

## **Устойчивое обеспечение сохранности цифровой информации в различных контекстах**

Обеспечение сохранности цифровой информации можно рассматривать как экономическую деятельность с общими для всех цифровых материалов и зависимыми от контекста атрибутами, которые влияют на принятие решений в течение всего жизненного цикла ресурса. Однако сами стратегии сохранения представляют собой серию решений, принимаемых в конкретных ситуациях, связанных с другими видами деятельности – от исследований и разработок до публикации и составления годовой сметы текущих расходов. Вся эта деятельность проводится в специфических социальных контекстах. В этой главе мы посмотрим, какие решения в области сохранения принимались в различных контекстах в реальных обстоятельствах. Наш анализ основан на результатах исследования двух десятков примеров – «кейсов», в которых контент распределялся по следующим четырем типам: научная коммуникация, результаты исследований, коммерческий контент в области культуры и веб-контент, являющийся результатом коллективного творчества. Эти типы контента особенно важны для естественных и гуманитарных наук, исследований и образования, государственной политики, культурного наследия и творческих индустрий. Кроме того, это исследование демонстрирует различный уровень развития стратегий сохранения.

### **<...> 1. Научный дискурс**

Научный дискурс – часто называемый научной коммуникацией – это опубликованные результаты научных исследований: идеи, теории, аналитические данные, оценки результатов предшествующей научной работы, выводы, которые все вместе и формируют научную информацию. Аудитория у научной литературы сравнительно однородная, концентрированная, с высокоспециализированным кругом чтения, разделяющая общее мнение относительно ценности научной информации в целом, а также вклада индивидуальных личностей в развитие научных знаний. Участники процесса четко определены: это ученые, которые являются и создателями, и пользователями информации; исследовательские институты, которые их поддерживают и, в общем и целом, финансируют научную деятельность, научные публикации и обеспечение сохранности в научных библиотеках; издательства, как коммерческие, так и некоммерческие, представляющие интересы ученых как авторов; библиотеки, которые представляют интересы ученых как пользователей и представляют распределенный спрос на доступ к информации по различным дисциплинам разных поколений пользователей.

Многие века исследовательское и образовательное сообщество занимались сохранением научного дискурса, и необходимость сохранять информацию в цифровых форматах широко признана как естественное продолжение функции сохранения ее традиционных предшественников. Кроме того, сложилось распределение ответственности за обеспечение сохранности как наследие печатной эры (иногда не очень устойчивое). Этот баланс интересов сторон нарушен цифровыми технологиями двояким образом. Во-первых, нарушены традиционные функции и ответственность участников, что объясняется неконкурентным характером цифровых ресурсов и имеет результатом разбалансированность стимулов и проблему неоплачиваемого использования...

Во-вторых, возникли новые жанры научной коммуникации, такие как онлайн-пространство для совместной работы, академические блоги, веб-сайты, электронные публикации и даже микроблоги в социальных сетях (например, в Twitter). За исключением тех областей, в которых сохранилась традиция выпуска «серой» литературы и препринтов – например, в физике и экономике. Нигде больше нет единого мнения относительно долговременной ценности этой информации и не разработаны практические методы отбора этих новейших материалов для долгосрочного хранения.

С целью дальнейшего анализа будем различать два типа научной коммуникации: *традиционные* и *новейшие* жанры. Мы осознаем, однако, что в реальности трудно провести такое разделение, и оно способно дезориентировать исследователя.

### ***1.1. Ценность и отбор***

Традиционные формы научной коммуникации по природе своей хорошо разработаны, и в цифровых форматах необходимо повторять все конвенции, принятые для печатных материалов, такие как ссылки и отсылки, указатели, таблицы, весь справочный аппарат, специфичный для определенных дисциплин. Ученые требуют не только тщательной разработки функциональности научных журналов или монографий в цифровых форматах, но также хотят получить гарантии, что научный документ представлен во всей полноте. Эти документы сохраняются сетью библиотек, неформально распределившей ответственность за сбор этой информации. Благодаря высоким стандартам точности, надежности, функциональности в использовании, которые требует научное сообщество, спрос на научную коммуникацию обычно реализуется по двум направлениям, что означает, что необходимо сохранять один и тот же объект в двух версиях.

Первое направление – это читательский спрос на *доступный цифровой объект*, версию, которая доступна ему в онлайн-режиме и

сопровождается всевозможными дополнительными ценными услугами и инструментами, например ссылками на информацию, которую читатель может найти в онлайн-версиях журналов *Science* и *Nature* или через службу *ScienceDirect* издательства *Elsevier*<sup>1</sup>.

Другой тип спроса – это спрос института-хранителя на *архивную копию цифрового ресурса*. Это то, чем управляют издательства, библиотеки, представляющие третью сторону архивные службы, например *Portico*. (Существует еще одна служба обеспечения сохранности – *LOCKSS*, которая позволяет библиотекам самим архивировать свою информацию.) Спрос на услуги архивирования с созданием архивной копии исходит от самих поставщиков журналов (издатели или библиотека); он связан со спросом на страховочные услуги, призванные защитить предшествующие инвестиции в услуги по предоставлению доступа, такие как *ScienceDirect* или *JSTOR*. Эти архивные копии иногда существенно отличаются от копий, оптимизированных в целях обеспечения доступа.

**Неопределенность будущей ценности.** Многие новые формы научной коммуникации, такие, например, как академические блоги и научный видеохостинг *SciVee*, могут быть утрачены безвозвратно, так как до сих пор не определена их долговременная ценность, и очень немногие институты готовы взять на себя ответственность за их длительное хранение<sup>2</sup>. Однако в краткосрочной перспективе эти ресурсы используются очень интенсивно: например, многие ученые пишут блоги. Однако за пределами этих дисциплин (главным образом, естественных и общественных наук, с давними традициями препринтов и серой литературы) научная информация, представленная в новых формах, не собирается и не сохраняется. В настоящее время интерес к новейшим формам научной коммуникации проявляют лишь частные лица, которые могут не представлять себе (или, наоборот, представлять), какие существуют угрозы устойчивости (в случае блогов, твитов и пр.), и не могут существенно повлиять на проблему (например, в рамках проектов, реализуемых по единовременному гранту).

По запросам ученых библиотеки принимают на себя ответственность за сохранение новых публикаций. Однако отбирать отдельные публикации, в которых заинтересован отдельный ученый, нерационально и не соответствует задачам научной коммуникации в целом. Средство преодоления этого противоречия в профессиональном сообществе – это сотрудничество с учеными с целью совместной оценки политики в

---

<sup>1</sup> 2014 AAAS Annual Meeting. – Mode of access: <http://www.sciencemag.org>; *Nature.com*. – Mode of access: <http://www.nature.com>; <http://www.sciencedirect.com>.

<sup>2</sup> Making science visible. – Mode of access: <http://www.scivee.tv>.

отношении долгосрочной ценности новейших средств научной коммуникации. Несколько пилотных внутридисциплинарных проектов по сбору информации вряд ли обойдутся очень дорого. Научные материалы, представленные в Интернете, могут скачиваться с сайтов, что не потребует столько средств, сколько приобретение печатных изданий или подписка на академическую литературу. В данном случае, стратегия малых инвестиций поможет снизить риск безвозвратной утраты и может оказаться весьма уместной.

**Рекомендация № 1.** Библиотеки, ученые и профессиональные общества должны выработать критерии отбора новейших жанров цифровой информации в научной коммуникации, а также стратегию обеспечения сохранности прототипа и стратегию доступа.

### ***1. 2. Стимулы***

Если для печатных публикаций существует четкое распределение сфер ответственности за сохранение и определены институты, уполномоченные действовать в интересах сегодняшних и будущих пользователей, то для цифрового мира инфраструктуры сохранения нет. Существующие законы об авторском праве не указывают на то, кто технически обеспечивает сохранность цифровых материалов, например отвечает за создание более трех копий объекта. Более того, закон об авторском праве может быть опрокинут контрактами. Ученый делегирует издателям права на опубликование и распространение информации, а издатели добавляют услуги по рецензированию и редактированию и предоставляют обновленные материалы библиотекам по лицензии, где ученые пользуются ими. У издателей имеются серьезные стимулы, чтобы сохранять эти материалы, таким образом удерживая лучших авторов и поддерживая эффективность коллективных цифровых ресурсов. Тем не менее, издательства – это не те организации, которые должны заниматься сохранением научного наследия на благо общества в долговременной перспективе. Для научного наследия долговременная перспектива означает несколько веков, а это значительно больше времени жизни большинства, хотя и не всех, издательских компаний.

***Разбалансированность или конкуренция стимулов.*** Несовпадение стимулов издателей и библиотек представляет потенциальный риск для устойчивости электронных журналов. В настоящее время при архивировании электронных журналов эта проблема зачастую решается путем передачи издателем ответственности третьей стороне, которая берет на себя сохранение материалов в долговременной перспективе (например, JSTOR, Portico, KB), в некоторых случаях при финансировании от издательства. Издатели не слишком противились идее «темного архива» – архива, не

обеспечивающего доступа пользователям, но действующего в интересах общества. Такие модели в случае их экономической устойчивости дают адекватное решение проблемы сохранности. Однако важно отметить, что прямая передача ответственности за сохранность не гарантирует устойчивости сохранности. Хотя у организации, желающей заниматься сохранностью, есть стимул к этому, могут возникнуть вопросы по поводу того, кто будет платить за сохранение в долгосрочной перспективе.

Архивы третьих сторон могут сыграть здесь важную посредническую роль. Но не менее важную роль могут играть сами ученые – как коллективную, так и индивидуальную, – если сохранят за собой права на свои труды и передадут другой стороне неисключительные права на их сохранение. Ученые, представляющие спрос, накапливают потенциал для выдвижения требований к сохранности как части соглашения о доступе. Примером может служить британская модельная лицензия на сохранение журналов NESLi2, которую использует JISC при обсуждении лицензионных условий на использование электронных журналов от имени высших учебных заведений Великобритании. В лицензию включены положения, согласно которым издательства должны депонировать копии журналов в соответствующих архивных организациях. Не совсем ясно, может ли эта модель быть успешно распространена на более обширное и распределенное сообщество в сфере высшего образования, например, в Соединенных Штатах, однако, в целом, заинтересованные игроки, действующие на стороне спроса, должны коллективно выражать свои потребности стороне, представляющей предложение услуг по сохранению информации.

Отдельные ученые – поставщики научного знания первого порядка – могут облегчить задачу обеспечения сохранности, оговорив, что они (или, в необходимых случаях, институты, которые они представляют) владеют бессрочной неисключительной лицензией на свои работы, которая не может быть передана третьей стороне. Гарвардский университет, Массачусетский технологический институт (MIT), Университет Канзаса и другие выдают такие лицензии от имени своих ученых. Если такая политика была бы распространена повсеместно и если бы подобная практика была принята, большую часть научных работ могли бы сохранять, помимо издательств, другие организации, главным образом репозитории этих институтов, библиотеки, третья сторона. Коллективная переговорная сила поможет защитить права ученых в самых широких масштабах, а не только права отдельных институтов. Однако эффект будет таким же, если большинство институтов примет такие же или аналогичные лицензии. Индивидуальное использование таких лицензий также позволит уменьшить препятствия к

сохранению новейшей информации, поскольку библиотеки и другие доверенные институты получают это право. И вновь отметим, что преодоление разрывов в стимулировании тем или иным способом не решает вопроса о финансировании процесса. Соответствующее финансирование может быть получено в рамках как централизованной, так и распределенной модели, и в некоторых случаях это уже произошло. Тем не менее, преодоление разрывов в стимулировании позволит сделать так, чтобы одни и те же его участники имели право сохранять информацию и выигрывали от этого процесса.

**Рекомендация № 2.** В целях долгосрочного обеспечения сохранности издательствам, имеющим право на сохранение информации, следует сотрудничать с третьей стороной – архивами или библиотеками.

**Рекомендация № 3.** Ученым следует рассмотреть возможность передачи неисключительных прав на публикацию и сохранение, что позволит децентрализовать и распределить сохранение вновь возникающих научных дискурсов (*Устойчивая экономика для цифровой планеты: обеспечение долговременного доступа к цифровой информации. Итоговый отчет Рабочей группы по устойчивому обеспечению долговременной сохранности и доступа к цифровой информации [Перевод с англ.]. – М.: МЦБС, 2013. – С. 104–113*).