

Л. И. БОРОДКИН

ИСТОРИЧЕСКАЯ ИНФОРМАТИКА В КОНТЕКСТЕ ТЕХНОЛОГИЙ E-HUMANITIES

Историческая информатика, развиваясь в русле современных тенденций информационного общества, формирует новые приоритеты. Первое десятилетие XXI в. принесло немало инноваций в информационную структуру науки и образования. Информационно-коммуникационные технологии (ИКТ) играют все более активную роль и в формировании новой инфраструктуры современных исторических исследований, университетского образования. Приоритетная тематика включает теперь 3D-реконструкции, исторические ГИС-приложения, информационную архитектуру ¹, методологию e-Science, новые Web-сервисы, применения гридтехнологий в задачах разработки распределен-

¹ Информационная архитектура — совокупность методов и приемов структурирования информации (знаний или данных). Чаще всего термин «информационная архитектура» применяется в веб-разработке; в этом контексте информационная архитектура связывается прежде всего с принципами систематизации информации и навигации по ней с целью помочь пользователям более успешно находить и обрабатывать нужные данные.

Бародкін Леанід Іосіфавіч — загадчык кафедры гістарычнай інфарматыкі Маскоўскага дзяржаўнага ўніверсітэта імя М. В. Ламаносава, доктар гістарычных навук, прафесар.

ных электронных ресурсов (e-Resources), внедрение цифровых медиа и дистанционных технологий в практику исторического образования.

Появление здесь ряда новых для гуманитарной аудитории терминов и понятий требует комментария: ведь они в ближайшие годы войдут в активный лексикон исторической информатики, а затем — и в общенаучный лексикон историков. Этот процесс уже идет, его инициировали известные голландские специалисты по исторической информатике, авторы программной работы о прошлом, настоящем и будущем этой развивающейся междисциплинарной области — Онно Боонстра, Питер Доорн, Лен Брёре 2.

* * *

Наиболее характерной чертой развития ИКТ в 2000-х гг. является формирование новых компонентов электронной среды взаимодействия. Этот процесс породил новые методологические подходы, программные и аппаратные решения, новые термины и понятия. Широкое распространение получили новые категории: e-Science, e-Social Science, e-Humanities, e-data, e-Research, e-Publications, e-Culture, e-Learning, e-Business, e-Banking, e-Medicine, e-Government и т. д., связанные с развитием распределенной инфраструктуры информационных технологий и ресурсов. Наверное, излишним будет напоминание, что «е» означает здесь «electronic». Нас интересует прежде всего содержание термина e-Science, возникшего изначально в Великобритании, где крупные исследовательские проекты в этой области начались в 2001 г. Именно там было дано первое определение e-Science, получившее в дальнейшем широкое распространение: «научно-технологическая область, в которой всевозрастающую роль играет распределенное глобальное взаимодействие посредством сети интернет, с использованием очень больших коллекций данных, компьютерных ресурсов тера-уровня и высококачественной визуализации, доступных индивидуальному пользователю».

2 Boonstra O., Breure L. and Doorn P. Past Present and Future of Historical Information Science. Amsterdam : NIWI-KNAW, 2004. См. об этой книге : Бородкин Л. И. Историческая информатика в точке бифуркации : движение к Historical Information Science // Круг идей : алгоритмы и технологии исторической информатики. Москва; Барнаул, 2005; Гарско ва И. М. Историческая информатика и квантитативная история: преемственность и взаи-модействие // Анализ и моделирование социально-исторических процессов. М. : КомКнига, 2006.

лю» 3. Надо сказать, что в русском языке термин e-Science существует пока преимущественно в англоязычном варианте 4.

Иногда возникает вопрос: в чем отличие концепции e-Science от хорошо известной нам глобальной сети интернет? Ведь пользователь интернета имеет доступ со своего компьютера к разнообразным ресурсам глобальной сети, размещенным на сайтах всех стран мира; при этом набор веб-сервисов непрерывно расширяется. Дело в том, что кроме «обычной» информации, размещенной в интернете, специалисты, работающие в сфере науки и образования, нуждаются в доступе к крупномасштабным информационным массивам, базам данных, имеющим объемы памяти, измеряемые терабайтами. Работа с такими массивами требует вычислительных мощностей с производительностью уровня терафлоп. Обычный пользователь интернета работать в этом режиме не сможет. Сегодня потребность в таких крупномасштабных ресурсах и компьютерных мощностях испытывают в основном ученые-естественники, но в ближайшем будущем к этому рубежу подойдут и разработчики проектов в различных областях социально-гуманитарного знания 5. Задача e-Science, таким образом, — создание организационных и технологических структур, разработка соответствующего программного обеспечения для функционирования новой информационной среды с распределенными ресурсами (информационными и вычислительными), обеспечивающими доступ к ним индивидуальных пользователей, исследовательских групп, лабораторий и институтов (эти сообщества называют виртуальными организациями, *collaboratories*).

Основное русло реализации задач e-Science прокладывают грид-технологии. Эта концепция (нередко ее называют *Grid Computing* — распределенные сети, или «решетки» вычислительных ресурсов) соответствует одному из ведущих и перспективных направлений развития ИКТ. В последние годы начата разработка средств поддержки Grid на уровне стандартов и продуктов. Ситуацию можно сравнить с состоянием сети WWW в конце 1990-х гг., при этом специалисты предсказывают, что Grid произведет такую же революцию в области обработки данных, какую интернет произвел в сфере компьютерных коммуникаций. Так, корпо-

3 См., например : <http://www.escience.am>, <http://www.nesc.ac.uk>

4 В отличие, например, от термина e-Government, который получил распространение в русском языке как «электронное правительство».

5 Уже сейчас историками созданы огромные базы данных, например в области исторической демографии. Эти базы содержат миллионы (а в отдельных случаях — десятки миллионов) записей, основанных на первичных материалах переписей населения.

рация Oracle развивает свои продукты в направлении более широкого применения Grid-технологий. Grid-технологии используют новую архитектуру программного обеспечения, предназначенную для объединения большого количества недорогих модулей памяти и серверов, создающих распределенный виртуальный ресурс, транспарентный и легко доступный для пользователя (независимо от места нахождения его компьютера). Одна из наиболее сильных сторон Grid-технологий — новые возможности эффективного распространения информации и обучение с доступом к обширным ресурсам 6. Благодаря все более широкому применению современных ИКТ в образовании, доступ к гридинфраструктуре получают школы и вузы, исследовательские группы и широкий круг пользователей различных приложений, что приведет к качественным изменениям в учебном процессе, научных исследованиях и в конечном итоге в разных видах экономической деятельности. При этом специфика термина e-Social Sciences заключается не только в предметной области используемых распределенных информационных ресурсов, но и в акценте на взаимодействие ученых-обществоведов со специалистами по ИКТ 7.

В этой связи представляет интерес деятельность Национального центра по развитию e-социальных наук (NCeSS) 8, основанного пять лет назад Британским Советом по экономическим и социальным исследованиям (ESRC). Задача этого Центра — анализ особенностей использования гридтехнологий в

социальных науках, разработка средств их поддержки, подготовка рекомендаций по стратегии развития е-социальных наук. NCESS предоставляет также консультационные услуги, обучение и доступ к онлайн ресурсам — для помощи исследователям из различных областей социального знания в их адаптации к возможностям и реалиям е-социальных наук. NCESS включает координационный хаб (центр) в Манчестерском университете, семь исследовательских групп и 12 небольших исследовательских проектов, ориентиро-

6 Здесь возникает известная аналогия с единой энергетической системой, включающей множество производителей энергоресурсов, поставляющих электроэнергию в общую сеть, к которой подключены миллионы потребителей («пользователей»), получающих доступ к сети, когда они в этом нуждаются (и не интересующихся, откуда конкретно пришла к ним энергия).

7 О подходах к созданию структур новой информационной среды в российской науке см.: Паринов С. И. E-Science — онлайн-будущее науки. Приложение к журналу «Информационные технологии». № 9. 2007.

8 Национальный центр по развитию е-социальных наук [Электронный ресурс]. — Режим доступа : <http://www.ncess.ac.uk/>

ванных как на количественные, так и на качественные подходы тех или иных социальных наук. Так, один из этих проектов (Collaboratory for Quantitative e-Social Science — CQeSS), действующий в Ланкастерском университете, имеет целью содействие квантитативным исследованиям в различных областях социальных наук, включая развитие исследовательского аналитического инструментария, необходимого для обработки больших массивов количественных данных и доступного в режиме «распределенного использования». Речь идет о таких гридразработках, как SABRE, GROWL, SAKAI 9.

Что касается гуманитарных наук, то перспективы использования грид-технологий связываются здесь во многом с расширяющимися возможностями онлайн-доступа к огромным массивам оцифрованных книг, журналов, изобразительных материалов и т. д. Эффективность использования этих ресурсов резко возрастает, когда исследователь может связывать не только различные тексты, но и тексты с изобразительными источниками, историческими картами, записями интервью, видеоматериалами и т. д., размещенными на различных сайтах. Сегодня в Великобритании осуществляется поддержка четырех проектов по развитию виртуальной среды исследований в области гуманитарных наук, девяти регулярных научных семинаров и семи исследовательских проектов в конкретных областях гуманитарного знания (археология, музейное дело, классические исследования, история Византии) 10. В этих проектах используются различные технологии e-Science, включая обработку изображений, текста древних манускриптов, интеграцию национальных баз данных по материалам археологических экспедиций, компьютерное моделирование средневековых битв, использование 3-D сканирования для анализа поверхности объектов музейного хранения, ГИС. Существенно, что работа над этими проектами идет в тесном сотрудничестве гуманитариев и специалистов из области ИКТ. Это междисциплинарное взаимодействие необходимо еще и потому, что массивы данных, анализируемые гуманитариями (в частности, историками), имеют более сложную, «размытую» («fuzzy») структуру в сравнении с данными, полученными в естественнонаучных исследованиях.

9 Quantitative Research — Национальный центр по развитию е-социальных наук [Электронный ресурс]. — Режим доступа : <http://www.ncess.ac.uk/research/quantitative/>

10 Грантовая поддержка этих проектов проводится усилиями Arts and Humanities Research Council (AHRC) и UK e-Science Core Programme.

Подобные проекты ведутся не только в Великобритании 11. П. Доорн и П. Алькховен отмечают заметную роль голландской программы CATCH, развивающей возможности доступа к оцифрованным объектам культурного наследия в рамках активного сотрудничества между ИТ-специалистами, институтами культурного наследия и учеными-гуманитариями. Участники десяти проектов этой программы разрабатывают соответствующее программное обеспечение, проводят эксперименты, используя грид-технологии хранения данных в сотрудничестве с голландским Big Grid Project. Для усиления возможностей доступа к оцифрованным данным гуманитарного профиля Королевская Академия наук и искусств Голландии (KNAW) и Голландский научный фонд (NWO) создали новый институт — DANS (Data Archiving and Networked Services). DANS функционирует как сетевая организация, с центром, ответственным за организацию инфраструктуры данных 12. В Германии развивается так называемая D-Grid инициатива, в рамках которой создается TextGrid программа, ориентированная на совместную деятельность сетевого сообщества ученых, проводящих редактирование, аннотирование, анализ и публикацию текстов. Дальнейшая цель программы — разработка технологий семантических сетей (Semantic Grid) 13.

На европейском уровне подобные институты и программы работают вместе под эгидой DARIAH (Digital Research Infrastructure for the Arts and Humanities). Цель этого Центра — обеспечить координацию усилий по сохранению культурного наследия в Европе и обеспечению доступа к материалам гуманитарных исследований 14.

Эти разработки дают основания для введения термина e-Humanities (менее принятого пока, чем e-Science)

15. В работе Й. Кирца рассмотрены десять направлений разработок, развитие которых приведет к повышению роли e-Humanities в развитии гуманитарных наук 16:

11 Alkhoven P., Doorn P. New Research Perspectives for the Humanities // International Journal of Humanities and Arts Computing. 2007. Vol. 1, № 1. P. 35—47.

12 Ibid. P. 45. Отметим, что директором-организатором DANS является д-р П. Доорн, Президент АНС.

13 Ibid.

14 Ibid.

15 Ibid. P. 35.

16 Kitz J. E-based Humanities and E-Humanities on a SURF platform. (A report commissioned by SURF-DARE), KRA publishing research, 1 June 2004, Amsterdam.

- 1) создание больших коллекций оцифрованных материалов (таких, например, как корпус источников);
- 2) смена акцентов в образовательных программах, ориентированных на овладение навыками работы со стандартными пакетами: предлагается изучать компьютерные методы и технологии, связанные с решением исследовательских задач, характерных для соответствующей области гуманитарного знания;
- 3) helpdesks: их роль на гуманитарных факультетах должна измениться, они должны содержать больше возможностей для работы студента-гуманитария в новой исследовательской информационной среде, обеспечивать использование специализированного программного обеспечения;
- 4) установление более тесных связей гуманитарной информатики (humanities computing) с «базовой» информатикой;
- 5) обеспечение более широкого доступа к массивам электронных данных; они должны быть приложенными к соответствующей публикации; архивы машиночитаемых данных должны принимать их на хранение; необходимо разработать стандарты метаданных для упорядочения электронных библиотек, архивов и других ресурсов;
- 6) создание библиотек программного обеспечения, прошедшего апробацию в исследовательских проектах; они должны быть хорошо документированными;
- 7) развитие веб-сервисов, которые повысят возможности использования данных и ресурсов в различных исследованиях;
- 8) разработка схем метаданных становится критическим аспектом развития e-Humanities в ближайшие годы;
- 9) разработка основанных на онтологиях авторских программных средств (а также программного обеспечения, созданного в коллективных проектах) — важная составляющая процесса; структура метаданных при этом должна быть согласована с соответствующей структурой, принятой в архивах машиночитаемых данных;
- 10) развитие удобного интерфейса и средств визуализации — для повышения эффективности исследований мультимедийных данных, демографических и социально-экономических рядов, материалов археологических экспедиций и т. д.

Как отмечают П. Доорн и П. Алькховен, наиболее высокие оценки экспертов получили пп. 1), 2) и 4), что подчеркивает приоритетность создания коллекций оцифрованных материалов, трансформации об

разовательных программ в области ИКТ-приложений и более тесной связи «отраслевой» гуманитарной информатики с «базовой» информатикой 17. Профессор Ш. Росс, директор Центра гуманитарной информатики Института информации Университета Глазго также рассматривает в качестве важнейшей задачи развития e-Humanities в ближайшем десятилетии — обеспечение хранения оцифрованных данных и онлайн-доступа к ним 18.

* * *

Заметной тенденцией является существенное обновление веб-сервисов, которые в последние годы получили новый импульс к развитию в связи с формированием концепции Web 2.0. Появление термина Web 2.0 принято связывать со статьей Тима О'Рейли «What Is Web 2.0» от 30 сентября 2005 г.¹⁹, впервые опубликованной на русском языке в журнале «Компьютерра» (№37 (609) и 38 (610) от 14 и 19 октября 2005 г. соответственно) и затем выложенной под заголовком «Что такое Веб 2.0» на веб-сайте «Компьютерра online»²⁰. В этой статье Т. О'Рейли связал появление большого числа сайтов, объединенных некоторыми общими принципами, с общей тенденцией развития интернет-сообщества и назвал это явление Web 2.0 в противовес «старому» Web 1.0. Несмотря на то что значение этого термина до сих пор является предметом многочисленных споров, те исследователи, которые признают существование Web 2.0, подчеркивают фактор широкого участия «обычных» пользователей глобальной сети в формировании и развитии коллективных интернет-ресурсов и выделяют несколько основных аспектов и технологий этого явления: Wiki — технология коллективного создания гипертекста; Flickr — социальный фотосервис; YouTube — социальный видеосервис; Google Maps — карты и спутниковые снимки Земли; Blogs — сетевые дневники; Del.icio.us — онлайн закладки; Netvides — сетевой десктоп и др.

17 Alkhoven P., Doorn P. Op. cit. P. 38—39.

18 Ibid. P. 42.

19 <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

20 <http://www.computerra.ru/think/234100/>

* * *

Важное место в тематике зарубежных ассоциаций АНС в последние годы занимают вопросы разработки научно-образовательных ресурсов, использования новых ИКТ в обучении студентов-историков (включая технологии дистанционного обучения, Веб 2.0 и др.). Так, этой тематике была посвящена очередная ежегодная конференция американской ветви АНС (2008). Конференция имела характерное название: Web 2.0 / History 2.0: Making History Together. В центре внимания участников были следующие вопросы:

- Что включает «исторический сегмент» Web 2.0?
 - В чем его отличия от соответствующего сегмента Web 1.0?
 - Что заставляет нас выходить за пределы возможностей Web 1.0?
-
- Каковы приложения Web 2.0 в исторических исследованиях и образовании?
 - Каковы «плюсы» и «минусы» Web 2.0 в этих приложениях?

Рассмотрение данных вопросов потребовало обращения к конкретным аспектам сетевого взаимодействия, указанным организаторами конференции: программы «Open source», используемые для интерпретации истории; тэги в исторических приложениях; технологии Wiki, блоги и RSS в исторических приложениях; право собственности на данные; базы данных и исторический анализ; картографирование истории (документация, основанная на возможностях Google); создание звуковых и видеопередач (подкастов), цифровое видео (например, YouTube), виртуальные миры в исторических приложениях.

Представленная тематика активно обсуждалась на ежегодной британской конференции АНС—УК в ноябре 2007 г., где акцент был сделан на использовании ИКТ в образовании студентов-историков всех трех

уровней (бакалавриат, магистратура, PhD). На шести секциях этой конференции обсуждались следующие вопросы: 1) компьютерные технологии в обучении аспирантов-историков; 2) веб-ресурсы национальных архивов: поиск в онлайн-каталогах; 3) креативное использование ИКТ в обучении историков — бакалавров и магистров; 4) критика источников: «цифровая дипломатика» в учебном плане историков; 5) аккумуляция и дифференциация профессиональных знаний и навыков при переходе с уровня бакалавра на уровень магистра: роль ИКТ; 6) современное содержание специализации по профилю «History and Computing» в британских университетах.

Одна из задач, стоящих перед специалистами по исторической информатике, работающими в университетах, — включить рассмотрение отмеченных выше тенденций в учебном плане специализации по профилю «Историческая информатика».

* * *

Как уже отмечалось выше, одно из актуальных направлений исторической информатики, активно развивающихся в последние годы, связано с виртуальной реконструкцией объектов историко-культурного наследия, основанной на технологиях компьютерного 3D-моделирования. Эта проблематика была в центре внимания XII международной конференции ассоциации «История и компьютер» (Москва, 2010 г.). Представленные на конференции доклады содержали подробный анализ накопленного историками разных стран опыта виртуальной реконструкции архитектурных памятников различных эпох и культур. Наибольший интерес вызвали демонстрации 3D-реконструкции пространственной структуры монастырских комплексов Москвы и Енисейска, а также памятников античной архитектуры Херсонеса. Очевидно, происходит переход от научно-популярной, «экскурсионно-туристической» стадии развития разработок реконструкции объектов историкокультурного наследия к академичной, всесторонне аргументированной реконструкции, основанной на тонких источниковедческих методиках синтеза различных документов, характеризующих эволюцию реконструируемых объектов.

Завершая предпринятый в данной работе обзор новых тенденций развития исторической информатики, можно заметить, что содержание отмеченной нами ранее «точки бифуркации» в эволюции этой междисциплинарной области 21 получило более конкретную характеристику.

Артукул паступіў у рэдакцыю 14 красавіка 2010 г.

21 Бородин Л. И. Историческая информатика в точке бифуркации : движение к Historical Information Science // Круг идей : алгоритмы и технологии исторической информатики. М.; Барнаул, 2005.