

Исторические информационные системы : теория и практика

С. И. Корниенко
Д. А. Гагарина
Н. Г. Поврозник



Издательский дом
Высшей школы экономики
МОСКВА, 2021

УДК 004:930
ББК 16.33+63.3
К67

Рецензенты:

доцент Школы исторических наук факультета гуманитарных наук
Национального исследовательского университета
«Высшая школа экономики», к.и.н.

Д.А. Добровольский;

заведующая кафедрой документоведения, архивоведения
и истории государственного управления

Уральского федерального университета, д.и.н., доцент

Л.Н. Мазур;

доцент кафедры исторической информатики исторического факультета
Московского государственного университета им. М.В. Ломоносова,
директор Института цифровых гуманитарных исследований
Сибирского федерального университета, к.и.н.

А.Ю. Володин

Корниенко, С. И., Гагарина, Д. А., Поврозник, Н. Г. Исторические информационные системы: теория и практика [Текст] : монография / С. И. Корниенко, Д. А. Гагарина, Н. Г. Поврозник ; Нац. исслед. ун-т «Высшая школа экономики». — М. : Изд. дом Высшей школы экономики, 2021. — 231, [1] с. — 600 экз. — ISBN 978-5-7598-2321-6 (в обл.). — ISBN 978-5-7598-2426-8 (e-book).

Исторические, или историко-ориентированные, информационные системы — значимый элемент информационной среды гуманитарных наук. Его выделение связано с развитием исторической информатики и историко-ориентированного подхода, формированием информационной среды, практикой создания исторических ресурсов.

Книга содержит результаты исследования теоретических и прикладных проблем создания и внедрения историко-ориентированных информационных систем. Это первое комплексное исследование по данной тематике. Одни проблемы в книге рассматриваются впервые, другие — хотя и находили ранее отражение в литературе, но не изучались специально.

Издание адресовано историкам, специалистам в области цифровой истории и цифровых гуманитарных наук, а также разработчикам цифровых ресурсов, содержащих исторический контент или ориентированных на использование в исторических исследованиях и образовании.

УДК 004:930
ББК 16.33+63.3

Опубликовано Издательским домом Высшей школы экономики
<http://id.hse.ru>

doi:10.17323/978-5-7598-2321-6

ISBN 978-5-7598-2321-6 (в обл.)
ISBN 978-5-7598-2426-8 (e-book)

© Корниенко С.И., Гагарина Д.А.,
Поврозник Н.Г., 2021

Оглавление

Введение	6
Глава 1. Исторические информационные системы: понятие, классификация и оценка.....	10
1.1. Исторические информационные системы как информационные ресурсы.....	10
1.2. Исторические информационные системы: понятие и эволюция в свете развития исторической информатики.....	16
1.3. Историко-ориентированные базы данных	23
1.4. Классификация исторических информационных систем	28
Классификация по назначению.....	29
Классификация по областям исторического знания	34
Классификация по типу исторических источников	35
1.5. Каталог историко-ориентированных информационных систем.....	37
1.6. Анализ и оценка исторических информационных систем ..	41
Глава 2. Создание исторических информационных систем	50
2.1. Принципы, возможности и ограничения различных подходов.....	50
Источнико-ориентированный подход	50
Проблемно-ориентированный подход.....	53
Смешанный подход	54

2.2. Модели исторических информационных систем	54
2.3. Системы на основе реляционной модели	61
2.4. XML-разметка текстов и исторические системы.....	67
2.5. Привлечение пользователей к созданию исторических информационных систем и ресурсов	82
2.6. Онлайн-сервисы создания исторических информационных систем.....	88

Глава 3. Информационные системы в сохранении и репрезентации историко-культурного наследия 95

3.1. Информационные системы в библиотеках	96
3.2. Информационные системы в архивах.....	107
3.3. Информационные системы в музеях. Виртуальные музеи	117
3.4. Пространственная репрезентация исторического и культурного наследия	129
3.5. Взаимодействие государства, организаций хранения, исследовательских учреждений и других институтов.....	138
Международное сотрудничество.....	141
Каталоги и системы документирования цифрового историко-культурного наследия	144
«Рожденный цифровым». Веб-архив как тип информационной системы	147

Глава 4. Информационные системы как средство поддержки научных исторических исследований..... 151

4.1. Компьютеризированные исторические исследования и исследовательский инструментарий систем.....	151
4.2. Методика проведения исторических исследований различного типа	162
Конкретно-историческое исследование.....	163
Источниковедческое исследование	165
Историографическое исследование	167

4.3. Информационные системы в изучении истории России.....	167
Системы по истории земского самоуправления в России.....	168
Системы на основе губернской периодики.....	172
Парламентская история позднеимперской России.....	178
Глава 5. Информационные системы в историческом образовании	183
5.1. Педагогика Digital Humanities	183
5.2. Дидактические возможности исторических информационных систем.....	185
5.3. Информационно-образовательные исторические системы	192
5.4. Обучение разработке исторических информационных систем и баз данных	198
Заключение	202
Библиографический список	206
Список упоминаемых в тексте интернет-ресурсов	221
Информация об авторах	230

Введение

Существенную часть информационных ресурсов гуманитарных наук и образования составляет историческая информация. Со временем увеличивается ее объем, растет количество и разнообразие источников данных — исторических источников. В этих условиях важными становятся подходы к извлечению, организации, хранению и обработке исторической информации, выбору соответствующих средств и методов, а также обеспечение их научности, строгости и верифицируемости. В значительной степени указанные задачи позволяют решать исторические информационные системы, которым посвящена представляемая книга.

Информационные системы параллельно с решением вычислительных задач практически сразу становятся одним из главных направлений развития ЭВМ. Их специфика требует структурированной организации, основой чего наиболее часто являются базы данных. К концу 1980-х — началу 1990-х годов базы данных и информационные системы начинают исследоваться в аспектах исторической науки и исторической информации.

Исторические, или историко-ориентированные, информационные системы (мы будем употреблять эти понятия как синонимичные) — значимый тип исторических ресурсов и важный элемент информационной среды гуманитарных наук и образования. Выделение этого класса систем связано с несколькими тенденциями: развитием исторической информатики, цифровой истории и историко-ориентированного подхода; формированием информационной среды истории и гуманитарных наук в целом, инфраструктуры цифрового истори-

ческого и культурного наследия; практикой создания исторических электронных ресурсов, в частности баз и банков данных.

Представленное издание содержит результаты исследования методологических, теоретических и прикладных проблем создания и использования историко-ориентированных информационных систем. Монография является первым комплексным исследованием по заданной теме. Одни проблемы в книге ставятся и рассматриваются впервые, другие — хотя и находили ранее свое отражение в соответствующей литературе, не изучались специально и столь глубоко.

В главе 1 монографии рассмотрены изменения в информационной среде исторической науки, дана характеристика информационных ресурсов различного типа. На основе анализа историографии и многолетней практики авторов сформулировано и обосновано определение исторических информационных систем; разработана методика их описания, классификации и оценки; проанализированы роль и место систем различного типа. Значимой частью работы является каталог «Историко-ориентированные информационные системы», который, с одной стороны, стал организационной основой исследования, а с другой — его практическим итогом и апробационной площадкой для полученных теоретико-методических результатов. Характеристика ресурса также приведена в этой главе.

В главе 2 комплексно рассмотрены вопросы разработки исторических информационных систем; проанализированы существующие подходы, их возможности и ограничения; разработаны обобщенные модели. Наибольшее внимание уделено системам на основе реляционной модели данных и технологий разметки текста XML-TEI. С точки зрения исторической специфики в этой главе описаны две важные текущие тенденции создания цифровых (в том числе исторических) ресурсов — краудсорсинг для реализации крупномасштабных проектов и массовое развитие доступных сервисов.

Глава 3 посвящена возможностям информационных систем в работе с историко-культурным наследием; проанализированы потенциал, роль и место рассматриваемого вида ресурсов в библиотеках, архивах и музеях. Глава содержит как технологические аспекты,

в частности создание 3D-моделей и виртуальных экспозиций, так и инфраструктурные: взаимодействие государства, организаций хранения и исследовательских учреждений при создании информационных систем, международное сотрудничество, документирование цифрового историко-культурного наследия.

В главе 4 издания описаны уровни поддержки и методики проведения конкретно-исторических, источниковедческих и историографических исследований на основе историко-ориентированных информационных систем, а также приведены примеры таких исследований. Кроме того, в главе уделено внимание возможностям информационных систем в решении специфических задач исторической науки.

Глава 5 полностью посвящена педагогическим, методическим и дидактическим вопросам. Приведены примеры, анализ и методика применения систем, имеющих изначально образовательный характер, а также варианты адаптации системы поддержки научных исследований, каталогов исторических источников и ресурсов другого типа. Кроме того, описаны методы и форматы обучения студентов разработке информационных систем с учетом исторической специфики.

Результаты, описанные в книге, были получены в ходе реализации ряда научных проектов, работа над которыми велась с 2003 г. в Центре цифровой гуманитаристики (до 2016 г. — Лаборатория исторической и политической информатики) Пермского университета, а в настоящее время продолжается на кафедре гуманитарных дисциплин НИУ ВШЭ — Пермь. Одним из наиболее значимых в этой серии стал проект «Историко-ориентированные информационные системы: методологические, теоретические и прикладные проблемы создания и использования», поддержанный Российским фондом фундаментальных исследований и реализованный в 2013–2015 гг. (грант РФФИ № 13-06-00655). Пользуясь случаем, авторы выражают благодарность фонду за многолетнюю поддержку проектов по созданию исторических информационных систем и проведению научных исследований на их основе.

Такое широкое по отражаемой проблематике исследование стало результатом работы целого коллектива. Кроме авторов книги в нем принимали участие сотрудники Центра цифровой гуманитаристики Пермского университета и кафедры гуманитарных дисциплин НИУ ВШЭ — Пермь. Авторы выражают искреннюю благодарность за предоставление материалов для монографии и общий вклад в работу молодым ученым О.В. Власовой, Ю.В. Яценко, А.Р. Ехлаковой, И.Д. Исмакаевой.

Представленное издание не могло бы состояться без внимательного прочтения вариантов рукописи нашими коллегами. Авторы признательны за советы и замечания А.Ю. Володину (МГУ им. М.В. Ломоносова), Л.Н. Мазур (УрФУ), А.А. Мироносу (ННГУ им. Н.И. Лобачевского). Благодаря критическому отношению к рецензируемой рукописи и детальным рекомендациям Д.А. Добровольского (НИУ ВШЭ) текст приобрел большую аккуратность, были уточнены некоторые классификации и обновлены описания отдельных ресурсов.

Глава 1. Исторические информационные системы: понятие, классификация и оценка

В этой главе рассматриваются состояние и тенденции развития исторических информационных систем на уровне понятийного аппарата, содержания, используемых исторических источников и других параметров. Даются классификация и характеристика некоторых типов систем. Предложена методика оценки историко-ориентированных информационных систем.

1.1. Исторические информационные системы как информационные ресурсы

Информатизация привела к коренным изменениям в информационной среде и структуре исторической науки. Наиболее отчетливо и полно это выразилось в возникновении междисциплинарных направлений на стыке гуманитарных наук, информатики и математики — гуманитарного компьютеринга (*Humanities Computing*), клиометрии, количественной истории, исторической информатики (*Historical Information Science*), цифровых гуманитарных наук (*Digital Humanities*), цифровой истории и др.

Гуманитарии используют компьютеры для решения профессиональных задач практически с начала появления ЭВМ — с 1950-х годов. Наиболее ярким примером в области компьютерной лингвистики и использования информационных технологий (ИТ) для обработки исторических текстов является проект *Index Thomisticus*, разра-

ботанный священником и ученым Роберто Бусой¹. В это же время появляются исторические исследования с использованием ЭВМ, в основном для автоматизации обработки статистических данных [Thomas, 2004]. Эта связь между количественными историческими исследованиями и источник-ориентированной обработкой данных (которая впоследствии стала основой источник-ориентированного подхода к созданию баз данных) проанализирована в статье М. Таллера [Thaller, 1995]. По мнению автора, несмотря на то, что они часто рассматриваются как противоречия, фактически исключающие друг друга, на самом деле корни источник-ориентированного подхода лежат в традиции количественных исследований.

До 1980-х годов число гуманитарных исследований с применением компьютеров было невелико; по мнению ряда авторов, они носили случайный характер [Alkhoven, Doorn, 2007; Feeney, Ross, 1993]. С 1980-х годов, с появлением микропроцессорной техники, компьютеры проникают уже во все гуманитарные науки, количество проектов стремительно растет. Но рост и значение ИТ неравномерно распределяются среди разных направлений гуманитарных наук, в некоторых областях роль компьютеров становится более важной, чем в других [Alkhoven, Doorn, 2007]. Неравномерность наблюдается и в отношении типов изучаемых документов — текстовых, статистических, визуальных, звуковых, устных источников и материальных объектов.

Развитие исторической информатики и других направлений, в настоящее время объединенных Digital Humanities, связано не только с научными исследованиями, но и с созданием электронных ресурсов. Необходимость их разработки определяется всем процессом цифровой трансформации общества, переходом к цифровой культуре. С ростом числа компьютеров, а затем с появлением и распространением Интернета меняются способы создания, поиска, обработки и публикации источников [Alkhoven, Doorn, 2007]; увеличиваются

¹ Index Thomisticus [Электронный ресурс]. URL: <http://www.corpusthomicum.org/it> (дата обращения: 10.07.2020).

количество и разнообразие цифровых информационных ресурсов, их удельный вес в информационной среде общества.

Заметной частью новой инфраструктуры гуманитарных наук в целом и исторической науки в частности являются предметно-ориентированные и адаптированные к этим областям информационные ресурсы и инструментарий. Исторические информационные ресурсы разнообразны по проблематике и назначению. Они могут быть продуктом академической и публичной истории, разрабатываться профессиональными историками, историками-любителями, студентами и даже школьниками. Это как продукты для научного изучения, продукты-результаты такого изучения, так и продукты для исторического образования и просветительских целей [Корниенко, Власова, Гагарина, 2012].

Среди исторических информационных ресурсов существенное место занимают исторические (историко-ориентированные) информационные системы. Этот вид ресурсов как форма и средство сохранения, организации, обработки и визуализации исторической информации позволяет эффективно решать проблемы учета своеобразия, содержательных и формальных особенностей исторических данных, знаний и науки в целом. Будучи сегодня одним из основных компонентов информационных ресурсов исторической науки и существенной частью информационных ресурсов социально-гуманитарных наук, исторические информационные системы, а также процессы их создания и использования отражают все характерные черты их природы и применения. В этом смысле приобретает определенную важность понимание тех свойств и характерных особенностей исторических информационных систем, которые вытекают из их принадлежности к информационным ресурсам вообще и историческим ресурсам в частности. Это относится прежде всего к определению понятия *«исторический информационный ресурс»*. Для определения, классификации и анализа указанного вида ресурсов недостаточно интуитивного понимания исторической специфики и понимания информационных ресурсов вообще [Корниенко, Власова, Гагарина, 2012]. Как и в отношении ресурсов в целом, при-

менительно к историческим информационным системам такой уровень понимания не обеспечивает нужной четкости определения характера, содержания, установки границ и критериев для отнесения ресурсов к историческим. Более того, при определенном подходе можно расширить границы содержания рассматриваемого понятия на весь круг ресурсов вообще. Очевидно, что с течением времени любые результаты, остатки и свидетельства деятельности человека станут историческими свидетельствами, т.е. историческими источниками. Такой подход к указанному понятию правомерен, поскольку он наиболее полно раскрывает то общее, что характеризует исторические информационные ресурсы как информационные ресурсы вообще. Однако в этом одновременно состоит и недостаток данного подхода. При нем и на формальном, и на содержательном уровнях исчезает специфика исторических информационных ресурсов. Поэтому существует и более строгий подход, основанный на учете природы и предназначения таких ресурсов. В этом случае к историческим информационным ресурсам относят только те, которые созданы специалистами в области исторических исследований или обучения. В то же время такая четкость и определенность оставляют за рамками рассмотрения значимую для исторической науки и образования часть ресурсов, например цифровые музеи, коллекции, библиотеки и тематические сайты междисциплинарного характера, а также любительские ресурсы.

Преодолеть как излишне расширительное толкование, так и чрезмерную строгость в определении понятия «исторический информационный ресурс» позволяет подход, при котором основным критерием служит отношение к предметной области — истории. Иными словами, целесообразно относить к историческим те ресурсы, в которых тем или иным образом присутствует историческая тематика. Такой критерий дает возможность определения исторических информационных ресурсов, минуя имеющиеся противоречия в подходах к этому понятию в научной литературе [Антопольский, 2004, с. 37–45] и законодательстве, а также дискуссии о предмете изучения истории как науки.

Представляется, что указанный подход и вытекающее из него определение достаточно продуктивны для оценки качества, исследовательской и образовательной эффективности использования, дальнейшего развития методик и инструментов создания и применения исторических ресурсов.

Как еще в начале 2000-х годов отмечал А.Б. Антопольский, разработка стандартов описания информационных ресурсов, а следовательно, и качество их документирования во многом связаны с содержанием понятия «*информационные ресурсы*» и их классификациями [Антопольский, 2004, с. 37–45]. Разработка таких классификаций и стандартов описания связана с классификацией информационных ресурсов соответственно структуре и содержанию предметной области, для которой они предназначены. В описании исторических информационных ресурсов могут использоваться: атрибуты, общие для информационных ресурсов; исторические атрибуты, в том числе применяемые для описания исторических источников и литературы; назначение ресурсов; типы данных ресурсов; тематические атрибуты [Корниенко, Власова, Гагарина, 2012].

Существенное значение с точки зрения возможностей использования ресурсов имеет классификация по типу данных, которым представлена информация, и по соответствующим информационным технологиям. В этом смысле информационные ресурсы подразделяются на текстовые, гипертекстовые, числовые, графические (в том числе 3D, 4D), аудио-, видео-, мультимедиа-, интернет-ресурсы и др. Такая классификация может использоваться, в частности, для электронных версий источников и объектов культурного наследия — в зависимости от типа источника его электронная версия может быть представлена в виде текстовой, графической (в том числе 3D-графической), аудио- или видеoinформации. Те или иные ресурсы, как правило, являются комбинацией применения различных видов технологий [Корниенко, Власова, Гагарина, 2012].

Еще один аспект анализа исторических информационных систем как информационных ресурсов связан с их значительным распространением в Интернете. Исторические информационные системы

как интернет-ресурсы представлены сегодня историко-ориентированными личными и коллективными, многопрофильными и проблемно-тематическими сайтами, порталами, сервисами, форумами, хранилищами данных, веб- и цифровыми историческими архивами, поисковыми системами. Наличие исторического характера и содержания в блогах, электронной переписке, тематических группах и на страницах в социальных сетях превращает содержащие их информационные системы в уникальные современные исторические интернет-ресурсы, обогащает их многообразие и возможности, повышает значимость.

Классификация исторических интернет-ресурсов может проводиться по разным критериям: по цели, содержанию, институциональному принципу, а также по совокупности критериев. Интерес представляют такие группы исторических интернет-ресурсов, как ресурсы, созданные для организации, хранения, обработки и представления массивов исторических источников и публикаций (национальные и международные электронные архивы; электронные библиотеки, каталоги, коллекции; виртуальные музеи; электронные исторические журналы и сайты традиционных периодических изданий); проблемно-тематические исторические интернет-сайты и порталы; сайты исторических научных и образовательных учреждений, а также учреждений хранения и национальной памяти (сайты исторических и гуманитарных факультетов университетов и исследовательских учреждений; сайты музеев, архивов, библиотек) [Корниенко, Власова, Гагарина, 2012].

Изучение исторических информационных систем на теоретическом и прикладном уровнях показывает, что к настоящему времени накоплен определенный опыт их классификации и описания, однако комплексных и общепринятых в среде историков систем классификации и стандартов описания на сегодняшний день не существует. В качестве основной проблемы их создания следует назвать значительное функциональное и содержательное разнообразие, а также быстрое технологическое усложнение. Преодолеть эти препятствия в отношении исторических информационных систем можно на основе анали-

за основных понятий, эволюции теоретических оснований, практики создания историко-ориентированных систем, разработки критериев для их классификации, выделения основных видов, типов и характеристик. Решению этих задач в значительной степени и посвящены следующие разделы.

1.2. Исторические информационные системы: понятие и эволюция в свете развития исторической информатики

В информатике в широком смысле под информационной системой понимают программный комплекс, к функциям которого относятся обеспечение хранения информации в памяти компьютера, выполнение специфических преобразований информации и предоставление интерфейса [Кузнецов, 2005]. Понятие «*информационные системы*» охватывает достаточно разнообразные ресурсы.

Первые специализированные публикации по проблематике использования информационных систем в исторической науке появляются в конце 1980-х — начале 1990-х годов. Среди таковых работа И. Кропача, в которой предложено толкование термина «*историческая информационная система*» [Кропач, 1988]. Как было указано во введении, значительную роль в формировании рассматриваемого класса систем сыграли историческая информатика, историко-ориентированный подход, специфика информационной среды истории и гуманитарных наук в целом, а также практика создания исторических электронных ресурсов, в частности баз данных. Эти аспекты могут быть отправными точками при определении и обосновании указанного понятия, выделении его видовых признаков, разработке классификации, требований, стандартов описания и т.д.

На основе изучения литературы, анализа и обобщения опыта разработки и применения рассматриваемых систем нами было предложено следующее определение: *историко-ориентированные информационные системы* — это особый класс информационных систем,

предназначенных для хранения, организации исторической информации, обеспечения доступа к ней и ее аналитической обработки в соответствии с потребностями исторических исследований, образования и популяризации исторических знаний². Данное определение соответствует имеющимся в литературе общим определениям понятия «*информационная система*» и отражает сущность историко-ориентированного подхода. Оно позволяет при всем многообразии историко-ориентированных информационных систем выделять их в особый класс на основе достаточно четких критериев, учитывающих специфику информации, методов, технологий ее обработки и визуализации, назначения и функций, связанных с реализацией подходов и методов исторических исследований, решением научных и (или) образовательных, просветительских задач в области истории. На основе данного определения достаточно удобно выстраивается иерархия понятий, связанных с историко-ориентированными информационными системами. Оно корректно соотносится с понятиями верхнего уровня — «*информационные ресурсы*», «*информационные системы*», «*базы данных*», «*исторические ресурсы*», а также с понятиями нижнего уровня — «*исторические базы данных*», «*просопографические информационные системы*», «*историко-ориентированные геоинформационные системы (ГИС)*» и др. В предложенное определение вполне укладываются и различные потенциальные параметры функционального назначения систем как информационных ресурсов: *научно-исследовательские, образовательные, просветительные, политические, формирование и сохранение исторической памяти* и др.

На исторические информационные системы важно посмотреть и с точки зрения развития исторической информатики.

Историческая информатика (Historical Information Science) — междисциплинарное направление, изучающее теоретические и при-

² Это определение было впервые нами предложено в: *Кирьянов И.К., Корниенко С.И., Гагарина Д.А., Рябухин И.В.* Информационный ресурс по парламентской истории России начала XX в. // *Власть*. 2010. № 12. С. 83–86. В текущем варианте добавлено «...и популяризации исторических знаний», чего не было в исходной редакции.

кладные вопросы использования информационных технологий в исторической науке и образовании. Это направление приобретает свои очертания в середине 1980-х годов с распространением персональных компьютеров. Л. МакКрэнк, анализируя развитие данной области в период с 1984 по 2000 г., пишет, что историческая информатика вырастает из количественных методов в истории и социальных науках [McCrank, 2002]. Сам термин приходит на смену использовавшемуся ранее понятию «*история и вычисления*» (*History and Computing*), введенному по названию одноименного журнала. В отчете «Прошлое, настоящее и будущее исторической информатики» отмечается, что смена термина была удачной, так как *история и вычисления* — это «неопределенно и запутанно»; при этом *историческая информатика* не является ни *историей*, ни *вычислением*. «Это самостоятельная наука, имеющая собственные методологические рамки», — отмечают О. Бонстра, Л. Бреуре, П. Доорн [Boonstra, Breure, Doorn, 2006]. Предметом исторической информатики является историческая информация, а также способы ее создания, проектирования, редактирования, анализа и представления с помощью информационных технологий [McCrank, 2002].

Российскими специалистами в области исторической информатики Л.И. Бородкиным, И.М. Гарсковой, В.Н. Владимировым и др. были поставлены существенные методологические и прикладные проблемы, связанные с критическим анализом пройденного исторической информатикой пути и определением направлений ее развития [Бородкин, 2005; Бородкин, Гарскова, 2011; Владимиров, 2006; Гарскова, 2018]; среди них также создание и использование историко-ориентированных систем.

Несмотря на стремительное развитие, теоретические обоснования, очевидную практическую значимость и неоднократно описанные в научной литературе перспективы, в настоящее время термин «*Historical Information Science*» практически не используется в англоязычной литературе. Во многом область исследований и разработок, ранее относившихся к исторической информатике, сейчас рассматривается в русле *цифровой истории* (*Digital History*), а также *циф-*

ровой гуманитаристики (*Digital Humanities*) как зонтичного термина (*umbrella term*), объединяющего использование информационных технологий во всех гуманитарных науках. В русскоязычной литературе термин «историческая информатика» используется и сейчас, что связано с активной деятельностью Ассоциации «История и компьютер» и кафедры исторической информатики МГУ им. М.В. Ломоносова. При этом международная ассоциация *History & Computing* прекратила работу, проведя последнюю конференцию в Амстердаме в 2005 г., а после 2010 г. перестал издаваться и журнал *History and Computing*. Отмеченные изменения в эволюции направления, определяемого за рубежом понятием «*Historical Information Science*», а в России — «историческая информатика», привели к неоднозначному пониманию в нашей стране соотношения этого и других терминов, обозначающих в настоящее время процессы информатизации в исторических науке и образовании. Так, И.М. Гарскова в указанной монографии, констатируя факт «сходства в траекториях развития “исторического компьютеринга” на Западе и исторической информатики в России до “точки бифуркации” в середине 2000-х годов», говорит об «их кардинальном различии в последнее десятилетие, отмеченном трансформацией “исторического компьютеринга” в “цифровую историю” после нее» [Гарскова, 2018].

Однако такой взгляд на соотношение современных терминов, отражающих определенные направления в областях исторического и гуманитарного компьютерингов, не является единственным. Не все ученые разделяют мнение о столь резком противопоставлении понятий и соответствующих направлений, разрыве преемственности и утрате сходства. Еще в 1988 г. М. Таллер заметил, что предмет и методы историков настолько разнообразны, что историки часто оказываются ближе связанными с исследователями из других сфер, чем с другими направлениями исторической науки [Thaller, 1988].

Нельзя не признать, что историческая информатика не нашла достаточного и заслуженного признания в среде традиционных историков. На это обращали внимание исследователи уже в 1990-х и 2000-х годах; очевидно, это можно констатировать и на текущем

этапе. Так, в 1994 г. В.А. Спек отмечал: «...факт остается фактом: профессия по-прежнему разделена между небольшим меньшинством историков, которые используют компьютеры в качестве инструментов для анализа исторических данных, и подавляющим большинством, которые, хотя и могут использовать компьютер для набора текста, не думают о том, что он может стать методологическим активом» [Speck, 1994³]. А так Л. МакКрэнк пишет об ученых-историках в 2002 г.: «Как люди такого характера, интеллекта, образования и позиции влияния могут быть настолько устойчивы к информационным технологиям? Могут ли они не видеть возможности, которые предлагают технические достижения нашего времени?» [McCrank, 2002]. Достаточно подробный обзор противостояния традиционных историков и тех, кто использовал количественные методы и технологии, приведен в статье У.Дж. Томаса «Вычисления и историческое воображение» [Thomas, 2004].

Не вдаваясь в причины указанных изменений и проблем, а также в соотношение терминов «*историческая информатика*», «*цифровая история*», «*цифровая гуманитаристика*» (как было отмечено, существуют разные подходы и достаточное количество литературы по данному вопросу [Бородкин, 2005; Бородкин, 2012; Владимиров, 2006; Гарскова, 2014]), отметим, что формирование и методология исторической информатики как самостоятельного междисциплинарного направления напрямую проецируются на понятийный аппарат и идеологию создания и использования историко-ориентированных информационных систем. Именно в связи с этим становление и развитие исторической информатики и ее подходов представляют ценность для понимания феномена историко-ориентированных информационных систем.

Выделим несколько важных, на наш взгляд, аспектов.

Во-первых, точно так же как неологизм «*историческая информатика*» не является просто сочетанием *истории* и *информатики*

³ Цит. по: Anderson I. History and computing // Making History [Электронный ресурс]. URL: http://www.history.ac.uk/makinghistory/resources/articles/history_and_computing.html (дата обращения: 27.08.2018).

(здесь скорее в смысле *Information Science*, а не *Computer Science*), давая больше, чем сумму этих частей [McCrank, 2002], *историческая (историко-ориентированная) информационная система* — это больше, чем просто сумма *истории* (исторических данных, знаний) и *информационных систем*.

Во-вторых, *историческое* в исторической информатике понимается в самом широком смысле, оно «не ограничивается историей как дисциплиной, также история является краеугольным камнем других гуманитарных наук» [Boonstra, Breure, Doorn, 2006].

В-третьих, исторические информационные системы, как и историческая информатика, используют практически все виды исторических источников — числовые, текстовые, материальные, устные, аудиовизуальные и др. В некотором смысле историческая информатика даже способствует более активному внедрению в научный оборот нетекстовых источников. С точки зрения хронологии, географии и проблематики исторического исследования историческая информатика также не имеет ограничений. Наконец, «каждая технология и методология кажутся восприимчивыми к принятию историками, адаптированными к историческим исследованиям... Никаких ограничений, кроме самоналоженных, не видно» [McCrank, 2002]. Эта «всеядность» исторической информатики определяет разнообразие историко-ориентированных информационных систем и множество оснований для их классификации (к чему мы вернемся позже в этой главе).

В-четвертых, историческая информатика не просто исходит из понимания истории как науки и как направления, имеющего научную базу и доказательные методы документации и интерпретации (не принимая постмодернистские трактовки истории и «рассказы об истории»), но и усиливает *научность истории*. Л. МакКрэнк отмечает в связи с этим, что информатику можно рассматривать как «судебную экспертизу истории»; что задача исторической информатики — не просто информирование, а «создание основы для приемлемости, надежности и достоверности информации и ее интерпретации в контексте, как ближайшем и непосредственном, так и в более отда-

ленно» [McCrank, 2002]. В этом смысле особенно заметны роль и наследие квантитативной истории и математических методов в истории в формировании исторической информатики. Несмотря на то, что методологически, содержательно и технологически историческая информатика выходит за пределы квантитативной истории, тем не менее в требованиях системности, верифицируемости, объективности и доказательности — не уступает ей.

Эти повышенные (по сравнению со средними по истории в целом) ожидания и требования научности исторической информатики определяют принципы и подходы к проектированию, наполнению и дальнейшему использованию исторических информационных систем. Хотя при заполнении базы данных, разметке исторического источника, нанесении объектов на географическую карту или реконструкции памятника исследователь также вынужден принимать субъективные решения, есть значимое отличие в их характере по сравнению со многими другими направлениями исторической науки: эти операции с историческими сущностями более четко фиксируются и, как правило, предполагают простые инструменты корректировки при уточнении данных либо даже позволяют заложить в модель информацию о достоверности тех или иных источников. Одновременно акцент исторической информатики на научности истории делает особо значимыми информационные системы для поддержки научных исторических исследований. Д.А. Гутнов и В.А. Перевертень выделили в отдельный класс компьютерные системы, созданные для поддержки информационно-познавательной деятельности историков. Для обозначения этого понятия авторы предложили использовать термин *«историко-исследовательские информационные системы»* [Гутнов, Перевертень, 1994]. К этому типу систем и их использованию в исследованиях мы вернемся в главе 4.

Таким образом, исторические информационные системы определенно открывают новые перспективы организации и массовой обработки исторического материала, большей формализации исторического знания и применяемых в нем доказательств.

1.3. Историко-ориентированные базы данных

Выделение класса и термина «историко-ориентированные информационные системы» связано с практикой создания исторических баз данных, которые долгое время были и все еще часто остаются технологической основой информационных систем. Развитие понятийного аппарата, подходы к созданию и классификация, а также другие теоретические и прикладные вопросы тесно связаны и во многом схожи для исторических баз данных и информационных систем. Можно отметить и такой признак, как типичное объединение двух видов ресурсов в одну секцию на профильных конференциях по исторической информатике и цифровой гуманитаристике.

Феномен внимания историков и других гуманитариев к технологии баз данных интересен сам по себе. «База данных становится центром творческого процесса в компьютерную эпоху», — пишет Л. Манович в 1999 г. Анализируя базы данных как культурную форму, он приходит к выводу, что база данных и нарратив являются «естественными врагами»: база данных представляет предметную область как неотсортированный список; нарратив же, напротив, создает линейное, а значит, последовательное и, возможно, причинно-следственное повествование, которое кажется неупорядоченными событиями. «Конкурируя на одной и той же территории человеческой культуры, каждый из них претендует на исключительное право сделать смысл из мира» [Manovich, 1999]. В некотором роде базы данных являются новой, текущей и будущей формой знаний, заменяя повествование в изучении истории и литературы [Drucker, Kim, Salehian, Bushong, 2013].

Базы данных относятся к эффективным средствам структурирования и организации информации. В зависимости от модели, которая положена в основу базы данных, она может быть иерархической, сетевой или реляционной. Возможность организации данных в базе на основе разных моделей позволяет наиболее адекватно представлять

информацию об отображаемых объектах и эффективно работать с ней. Проектирование и создание информационных систем на основе реляционных баз данных будет подробно рассмотрено в главе 2; задача этого раздела — показать влияние исторических баз данных на формирование теории и понятийного аппарата информационных систем, используемых в области истории.

Наиболее значимо и комплексно исторические базы данных рассмотрены в исследованиях М. Таллера. В его работах, изданных в период с 1980-х по 2000-е годы, зафиксированы возникновение исторических баз данных, их эволюция, состояние и значимость на разных этапах информатизации истории.

Технология баз данных была и продолжает оставаться одной из наиболее востребованных в исторической информатике, цифровой истории и цифровых гуманитарных науках. На ее основе создаются электронные каталоги, цифровые архивы и библиотеки, тематические исторические ресурсы, коллекции машиночитаемых источников, публикаций и др. Созданные на основе технологии баз данных исторические информационные системы позволяют оптимальным образом сохранять, организовывать, визуализировать историческую информацию и обеспечивают необходимый доступ к ней.

Ч. Харви и Д. Пресс в обобщающем труде «Базы данных в исторических исследованиях. Теория, методы и приложения», изданном в 1996 г., говорят о том, что «проектирование и разработка баз данных играют центральную роль в трансформации методов исследований в истории». Авторы отмечают значимость этой технологии для разработки систем хранения первичных и вторичных источников, создания исследовательских ресурсов, связывания записей (*record linkage*), организации, поиска и сортировки исследовательских материалов, анализа и представления исторической информации в целом [Harvey, Press, 1996].

С началом внедрения систем управления базами данных (СУБД) перед историками встал вопрос об использовании программ общего назначения или разработке собственных решений. Данный вопрос решался индивидуально: «некоторые историки экспериментировали

со стандартным программным обеспечением для баз данных: недружелюбными иерархическими СУБД, которые требовали программирования для навигации и поиска данных, или еще не полностью развитыми реляционными базами данных» [Boonstra, Breure, Doorn, 2006, p. 26].

С развитием и распространением реляционных баз данных, в силу своей простоты, адаптивности и универсальности, они быстро стали популярными и у историков [Denley, 1994; Harvey, Press, 1996]. Однако, как было показано в работах М. Таллера и других ученых, реляционная модель имеет серьезные ограничения при структурировании исторических источников. Так, исторические данные плохо вписываются в поля фиксированной длины, которые часто оказываются пустыми из-за отсутствия информации или содержат более одного значения; историки постоянно имеют дело с десятичными системами счисления, изменяющимися топонимами; одно и то же слово может означать географическую принадлежность или просто быть фамилией [Thaller, 1988].

Решением этих проблем стало создание М. Таллером в начале 1980-х годов специализированной СУБД Клио (др.-греч. Κλειω) [Таллер, 1995], которая отличалась поддержкой нечеткости исторической информации и гибкостью форматов данных. Создание специализированной системы, как пишут О. Бонстра, Л. Брере и П. Доорн, «ознаменовало новый период исторического компьютеринга» [Boonstra, Breure, Doorn, 2006, p. 26].

Существенным шагом в осмыслении рассматриваемых проблем являлась дискуссия о подходах к созданию историко-ориентированных баз данных, начатая М. Таллером и П. Доорном [Доорн, 1995; Таллер, 1993]. Ее результатом стало обоснование и утверждение двух основных подходов: *проблемно-ориентированного* и *источнико-ориентированного*, которые вполне можно рассматривать в качестве подходов к созданию историко-ориентированных систем в целом (что будет сделано в главе 2). Помимо указанных были предложены методо-ориентированный и пользовательско-ориентированный подходы, однако они не оказали определяющего влияния на развитие направления.

Постановка вопросов об источнико-ориентированном и проблемно-ориентированном подходах к созданию баз данных для исторических исследований означала видение специфики информационных ресурсов для исторической науки не только в том, что они содержат информацию особого рода — историческую, но и в переносе акцентов в понимании специфики исторических ресурсов вообще и историко-ориентированных баз данных в частности в область организации информации исторических источников и методов их обработки.

В этом смысле показательна позиция В. Леверманна, который через обоснование понятия *«источнико-ориентированная обработка данных»* развивает понятия *«источнико-ориентированный подход»* и *«источнико-ориентированная информационная система»* [Леверманн, 1994]. При этом содержание указанного понятия раскрывается словами М. Таллера о том, что источнико-ориентированная обработка данных, в противоположность так называемой *методо-ориентированной*, — «это стремление смоделировать в компьютере весь объем информации, содержащейся в историческом источнике; это попытка обеспечить возможность использования источника в самых разнообразных целях. Предоставляя историку средства для различных типов анализа, он не требует от него на этапе создания базы данных принятия решения о том, какие методы будут позднее использованы» [Таллер, 1993]. Отмечая достоинства и значение источнико-ориентированной обработки данных, В. Леверманн акцентирует внимание на том, что она ориентирована на максимально точное сохранение источника и предоставление широкого диапазона средств, позволяющих использовать различные аналитические подходы к ним [Леверманн, 1994].

Примечательно, что в предисловии редакции к сборнику научных трудов, в котором была опубликована цитируемая работа В. Леверманна, вводится и такое понятие, как *«специализированное источнико-ориентированное программное обеспечение»* [Круг идей..., 1994, с. 5]. Здесь же сформулирован вопрос о соотношении источнико-ориентированных технологий и информационных систем для обработки исторических данных. Отвечая на него, авторы указали, что проблема источнико-ориентированных технологий не сводится только к соответствующим системам, а связана с идеологией отноше-

ния к информации, в центре которой — сохранение инфраструктуры данных (структурной, семантической), возможность воспроизвести историографическую и культурную среду исследовательского процесса. Приведенные положения обогащают понимание историко-ориентированного подхода и историко-ориентированных информационных систем аспектами, связанными с программным обеспечением и технологиями. Кроме того, в них сформулированы и важнейшие принципы обеспечения специфики исторической информации, создания и эффективности историко-ориентированных информационных систем: сохранение инфраструктуры исторических данных, возможность воспроизведения историографической и культурной среды исторического исследования [Круг идей..., 1994, с. 7, 8].

Можно однозначно констатировать, что базы данных на много лет заняли нишу основного средства организации исторической информации и структурирования исторических источников. Историки и в настоящее время активно используют СУБД общего назначения для решения своих задач. Но нельзя не отметить, что в последнее десятилетие ведущим стандартом представления текстовых источников в гуманитарных науках (особенно в лингвистике и литературоведении) стал стандарт TEI (*Text Encoding Initiative*), позволяющий на основе XML проводить глубокую разметку текстов. Эта технология, с одной стороны, нивелирует недостатки реляционной модели и стандартных СУБД, а с другой — является более простой и доступной по сравнению с Клею.

Разработка стандарта TEI начинается еще в конце 1980-х годов, в 1990 г. выходит первое руководство (*Guidance*)⁴. Другим набирающим популярность средством представления гуманитарных данных становятся технологии NoSQL. Однако TEI очень медленно проникает в инструментарий историков, особенно российских. Несмотря на отмеченные недостатки, реляционные базы данных все еще преобладают. Подробнее технология XML и стандарт TEI как основа для создания исторических информационных систем будут рассмотрены в главе 2.

⁴ TEI: History [Электронный ресурс]. URL: <http://www.tei-c.org/about/history> (дата обращения: 10.04.2020).

1.4. Классификация исторических информационных систем⁵

Исторические информационные системы могут быть классифицированы по общим, специфическим историческим и специальным, свойственным только данному виду ресурсов основаниям [Корниенко, Кирьянов, Гагарина, Власова, 2014].

Классификация историко-ориентированных систем по общим основаниям определяется их принадлежностью к информационным ресурсам и информационным системам вообще. Такие классификации в основании имеют те или иные свойства и признаки последних. Существуют общие классификации информационных систем по архитектуре, по характеру обработки данных, по масштабам охватываемых задач, по сфере применения и др. В связи с тем, что они в достаточной мере рассматриваются в соответствующей научной и учебной литературе, остановимся подробнее на классификациях, связанных с принадлежностью историко-ориентированных информационных систем к специальным, специфическим историческим видам и типам ресурсов.

Так, в отечественной исторической информатике уже в середине 1990-х годов разрабатывались классификации баз данных, в основание которых были положены тип данных и характер источников (нарративные, статистические, формулярные) [Аханчи, 1994; Аханчи, 1995]. Другим важным основанием является разделение информационных систем по моделям, положенным в основу их создания (источнико-, проблемно-, методо-ориентированные и смешанные). Одной из наиболее часто используемых является классификация исторических информационных систем по назначению или характеру решаемых за-

⁵ Результаты, описанные в данном разделе, частично были ранее опубликованы; см.: Гагарина Д.А., Корниенко С.И., Поврозник Н.Г. Информационные системы в цифровой среде исторической науки // Электронный научно-образовательный журнал «История». 2016. Вып. 7 (51) [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: <http://history.jes.su/s207987840001638-0-1> (дата обращения: 10.07.2020).

дач. Например, выделяют просопографические системы и исторические ГИС. В основу классификации могут быть положены и такие характеристики, как хронология, тема, география, характер источников, данные которых они содержат (массовые или единичные источники).

Как уже отмечалось, описание и классификация имеют существенное значение для создания систем документирования и каталогизации историко-ориентированных информационных систем, оценки их места и роли, определения перспектив разработки и возможностей использования. Ниже рассмотрен один из наиболее существенных, на наш взгляд, подходов к классификации и выделены некоторые значимые виды систем.

Классификация по назначению

Классификация информационных систем по назначению или характеру решаемых задач дает представление об их разнообразии, возможностях использования в исследованиях и образовании. Анализ массива систем по данному основанию позволил выделить следующие их типы: геоинформационные, просопографические, фактографические, библиографические, историографические, энциклопедические, экспертные, статистические системы, а также научные журналы (примеры систем этих типов представлены в каталоге, описанном в следующем разделе). Большинство рассмотренных нами систем относятся к фактографическим, геоинформационным, историографическим, библиографическим и энциклопедическим ресурсам.

Развитие информационных технологий приводит к появлению новых типов систем. Так, все большую популярность приобретают системы исторических 3D- и 4D-реконструкций, системы реализации дополненной и виртуальной реальности (симуляторы, виртуальные музеи, экспозиции, экскурсии и пр). При классификации по назначению можно выделить библиотеки источников (*Primary Sources Libraries*), информационно-справочные/информационно-поисковые системы, экспертные системы, системы автоматизации научных исследований и информационно-образовательные системы.

Отдельно можно выделить особую группу *метасистем* (*полисистем*) — крупных проектов, реализуемых при сотрудничестве нескольких организаций (университетов, архивов, библиотек). Метасистемы представлены следующими типами.

1. Коллекции информационных систем и баз данных в рамках университета и других организаций, не связанные тематически. Примеры: Databases of Johns Hopkins Libraries, History⁶; Database Library. University of Wisconsin-Madison⁷.
2. Тематические порталы, содержащие ссылки на несколько тематических информационных систем и публикации по теме. Пример: Primary Sources on the Web: Finding, Evaluating, Using⁸.
3. Агрегаторы, содержащие каталог верхнего уровня, далее подкаталоги, далее ссылки на сторонние сайты с ресурсами (в частности, источниками). Такие системы реализованы на основе облачных технологий, вики-технологий и др. Примеры: EuroDocs: Online Sources for European History⁹, European History Primary Sources (EHPS)¹⁰.

Просопографические системы. Довольно распространенным типом информационных ресурсов являются *просопографические системы*, которые можно отнести одновременно к источник- или проблемно- и методо-ориентированным. В 1990-е годы отмечается всплеск интереса к просопографическим исследованиям, среди

⁶ Databases of Johns Hopkins Libraries, History [Электронный ресурс]. URL: <https://databases.library.jhu.edu/databases/subject/history> (дата обращения: 10.04.2020).

⁷ Database Library. University of Wisconsin-Madison [Электронный ресурс]. URL: <https://search.library.wisc.edu/search/database> (дата обращения: 10.04.2020).

⁸ Primary Sources on the Web: Finding, Evaluating, Using [Электронный ресурс]. URL: <http://www.ala.org/rusa/sections/history/resources/primarysources> (дата обращения: 10.04.2020).

⁹ EuroDocs: Online Sources for European History [Электронный ресурс]. URL: http://eudocs.lib.byu.edu/index.php/main_zage (дата обращения: 10.04.2020).

¹⁰ European History Primary Sources (EHPS) [Электронный ресурс]. URL: <http://primary-sources.eui.eu> (дата обращения: 10.04.2020).

главных причин которого Ю.Ю. Юмашева называет широкое распространение персональных компьютеров и появление доступных СУБД [Иванова, Юмашева, 2005]. Методика просопографического исследования эффективно реализуется с помощью реляционных баз данных и позволяет на их основе организовать информацию и быстро получить результаты. В это время появляется большое количество персональных просопографических баз данных, построенных преимущественно на основе источников XX в. анкетного характера. Во второй половине 1990-х — начале 2000-х годов выходит ряд публикаций Ю.Ю. Юмашевой, И.М. Гарсковой, Д.А. Гутнова, В.А. Перевертня и др. о просопографических базах данных [Гарскова, 1995; Иванова, Юмашева, 2005; Перевертень, 1997]. В этих работах рассматривается не только реализация конкретных проектов, но и методологические вопросы. В настоящее время доля просопографических информационных систем среди прочих видов заметно снизилась, но одновременно появляются крупные просопографические проекты с сотнями тысяч и миллионами записей. Таковыми являются, например, проект «Память народа», собравший информацию о местах первичных захоронений более 5 млн солдат и офицеров Великой Отечественной войны¹¹, и *Lives of the First World War*¹², в котором собрана информация о почти 8 млн военнослужащих Первой мировой войны.

Геоинформационные системы. Первые работы по изучению археологических, региональных, социально-экономических проблем с применением ГИС стали появляться в Северной Америке и Европе 30–35 лет назад. В 2000-е годы наблюдается рост интереса к методам и технологиям пространственного анализа и визуализации данных в исторических исследованиях и, соответственно, числа *исторических геоинформационных систем*. В книге «Past, Present and Future of Historical Information Science» методы визуализации и картографического моделирования выделяются как новые и перспективные в обла-

¹¹ Память народа [Электронный ресурс]. URL: <https://pamyat-naroda.ru> (дата обращения: 10.04.2020).

¹² *Lives of the First World War* [Электронный ресурс]. URL: <https://livesofthefirstworldwar.org> (дата обращения: 10.04.2020).

сти научного знания вообще и в исторической информатике в частности [Boonstra, Breure, Doorn, 2006]. В первой половине 2000-х годов выходит несколько обобщающих работ, среди которых «Past Time, Past Place: GIS for History» А.К. Ноулс [Knowles, 2002], «A Place in History: A Guide to Using GIS in Historical Research» Я.Н. Грегори [Gregory, 2003], а также «Историческая геоинформатика: геоинформационные системы в исторических исследованиях» В.Н. Владимиров [Владимиров, 2005].

Достигнутые на сегодняшний день теоретические и прикладные результаты по созданию и использованию геоинформационных систем в исторической и других гуманитарных науках, количество и уровень научных публикаций по этой проблематике позволяют говорить о таких сложившихся направлениях и дисциплинах, как историческая геоинформатика и пространственная история [Владимиров, 2005; Владимиров, Рыгалова, 2014; Bodenhamer, Corrigan, Harris, 2010; Knowles, 2002; Knowles, 2008; Gregory, 2003; Gregory, 2007; Gregory, 2014]. Их общие свойства и характеристики прежде всего связаны с определением понятия и особенностями геоинформационных систем как вида информационных ресурсов и технологий. Однако существует большое количество таких определений [Бугаевский, Цветков, 2002; Шайтура, 1998], что затрудняет однозначное понимание специфики, технологических и информационных особенностей ГИС, в том числе относящихся к области истории и гуманитарных наук. Мы будем ориентироваться на следующее определение: *геоинформационные системы* — это «информационные системы, обеспечивающие сбор, хранение, обработку, доступ, отображение и распространение пространственно-координированных (пространственных) данных» [Баранов, Берлянт, Капралов, Кошкарёв, Серапинос, Филиппов, 1999].

Исторические энциклопедии. Одним из важных типов историко-ориентированных ресурсов являются *исторические энциклопедии*. Историческая интернет-энциклопедия — это энциклопедия, действующая в веб-среде и содержащая массивы исторических фактов и их интерпретаций. Можно выделить два подтипа таких

ресурсов: те, авторство статей в которых принадлежит коллективу разработчиков, и те, содержание которых создается пользователями (вики-энциклопедии) [Яценко, Яценко, 2015].

Анализ 550 интернет-ресурсов и 50 мобильных приложений, отобранных с точки зрения информативности (рассматривались ресурсы, содержащие от 5000 до 20 000 статей объемом от 20 до 150 слов), визуализации, функциональных возможностей и доступности, проведенный Ю.В. Яценко и А.В. Яценко, показал, что существуют разные по структуре исторические энциклопедии. Часто встречается иерархический вид с несколькими уровнями, когда каждый последующий уровень является подразделом предыдущего информационного блока, а основной страницей системы является стартовая (главная) страница сайта. Другой вариант — перекрестные ссылки. Статьи могут быть упорядочены посредством алфавитного и тематического указателей. При этом преобладает смешанный (гибридный) вид, когда есть и корневая страница, и перекрестные переходы между статьями, организованные на основе гиперссылок. Кроме того, используются электронные версии печатных изданий. В этом случае поисковая система внутри интернет-энциклопедии может отсутствовать [Яценко, Яценко, 2015].

Большая доля энциклопедических ресурсов представляет информацию в виде текста, без сопровождения визуальными элементами [Платонов, 2012]. Количество мобильных приложений невелико, зачастую они содержат недостоверную информацию. В целом для исторических интернет-энциклопедий характерен устоявшийся набор отличительных атрибутов: таймлайн; развитая поисковая система с возможностью фильтрации и сортировки данных; тематические и алфавитные указатели; перекрестные ссылки, которые часто являются дополнением к основной поисковой системе. Такие веб-проекты часто оформлены в формате портала, они содержат расширения и интерактивные сервисы (например, элементы ГИС, инструментарий для работы с изображениями, в том числе их обработки). В Рунете ожидаемо преобладают энциклопедические ресурсы по истории России.

Классификация по областям исторического знания

О содержательных характеристиках исторических информационных систем позволяют судить такие атрибуты, как область истории, география и хронология. Классификация по ним показывает тематическое разнообразие изучаемого вида ресурсов.

По-прежнему, как было установлено П. Доорном еще в середине 1990-х годов [Доорн, 1995], наиболее популярными тематиками (в порядке убывания) являются: социальная история, экономическая история, история культуры, политическая история, демография, краеведение и археология. Тот факт, что большая часть современных систем предназначены для исследований по социально-экономической и политической истории, объясняется, с одной стороны, востребованностью данной проблематики, а с другой — наличием значительных массивов информации, требующей обработки. В то же время растет число ресурсов, содержащих информацию, связанную с историко-культурным наследием.

По хронологическим критериям большинство рассматриваемых информационных систем относятся к истории XX в., что связано с доступностью источников этого периода. Ранее, по данным на начало 1990-х годов, отмечалось преобладание ресурсов по истории XIX в., а затем XVII и XVIII вв. [Доорн, 1995].

В целом исторические информационные системы имеют разную временную протяженность — от нескольких месяцев до нескольких веков. Активно развиваются системы, охватывающие до нескольких тысячелетий, в том числе в рамках междисциплинарного направления *Big History* — истории развития Земли от Большого взрыва до современности. Для таких проектов характерно использование таймлайнов и картографических инструментов (например, как в интерактивном атласе TimeMaps¹³, содержащем более 1500 карт по различным периодам и регионам). Значительная хронологическая протяженность

¹³ TimeMaps [Электронный ресурс]. URL: <http://www.timemaps.com> (дата обращения: 10.04.2020).

присуща и более традиционным коллекциям источников, к которым относится проект Йельского университета Avalon¹⁴, разработанный в 2008 г. и содержащий полные тексты источников в области права, экономики, политики, дипломатии и государственного управления, начиная с 4 тыс. лет до н.э. Наиболее известным в области Big History является одноименный Big History Project¹⁵, разрабатываемый по инициативе Билла Гейтса и рассматривающий почти 14 млн лет истории. Еще один проект Big History — ChronoZoom — совместная разработка Калифорнийского университета в Беркли и Московского государственного университета им. М.В. Ломоносова¹⁶. Данный проект представляет инструменты для создания собственных ресурсов, их сохранения и репрезентации в рамках единой системы. Для перечисленных примеров и других подобных проектов свойственны сильная образовательная составляющая и наличие методических разработок, адаптирующих использование продукта в учебном процессе.

Классификация по типу исторических источников

Существующие историко-ориентированные информационные системы варьируются по типам исторических источников, использованных в формировании контента. В этом смысле можно говорить о системах, основанных на письменных, статистических, вещественных, изобразительных, аудио-, видео-, мультимедийных, картографических и других источниках, а также на их различных комбинациях.

Наиболее часто встречающиеся источники на всем протяжении развития историко-ориентированных информационных систем и баз данных — письменные, что связано как с их преобладанием в общем объеме исторических источников, так и с доступностью технологий

¹⁴ Avalon [Электронный ресурс]. URL: <http://avalon.law.yale.edu> (дата обращения: 10.04.2020).

¹⁵ Big History Project [Электронный ресурс]. URL: <https://school.bighistoryproject.com> (дата обращения: 10.04.2020).

¹⁶ ChronoZoom [Электронный ресурс]. URL: <http://www.chronozoom.com> (дата обращения: 10.04.2020).

для их перевода в цифровой формат. Вместе с тем характер используемых письменных источников и способы их репрезентации в системе существенно изменялись со временем. Историко-ориентированные базы данных конца 1980-х — 1990-х годов основывались преимущественно на массовых, прежде всего формулярных источниках: личных карточках, отчетных делопроизводственных документах, родословных книгах, формулярных списках, информация которых хорошо структурирована и удобным образом укладывается в формат реляционных баз данных.

Поскольку хранение, анализ и описание изображений в оболочке информационной системы гораздо легче текста, в ряде случаев текстовый источник заносится в информационную систему в виде нераспознанного изображения. Такие изображения могут быть сделаны на основе оригинала или копии (факсимильного издания) и представлены следующими форматами: jpeg, gif, pdf, djvu и др. Все чаще встречаются псевдоизображения, представленные в формате pdf, передающие, например, оригинальный шрифт источника, а также нераспознанный или плохо распознанный текст.

С развитием технологий круг источников, для представления информации которых используются информационные системы, расширяется. Одновременно меняется способ репрезентации источников в системе в направлении большей машиночитаемости, а также точности и детальности визуализации. В наибольшей степени это видно на примере письменных и вещественных источников. Применительно к текстовым источникам это может быть обозначено как переход от простой атрибуции к полнотекстовости с сохранением формальных и внешних особенностей документа и возможностью передачи многослойности текста. Сегодня все более характерным при репрезентации текстовых источников в информационной системе становится их представление в виде изображения, метаописания, распознанного и размеченного текста и, если необходимо, транскрипции, перевода¹⁷.

¹⁷ Например, Древнерусские берестяные грамоты [Электронный ресурс]. URL: <http://gramoty.ru> (дата обращения: 10.04.2020).

В перспективе следует ожидать изменения способов репрезентации и для других источников, в частности аудио-, видео- и мультимедийных данных, для которых сегодня по-прежнему преобладают атрибутивное описание и наличие соответствующего файла, не предполагающие возможности полнотекстового поиска. Значительным шагом в этом направлении стала разработка инструментария для анализа массивов аудиоданных. Так, в Королевском технологическом институте (Стокгольм, Швеция) разработана программа¹⁸ с открытым кодом для поиска по большим неструктурированным массивам аудиоданных. В ходе тестирования указанного инструментария на оцифрованных аудиоматериалах из архива Института языка и фольклора (Institute for Language and Folklore, ISOF) — а это порядка 13 тыс. часов звучания — была доказана эффективность разработанной методики анализа [Fallgren, Malisz, Edlund, 2018].

Также встречается значительное число информационных систем, источниковой базой которых стали статистические данные. Как правило, такие системы предназначены для анализа социально-экономических процессов и динамики демографического развития.

Наличие многочисленных систем, различающихся по такому критерию, как тип источника, создает возможности для комплексных исследований на основе разнотипных источников, сочетая текст, статистику и изображения.

1.5. Каталог историко-ориентированных информационных систем

Сайт «Историко-ориентированные информационные системы» — <http://digitalhistory.ru> — предназначен для отображения процесса и результатов исследования теоретических и практических вопросов создания таких систем, использования их в науке, образовании и для сохранения историко-культурного наследия.

¹⁸ SoX — Sound eXchange [Электронный ресурс]. URL: <http://sox.sourceforge.net> (дата обращения: 10.04.2020).

Kornienko, Sergey I., Gagarina, Dinara A., Povroznik, Nadezhda G.
Historical Information Systems: Theory and Practice [Text] / S. I. Kornienko, D. A. Gagarina, N. G. Povroznik ; National Research University Higher School of Economics. — Moscow : HSE Publishing House, 2021. — 232 pp. — 600 copies. — ISBN 978-5-7598-2321-6 (pbk.). — ISBN 978-5-7598-2426-8 (e-book).

Historical, or history-oriented, information systems are a significant element of the information environment in the humanities. Its allocation is associated with the development of historical informatics and a history-oriented approach, the formation of an information environment, as well as the practice of creating historical resources.

The book contains the results of a study of theoretical and applied problems of the creation and implementation of history-oriented information systems. This is the first comprehensive study on this topic. Some problems in the book are considered for the first time, others were not specially studied although they were previously reflected in the literature.

The publication is addressed to historians, specialists in the field of digital history and digital humanities, as well as developers of digital resources containing historical content or focused on use in historical research and education.

Научное издание

Корниенко Сергей Иванович
Гагарина Динара Амировна
Поврозник Надежда Георгиевна

**Исторические информационные системы:
теория и практика**

Зав. книжной редакцией *Е.А. Бережнова*
Редактор *В.Е. Красавцева*
Компьютерная верстка: *С.В. Родионова*
Корректор *О.В. Кириллова*
Дизайн обложки: *И.В. Ветров*

Все новости издательства — <http://id.hse.ru>

По вопросам закупки книг
обращайтесь в отдел реализации
Тел.: +7 495 772-95-90 доб. 15295, 15297
bookmarket@hse.ru

Национальный исследовательский университет
«Высшая школа экономики»
101000, Москва, ул. Мясницкая, 20
Тел.: +7 495 772-95-90 доб. 15285

Подписано в печать 22.11.2021. Формат 60×88/16
Гарнитура Minion. Усл. печ. л. 14,1. Уч.-изд. л. 10,4
Печать офсетная. Тираж 600 экз. Изд. № 2441. Заказ №

Отпечатано в АО «ИПК «Чувашия»
428019, г. Чебоксары, пр. И. Яковлева, 13
Тел.: +7 8352 56-00-23