные научные коллективы не имели прежде большого опыта интернет-публикаций, но получив его, вероятно, сталкивались с нетривиальными вопросами планирования и организации труда. Важно

иметь адекватное представление обо всех этапах, объемах работ. В идеальном случае необходим расчет трудозатрат — проделанных и предстоящих. Более всего может быть недооценена трудоемкость транслитерации публикуемого текста и корректуры. В изученных проектах значительное время уделялось подготовке описаний, систематизации, фотокопированию документов. Последнее сложно организовать в необходимые сроки. Несмотря на позитивную динамику технического переоснащения, кадрового расширения, возможности архивных служб копирования документов остаются явно недостаточными. За исключением двух-трех российских вузовских центров пока не созданы стабильно существующие депозитарные центры хранения электронных ресурсов. Отчасти эта задача решается при обращении к государственной регистрации авторского права на создаваемые базы данных и программы для ЭВМ. Закрепление авторского права предполагает частичное депонирование электронных материалов³⁵. Организаторам коллективов требуется вести не только перманентный поиск и сохранение финансирования, но и поддерживать творческий потенциал команды

в обозримой перспективе. Главнейшая роль в финансировании российских проектов создания программного обеспечения принадлежит сейчас РГНФ и РФФИ.

Результаты археографических публикаций средствами компьютерных программ с обеспечением многовариантного поиска данных направлены на увеличение объема новой информации: новых знаний об исторических документах и знаний о созданных технологиях для гуманитариев. Таким образом, круг интересов компьютерной археографии по соотношению с практикой может приобретать в одних случаях фундаментальный, в других — прикладной характер. Электронные публикации новых документов в совокупности с программными инструментами поиска данных, в конечном счете, должны качественно расширить исследовательский инструментарий историков. Даже благодаря нескольким десяткам проектов компьютерной археографии в ближайшие десятилетия может быть решена актуальная задача публикации крупных документальных архивных собраний и создание новых методик широкого доступа к ценным коллекциям.

ПРИМЕЧАНИЯ

- ¹ Валк С. Н. Судьбы «археологии» // Археологический ежегодник за 1961 г. М., 1962. С. 464–465.
- ² Нестерович Ю. В. Проблемы унификации, оптимизации терминологии и построения терминосистемы археологии // Беларускі археографічны штогоднік. Вып. 9. Минск, 2009. С. 75–92; Козлов В. П. Основы теоретической и прикладной археологии. М., 2008.
- ³ Кун Т. Структура научных революций. М., 1975.
- ⁴ Гарскова И. М. Источниковедческие проблемы исторической информатики // Российская история. 2010. № 3. С. 155; Владимиров В. Н. Историческая информатика: от становления к развитию // Крыніцазнаўства і спецыяльныя гістарычныя дысцыпліны: навук. зб. Вып. 3. Мінск, 2007. С. 102–105.
- ⁵ Медушевская О. М. Профессионализм гуманитарного знания в условиях междисциплинарности // Проблемы источниковедения и историографии : материалы II научных чтений памяти академика И. Д. Ковальченко. М., 2000. С. 351.
- ⁶ Репина Л. П. Теоретические новации в современной историографии // Харьковский историографический сборник. Т. 10. Харьков, 2010. С. 13.
- ⁷ Второй близкий термин и понятие «технотронная документальная публикация» получили слабое распространение. С конца XX в. практика любых публикаций не обходится без сложных технических устройств.
- ⁸ Бахтина О. Н., Носков А. В., Хасанов В. Я. Концепция создания компьютерного археографического атласа Томской области // Археологический ежегодник за 1998 г. М., 1999. С. 54; Письменное наследие и современные информационные технологии : сб. статей (Международ. науч. школа, Ижевск, 12–15.10.2009) / отв. ред. В. А. Баранов. Ижевск, 2009.
- ⁹ Иногда в российских, а чаще в украинских исследованиях используется синонимичный термин и аналогичное понятие «электронная археология». См.: Папакін Г. В. Сучасна електронна археологія: предмет, специфіка, стан дослідження // Спеціальні історичні дисципліни: питання теорії та методики. [Т]. 21. Київ, 2013. С. 45–59.
- ¹⁰ См. например: Валк С. Н. Археологическая «легенда» // Археологический ежегодник за 1971 г. М., 1972. С. 8–16.

- ¹¹ Термин «образ исторического документа» не является синонимом «электронного документа». Первый термин семантически передает тезис о невозможности с помощью какой-либо эдиционной технологии передать всю полноту информации документа. Ни одна из самых совершенных публикаций, включая высококачественные цифровые фотокопии, не сможет передать весь набор свойств материала, оттенки и глубину почерка, степень и форму возможных исправлений текста, на что археограф должен факультативно указать. См.: Лаппо-Данилевский А. А. Методология истории / подг. текста Р. Б. Казаков и др. М., 2010. Т. 2. С. 236–237.
- ¹² К аналоговому типу публикаций можно отнести не только «бумажную» публикацию, но и воспроизведение исторических документов в кино- и фотопродукции.
- ¹³ Козлов В. П. Теоретические основы археографии с позиций современности // Отечественные архивы. 2001. № 1. С. 10–33; Каштанов С. М. О предмете и объекте археографии (некоторые замечания по поводу статьи В. П. Козлова) // Отечественные архивы. 2002. № 3. С. 59–63.
- ¹⁴ Лаппо-Данилевский А. А. Указ. соч. С. 29–33.
- ¹⁵ Белоконь Е. А. О возможностях создания новых информационных технологий в археографии // Информационный бюллетень Ассоциации «История и компьютер». М., 1994. № 12. С. 64–69.
- ¹⁶ Боброва Е. В. Анализ археографического уровня подготовки документальных публикаций в российском сегменте Интернет // Информационный бюллетень Ассоциации «История и компьютер». М., 2002. № 30. С. 80–83; Конявская С. В. Издание средневековых памятников письменности по рукописям // Древняя Русь. Вопросы медиевистики. 2012. № 1. С. 39–46.
- ¹⁷ URL: <http://www.hist.msu.ru/ER/index.html>. Основные разработчики проекта Т. Я. Валетов и Л. А. Пухова.
- ¹⁸ Сравнение проектов компьютерной археографии или проектов электронной републикации источников представляет собой отдельную и важную проблему исследований. Имеет место ряд спорных подходов в сравнении различных по целям, задачам, предназначению электронных ресурсов. Очевидно, что эти различия и критерии сопоставления электронных ресурсов должны оговариваться.
- ¹⁹ Информационная полнотекстовая система «Повесточные сказки думных людей XVII–XVIII веков» [Электронный ресурс] / сост. и ред. А. В. Захаров. Дата обновления: 25.09.2008. URL: <http://zaharov.csu.ru/povestki.pl>
- ²⁰ Захаров А. В. Государев двор Петра I: публикация и исследование массовых источников разрядного делопроизводства. Челябинск, 2009.
- ²¹ Информационно-поисковая полнотекстовая система «Боярские списки XVIII века» [Электронный ресурс] / сост. и ред. А. В. Захаров. Дата обновления: 29.01.2014. URL: <http://zaharov.csu.ru/bspisok.pl>
- ²² Stancelis V. Lietuvos parlamentarų žodyno duomenų bazės projektas [Electronic resource] / URL: http://www.parlamentostudijos.lt/Nr7/7_kalba_stancelis.htm (дата обращения: 19.01.2012.).
- ²³ Zakharov A. Высшие чины государева двора и механизмы обеспечения службы в России 1690-х — 1700-х годов // Projets. Drerupa, Paris, 12–24 septembre 2005. P. 131–134. [препринт].
- ²⁴ Chartes de l'abbaye de Saint-Denis (VIIe — XIIIe siècle) [Electronic resource]. URL: www.saint-denis.enc.sorbonne.fr (дата обращения: 29.01.2013).
- ²⁵ Clavaud F. The digital edition of the medieval charters of the Abbey of Saint-Denis: first results and prospects // Digital Medievalist. 2012. № 8. [Electronic resource]. URL: <http://www.digitalmedievalist.org/journal/8/clavaud/#Guyotjeannin1999> (дата обращения: 28.01.2014).
- ²⁶ К транслитерации текстов были привлечены 19 студентов и аспирантов ПетрГУ, корректура осуществлялась М. Е. Проскуряковой и Е. Д. Сусловой. Электронная опись фонда составлена сотрудниками архива З. Н. Исидоровой, Г. А. Победимовой при участии О. А. Абеленцевой. Технический директор проекта О. В. Черняков.
- ²⁷ Чернякова И. А. Олонецкая воеводская изба — полнотекстовая электронная база данных для новейших исследований по истории Карелии раннего Нового времени [Электронный ресурс]. URL: <http://illmik.petrus.ru/illmik/IZBA.html> (дата обращения: 29.01.2014).
- ²⁸ Там же.
- ²⁹ Грум-Гржимайло Ю. В., Сабенникова И. В. Некоторые проблемы публикации архивных документов в электронных изданиях // Вестник архивиста. 2006. № 2–3. С. 309–310.
- ³⁰ Проект финансировался двухлетним грантом РГНФ и Институтом «Открытое общество». Научный руководитель А. В. Контев, В. Б. Бородаев, веб-программирование А. А. Данькин.
- ³¹ База данных «Архив Комиссии А. В. Безра» (по коллекции Центра хранения архивного фонда Алтайского края) [Электронный ресурс]. URL: www.uni-altai.ru/ArchiveBeyer/index.html (дата обращения: 29.01.2014).

- ³² Более сложные и масштабные по объему проекты, выходящие за пределы трех лет, практически невозможно спланировать из-за множества влияющих обстоятельств и неясной «отдачи» промежуточных результатов.
- ³³ Эдиционные задачи — выбор документов для публикации, транскрипция, создание научно-справочного аппарата. Эдиционная мотивация — движущие факторы выбора формы и объема публикации.
- ³⁴ Варфоломеев А. Г., Иванов А. С. Принципы электронных публикаций комплексов исторических документов со средствами палеографического, текстологического и дипломатического анализа // Тезисы конференции E1Manuscript-2008. [Электронный ресурс] URL: www.textualheritage.org/content/view/57/68 (дата обращения: 28.01.2013).
- ³⁵ После передачи функции по регистрации баз данных ФГУП «Информрегистр», с 2009 г. государственную регистрацию баз данных и программ для ЭВМ ведет ФГБУ «Федеральный институт промышленной собственности», требующий для учета базы данных выборку до 50 листов.
-