

Веб-архивы в исторических исследованиях: опыт онлайн истории органов исполнительной власти Томской области

Ж.А. Рожнева. Национальный исследовательский Томский государственный университет, Томск

Ключевые слова:

Интернет, веб-архив, онлайн-история, исторический анализ.

В настоящее время Интернет играет большую роль в жизни людей и общества во многих странах, в том числе и в России. Обширные массивы сетевых ресурсов, в первую очередь веб-сайты, отражают различные политические, экономические, социальные, культурные события и процессы, происходящие в обществе. Поскольку Всемирная паутина имеет уже более чем двадцатилетнюю историю, веб-материалы могут рассматриваться в качестве потенциальных исторических источников. Об интересе к их изучению в историческом аспекте свидетельствует появление таких новых исследовательских направлений, как веб история (web history), веб-сайт-история (website history), онлайн-история организаций (history of organization online).¹

Как известно, глобальная сеть с одной стороны обладает избыточными массивами информационных ресурсов, с другой – значительная их часть существует в неизменном состоянии относительно недолгий период времени. Являясь частью культурно-исторического наследия, веб-материалы, безусловно, нуждаются в сохранении. Первый проект по созданию веб-архива начал реализовываться в 1996 г. некоммерческой организацией «Архив Интернета», основанной американским предпринимателем Брюстером Кейлом. В данном архиве хранятся копии веб-страниц со всего мира. Посредством специального архивного сервиса, получившего название Wayback Machine (Машина времени), пользователи могут получить свободный доступ к 485 млрд веб-страниц.² С конца 1990-х гг. начали создаваться национальные, региональные и тематические веб-архивы. По данным исследовательской группы Португальского веб-архива, на сегодняшний день в мире насчитывается 76 инициатив по веб-архивированию.³ Основной целью создаваемых архивов является обеспечение сохранности веб пространства как части современной цивилизации. Таким образом, для нынешнего и будущих

поколений веб-архивы имеют очевидную ценность. Однако их широкое использование в гуманитарных исследованиях только начинается.

В России нет программы по созданию национального веб-архива. Российские архивы и библиотеки, за редким исключением, не сохраняют веб-материалы как часть историко-культурного наследия. Для отечественных историков доступна только живая сеть или российские веб-сайты, заархивированные в рамках глобальных проектов, таких как Архив Интернета (Internet Archive),⁴ который пока не является привычной и понятной научно-исследовательской средой.

В то же время очевидно, что изучение современного периода новейшей истории России политической истории будет неполным без использования подобных материалов, в том числе, распространение Интернета и его применение в сфере государственного управления, в коммуникации между гражданами и государством началось более пятнадцати лет назад. Сегодня в Сети размещаются веб-сайты органов государственной власти всех уровней – от федеральных до местных. Свои сайты имеют и другие акторы политического процесса – партии, общественные движения, неправительственные организации, отдельные политики. Очевидно, что современный историк должен уметь использовать подобные источники в своих исследованиях. При этом его профессиональные навыки могут подвергнуться определенному испытанию, поскольку включение архивных и «живых» веб-сайтов в научный оборот в качестве первичных источников требует выработки новых методологических подходов, учитывающих их отличие от традиционных исторических источников.

В данном исследовании представлен опыт онлайн-истории главных органов исполнительной власти Томской области. Высшим исполнительным органом государственной власти региона является Администрация Томской области, во главе которой стоит губернатор. Веб-сайты томских органов исполнительной власти появились в Интернете в конце 1990-х – начале 2000-х годов. На основе их анализа можно выявить проблемы, возникающие в историческом исследовании, если оно основывается на веб-материалах как главных источниках. На данном этапе исследования меня интересовали следующие вопросы:

- Как изменялись рассматриваемые веб-сайты с течением времени?
- Какие возможности они дают для реконструкции истории органов власти?
- Какие сложности возникают при работе с подобными историческими источниками?

Исследование базируется в первую очередь на заархивированных веб-сайтах в Архиве Интернета и на веб-сайтах в «живой сети». В качестве дополнительных источников использовались российские и региональные нормативные документы, материалы независимого мониторинга информационной открытости веб-сайтов государственных органов, а также интервью с бывшими и нынешними сотрудниками органов власти региона, ответственными за функционирование сайтов.

У веб-сайта Администрации Томской области были выявлены четыре последовательные версии, отличающиеся дизайном и структурой. Две последних версии доступны в «живой сети», в то время как первая и вторая существуют только в архивной форме. Действующий веб-сайт Администрации Томской области имеет доменное имя tomsk.gov.ru, которое является его идентификатором в «живой сети» и в Архиве Интернета с момента создания. Предыдущая современной версия сайта остается доступной в «живой сети» под доменным именем old.tomsk.gov.ru. Ранее по этому адресу располагались первая и вторая версии сайта, после своей замены. Кроме того, в Архиве Интернета имеется небольшое количество снапшотов, связанных с доменным именем new.tomsk.gov.ru, которое использовалось для временного размещения второй версии сайта. Таким образом, выявленные варианты доменного имени отображают развитие веб-сайта во времени и присущую ему версиюность.

Согласно постановлению губернатора, первый веб-сайт Администрации Томской области был введен в действие в июле 1999 г.⁵ В Архиве Интернета он был впервые захвачен в январе 2000 г. На его главной странице в зоне копирайта указан 1999 г. Однако на страницах второй версии сайта, которая официально заменила первую в марте 2007 г.,⁶ в зоне копирайта в качестве начальной даты указан 1998 г. В то же время сама вторая версия стала доступна в сети под доменом new.tomsk.gov.ru в августе 2006 г., то есть гораздо раньше официальной даты. На захваченных в ноябре 2006 г. страницах в зоне копирайта указан период 1998–2005 гг., что отодвигает дату появления второй версии сайта еще дальше. Третья версия сайта, судя по веб-архиву, была размещена в сети между 17 мая и 8 июня 2012 г. Наконец, четвертая, современная версия сайта появилась в сети в июле 2015 г. При этом в последних двух версиях сайта датировка отсутствует.

В период с 1991 по 2012 г. губернатором Томской области был Виктор Кресс. Информация о его деятельности первоначально размещалась на сайте Администрации. Персональный сайт губернатора с доменным именем kress.tomsk.ru был создан примерно в 2002–2003 гг. У него выявлено две версии. Первая была захвачена в Архиве Интернета в июне 2003 г., однако в зоне копирайта на ее страницах в качестве начальной даты указан 2002 г. Вторая версия появилась в сети между июнем и декабрем 2009 г. На страницах сайта годы его существования не указаны. Сегодня обе версии доступны только в заархивированном виде, потому что после утверждения в должности в марте 2012 г. нового губернатора Томской области Сергея Жвачкина веб-сайт предыдущего губернатора сначала перестал обновляться, а затем в октябре 2012 г. исчез из Интернета в связи с прекращением обслуживания домена провайдером.

Информация о деятельности нового губернатора первоначально также отображалась на сайте областной администрации. Персональный сайт Сергея

Жвачкина появился в сети в декабре 2013 г. под доменным именем gubernator.tomsk.ru. Интересно, что сайт с таким доменным именем впервые был захвачен в Архиве Интернета в октябре 2002 г. и содержал объявление об услугах хостинга. За период с 2004 по 2012 г. его снапшотов в веб-архиве нет. Захват возобновился после начала работы сайта нового губернатора.

Таким образом, точно определить время появления в Интернете различных версий веб-сайта Администрации Томской области и губернатора пока не представляется возможным из-за противоречивости данных, которые зафиксированы в официальных документах и веб-архиве.

Содержание государственных веб-сайтов регламентируется федеральным и региональным законодательством, которое должно обеспечить своевременное и повсеместное распространение информации о деятельности органов власти. В то же время по итогам независимых мониторингов 2007–2013 гг. коэффициент открытости сайта Администрации Томской области колебался в пределах 30–54 %.⁷ В 2014 г. он составил 61 %.⁸ Веб-сайты губернаторов в мониторинге не участвовали. Исторический анализ веб-сайтов показал, что с течением времени их содержание постоянно расширялось. Это выразилось в усложнении рубрикации и более широком использовании мультимедийного контента. Можно выделить следующие группы материалов, которые позволяют реконструировать историю томской региональной администрации и института губернаторства с конца 1990-х годов по настоящее время:

1. Описание структуры администрации на уровне подразделений, должностных позиций, соответствующих полномочий и функций.
2. Информация о кадровой политике и персональном составе подразделений, включая фотографии и краткие биографии высокопоставленных сотрудников. Наиболее подробными являются данные о томских губернаторах, представленные на их сайтах.
3. Комплексы официальных документов, связанных с различными направлениями деятельности органов власти.
4. Новости, в том числе в мультимедийном формате, содержащие информацию о работе губернатора и региональной администрации, различных мероприятиях с их участием, а также важных событиях в жизни региона в целом.

Ретроспективный контент доступен исследователю в Архиве Интернета и в архивах действующих сайтов. Например, в архиве новостей третьей версии веб-сайта Томской администрации содержатся новости с апреля 2006 г., а в заархивированной первой версии – с 1999 года. На мой взгляд, исследуемые веб-сайты можно рассматривать как источники не только для реконструкции истории региональных органов власти, но и в более широком контексте, включая исторические изменения в государственной информационной политике или в технологиях веб-дизайна.

Признавая потенциальную возможность использования веб-сайтов томских органов исполнительной власти в исторических исследованиях, важно осознавать и связанные с этим проблемы.

Одной из самых сложных задач является выяснение особенностей функционирования сайта в информационной системе организации и способов формирования его контента. Это необходимо для того, чтобы адекватно оценить степень уникальности материалов, аутентичность и возможность их сохранения. Применительно к заархивированным веб-сайтам историк должен также понимать механизм веб-архивирования. Как удалось выяснить посредством интервью, в Администрации Томской области обязанности сотрудников по подготовке и размещению контента на сайте зафиксированы в должностных инструкциях и в положениях об отдельных подразделениях. Значительную часть содержания сайтов составляют документы, которые возникают в ходе текущей деятельности администрации и губернатора. Для их размещения делаются копии в нужном формате, а исходный документ продолжает храниться в системе управления документами, причем как в электронной, так и в бумажной форме. Другая часть контента специально создается или перерабатывается для публикации на сайте, например тексты разделов, новости, изображения, коллекции открытых данных. Именно эти материалы являются уникальными как для самой администрации, так и в интернет-пространстве. Однако их учет и организация хранения особо не регламентируются. Поскольку в администрации отсутствует практика архивирования веб-сайтов, возникает ситуация, когда наиболее уникальная часть контента является самой уязвимой с точки зрения долговременной сохранности.

Проблематичным, как уже говорилось, является определение времени создания различных версий сайтов, их размещения в Интернете и точная датировка отдельных материалов, поскольку на рассматриваемых веб-сайтах (за исключением текущей версии) отсутствуют даты обновления страниц. Особенно это касается фотографий и контента, изменение которого происходит с полной или частичной заменой предыдущего.

Упомянутая в литературе проблема разграничения объекта исследования в веб-архиве также является очень важной.⁹ Очевидно, что веб-сайты томских региональных органов власти, сохраняемые в Архиве Интернета, не являются полной копией исходных сайтов, потому что содержат только часть страниц и контента, особенно мультимедийного. Кроме того, переход по ссылкам очень часто связывает между собой страницы, захваченные в разное время, то есть историк не сможет выделить объект исследования, базируясь на одновременности его фиксации в архиве. Таким образом, заархивированные веб-сайты фактически распадаются на отдельные страницы, поскольку даже при формальном сохранении структуры ссылок сайт теряет свою целостность. Все это является неизбежным

и труднопреодолимым следствием самой технологии веб-архивирования. Возможным решением может стать анализ статистических данных веб-архива о количестве захваченных URL и выявление периодов с наибольшей плотностью захвата. Также для исторического исследования может оказаться полезным сравнительный анализ отдельных страниц на протяжении всего времени, в течение которого они захватывались.

Таким образом, веб-сайты безусловно можно рассматривать как самостоятельные исторические источники, значительно отличающиеся от традиционных. Они обладают не только технологической, но и содержательной спецификой. Применение к ним фундаментальных принципов традиционного критического анализа источников потребует от историков развития дополнительных компетенций и новых методологических подходов.

Литература:

1. Classen C., Kinnebrock S., Löblich M. Towards Web History: Sources, Methods, and Challenges in the Digital Age // Historical Social Research / Historische Sozialforschung. 2012. Vol. 37. No. 4 (142). Pp. 97–101; Brügger N. When the Present Web is Later the Past: Web Historiography, Digital History, and InternetStudies // Historical Social Research / Historische Sozialforschung. 2012. Vol. 37. No. 4 (142). Pp. 102–117.
2. Архив Интернета // Википедия. Свободная энциклопедия. https://ru.wikipedia.org/wiki/Архив_Интернета
3. List of Web archiving initiatives // Wikipedia. The Free Encyclopedia. URL: https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives#External_links (дата обращения 05.07.2015).
4. Internet Archive. URL: <https://archive.org> (дата обращения 06.07.2015).
5. Постановление Губернатора Томской области от 9 июля 1999 г. № 259 «Об официальном информационном web-сервере Администрации Томской области» // Портал Томская область. Томск, 2008–2015. URL: http://tomsk.news-city.info/docs/sistemso/dok_percli.htm (дата обращения: 20.04.2015).
6. Новость от 14 марта 2007 // Администрация Томской области. URL: http://tomsk.gov.ru/ru/press-centr/press-relizy/news_item/-/novost-zarabotal-odernizirovannyysayt-oblastnoy-administratsii (дата обращения 12.04.2015).
7. Мониторинг // Фонд свободы информации. Санкт-Петербург. URL: <http://old.svobodainfo.org/ru/node/6> (дата обращения 15.03.2015).

8. 2014: Высшие исполнительные органы государственной власти субъектов РФ. Рейтинг // Информетр. [Б.м.], 2014–2015. URL: <http://system.infometer.org/ru/monitoring/160/rating> (дата обращения: 15.03.2015).

9. Brügger N. Website History and the Website as an Object of Study // New Media & Society. February/March 2009. Vol. 11. No. 1–2. Pp. 115–132.