

МЕЖДУНАРОДНЫЙ ФОРУМ ПО ИНФОРМАЦИИ

2020

Том 45

№ 1

Выходит ежеквартально

Основан в 1975 г.

СОДЕРЖАНИЕ

Миллайон А. Дж., Мулесон Сэнди Х., Хадсон-Витале С.

Реструктуризация и формализация: научная коммуникация
как возможность устойчивого роста в информационных
организациях?

3

Добрески Б. Аналитический подход к информационным стандартам
на основе ценностей

13

Майерник М. Связывание научных ресурсов: построение жизненного
цикла их связи

22

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

д.филол.н. ГИЛЯРЕВСКИЙ Р.С.

ЧЛЕНЫ РЕДКОЛЛЕГИИ:

к.т.н. БЫКОВ В.А. (РОССИЯ), к.физ.-мат.н. ВАРДАНЯН Г. Г. (АРМЕНИЯ),

д.т.н., проф. ВОЙТОВ И. В. (БЕЛАРУСЬ),

МИРАЛИЕВ К. Х. (ТАДЖИКИСТАН), МОЛДОШЕВА Д. А. (КРГИЗИЯ)

РЕДАКТОРЫ:

КОБЗЕВА Л.В., ОВЧЕНКОВА Е.А.

Реструктуризация и формализация: научная коммуникация как возможность устойчивого роста в информационных организациях?*

А. Дж. МИЛЛАЙОН

(A. J. MILLION)

Университет шт. Мичиган, США

Хезер МУЛЕСОН СЭНДИ

(Heather MOULAISON SANDY)

Университет шт. Миссури, США

Синтия ХАДСОН-ВИТАЛЕ

(Cynthia HUDSON-VITALE)

Университет шт. Пенсильвания, США

Появляющиеся технологии революционизируют сферу научной коммуникации. Поэтому ученые все больше нуждаются в особой поддержке на всех этапах исследовательского процесса. На примере академической (научной) библиотеки, взятой в качестве единицы анализа, рассматриваются два понятия из теории «Диффузия инноваций» (DOI) Роджерса и литературы по организации инноваций, чтобы оценить возможность устойчивой работы по научной коммуникации в библиотеках. Это понятия организационной реструктуризации и формализации. Анализируются данные по сотрудникам Ассоциации научных библиотек (Association of Research Libraries, ARL) с соответствующими названиями занимаемой должности и трем документам о навыках в цифровом кураторстве. Результаты исследования предполагают, что информационные организации ARL реструктурированы, чтобы обеспечивать дополнительную поддержку исследований и чтобы навыки, связанные с должностью в научной коммуникации, становились более унифицированными. Делается вывод, что информационные специалисты в области научной коммуникации являются частью устойчивой практической области в рамках информационных организаций ARL, доминирующей последние десятилетия, и что подобная тенденция, скорее всего, продолжится, по крайней мере, в краткосрочной перспективе.

ВВЕДЕНИЕ

Появление новых технологий по поддержке издательства и коммуникации (включая социальные сети, онлайн службы управления библиографическими ссыл-

ками и платформы обмена) фундаментально изменило метод работы ученых [1] и сам исследовательский процесс, в который они вовлечены. Научная коммуникация становится все сложнее в результате новых и появляющихся технологий. Научная коммуникация, область, поддерживаемая информационными специалистами, определяется следующим образом: исследование того, как ученые любой области (например, физики, биологии, социальных и поведенческих наук, гуманитарных и технических наук) используют и распространяют информацию через формальные и неформальные каналы. Исследование научной коммуникации включает рост научной информации, отношения между научными областями и дисциплинами, информационные потребности

* Перевод Million A. J., Moulaison Sandy H., Hudson-Vitale C. Restructuring and formalizing: Scholarly communication as a sustainable growth opportunity in information agencies? // Proceedings of 81 st Annual Meeting ASIS&T. — P. 377-386. — <http://www.asist.org/wp-content/uploads/Final-81st-Annual-Meeting-Proceedings-1.pdf>

сти и использование информации отдельными группами пользователей, а также отношения между формальными и неформальными методами коммуникации [2, с. 13-14].

Научная коммуникация дает возможность осуществлять работу исследователей в университетах и организациях, занимающихся созданием нового знания. В США свыше половины федеральных ассигнований (в долл.) выделяются на гранты работающим в университетах исследователям [1], делая тем самым университеты центром создания нового знания.

ПОДДЕРЖКА УЧЕНЫХ

Поддержка научной коммуникации в университетских кампусах в век появляющихся технологий естественным образом падает на информационных специалистов, базирующихся в информационных организациях, таких как академические (научные) библиотеки. Традиционно научные библиотеки делятся опытом, который они предоставляют в технических и публичных сферах, но учитывая сложность технологий в сочетании с постоянно появляющимися комплексными потребностями их пользователей, должно возникнуть новое ответственное информационных специалистов [3]. В случае появления вопросов относительно научной коммуникации информационные специалисты не только курируют и поддерживают доступ к научной сфере, они также должны работать вместе с учеными, чтобы организовывать, предоставлять доступ, сохранять и обмениваться работами. Обеспечение особой поддержки ученых (например, [4]) в течение исследовательского процесса быстро становится сферой научной библиотеки.

По мере изменения потребностей исследователей и ученых параллельно развивается работа информационных специалистов. Работа по поддержке области «научной коммуникации сегодня отражает [потребность в применении растущего] созвездия средств, практик и навыков» [5, с. 365]. Поддержка информационных специалистов в области научной коммуникации, предоставляемая исследователям, варьируется в соответствии с приоритетами организаций, в которых они работают, областями знания ученых, которые они поддерживают, и их собственными навыками и подготовкой. В целях данной статьи рассматриваем работу, поддерживающую научную коммуникацию как помощь в цифровом кураторстве, управлении исследовательскими данными и открытом доступе и издательстве. Хотя знания и навыки по поддержке этой работы систематизированы в ряде документов (некоторые из которых анализировались в данной статье), мало что известно об устойчивости работы в сфере научной коммуникации в информационных организациях.

В статье оценивается роль информационного специалиста в научной коммуникации в стремлении понять устойчивость этой нетрадиционной работы. Чтобы оценить устойчивость, проведен обзор литературы, содержащей развивающиеся роли и задачи научных библиотек и роли, которую информационные специалисты могут играть в поддержке научной коммуникации в своих организациях. Затем представлены понятия *реструктуризации* и *формализации* из теории DOI Роджерса и литературы по организации инновации с предположением, что должностные позиции в научной коммуникации в последнее время рассеиваются среди информационных

организаций.* Вооруженные понятиями из теории DOI, мы обсуждали результаты анализа учета статистических данных о работе ARL и качественной оценки трех документов по компетенциям в научной коммуникации, свойственным цифровым кураторам. На основе наших наблюдений делается вывод, что библиотеки ARL реструктурированы, чтобы дополнять позиции в научной коммуникации благодаря их ценности, придаваемой информационными специалистами, а распределение документов о компетенции предполагает, что работа информационных специалистов становится более универсальной. По прогнозам DOI, реструктуризация и формализация в организациях связаны с более низкими темпами изменений. Таким образом, свидетельство предполагает, что, скорее всего, информационные специалисты в научной коммуникации останутся на ближайшее время поддерживающим дополнением в информационных организациях ARL. Завершается статья обсуждением ограничений нашего анализа и предложениями направлений дальнейшего исследования.

Устойчивость

В целях этой статьи определим *устойчивость* как стабильность определенных ролей и ответственностей внутри информационных организаций. Отчет Брандта, выпущенный ООН, [7] приводит другое определение, утверждающее, что *устойчивость* является «развитием, отвечающим потребности настоящего без ущерба способности будущих поколений удовлетворять свои собственные потребности» [7, с. 37]. Применительно к работе в информационных организациях это определение подразумевает, что устойчивость встречается тогда, когда отдельные лица ведут работу, которая не ставит под угрозу способности их преемников делать то же самое. Мы не уменьшаем ценность этого определения; однако оно не является полезным для применения в этой статье. Причина, по которой принимается первое определение, кроется в том, что мы непосредственно заботимся о сохранении научной коммуникации как ряда позиций и практик в информационных организациях ARL. Таким образом, нас интересует вопрос, будут ли краткосрочные практики специалистов по научной коммуникации, вероятно, опасны или контрпродуктивны в долгосрочной перспективе.

Вопросы исследования и обоснование

Сформированная на основе обзора данных исследования с 2012 г. рабочая группа по научным коммуникациям Совета университетских библиотек бассейна Тихого океана (Council of Prairie and Pacific University Libraries – COPPUL) обнаружила, что область научной коммуникации до сих пор не объединилась, чтобы сформировать сообщество практики в библиотеках [8]. Тем не менее, после десяти лет работы семинара по на-

* Это предположение основано на работе [5], посвященной истории семинара ACRL Scholarly Communication Roadshow Ассоциации университетских и научных библиотек (Association of College and Research Library, ACRL) и предшествующим усилиям по обучению библиотекарей относительно необходимости вовлекаться в работу по научной коммуникации.

учным коммуникациям ACRL авторы работы [5] обнаружили, что область значительно окрепла.

С учетом проводимой этими информационными специалистами работы и поддержки, необходимой исследователям и ученым из-за изменения технологий, в статье изучаются два исследовательских вопроса (RQ1 и RQ2):

RQ1. Заглядывая вперед — насколько устойчивой должна быть работа информационных специалистов в сфере научной коммуникации в информационных организациях ARL?

RQ2. В какой степени подобласть практики научной коммуникации (т.е. цифровое кураторство) проявилась в качестве связующей области практики?

Первый вопрос исследования (RQ1) изучал основную проблему, рассматриваемую в этом исследовании. Вопрос RQ2 сохраняет приверженность подобласти практики научных коммуникаций, которые должны рассматриваться при ответе на вопрос RQ1.

ОБЗОР ЛИТЕРАТУРЫ

Научная коммуникация — это область, необходимая для адаптации к современной технике, исследовательским подходам и методам коммуникации. Информационные специалисты, поддерживающие научную коммуникацию, появились в помощь исследователям в управлении этим новым ландшафтом. Как объясняют авторы [9], «деятельность библиотекарей по поддержке коммуникации увеличивается и изменяется в последние годы из-за усложнения природы современной цифровой научной сферы» [9, с. 137]. Например, автор [1] идентифицирует большие данные и большую науку как два изменения, глубоко повлиявшие на исследователей и совершаемую ими работу. Авторы канонического текста [10] также правильно идентифицируют отмеченный сдвиг в сторону вычислительного и использующего данные исследования в академической среде. Более того, меморандум Американского бюро по политике в области науки и техники «Расширение публичного доступа к результатам исследований, финансируемых на федеральном уровне» обогатил развитие библиотечных служб, поддерживающих открытый доступ и обмен научными результатами. Информационные специалисты по научной коммуникации научились управлять все более сложной областью научной коммуникации и связанными с ней технологиями и работают в целях поддержки ученых в использовании этих технологий через предлагаемые ими услуги. Автор [11] уточняет, что в рамках пересечения цифровых библиотек и научной коммуникации информационные специалисты все больше интересуются услугами, которые могут обеспечивать цифровые библиотеки [11, с. 415]; таким образом, можно сказать, что они похожи на библиотекарей технических служб и библиотекарей публичных библиотек прошлого.

Научная библиотека в развитии

В ответ на меняющиеся потребности своих пользователей и в поддержку новых сервисов, которые они сегодня должны предоставлять, научные библиотеки в корне меняются, переходя от фронтальной бинарной/обратной бинарной модели к конвергенции обеих моделей [12, 3]. Библиотечная реорганизация была основным способом, которым научные библиотеки позицио-

нируют себя для будущего доступа. В анализе литературы авторы [13] определяют пять шагов, предпринимаемых, как правило, библиотеками в ходе процесса реорганизации; библиотеки — 1) идентифицируют направления изменений; 2) проводят анализ и диагностику; 3) обсуждают и сообщают план изменений; 4) внедряют изменения; 5) проводят дальнейшую постоянную оценку. Назначение новых задач персоналу на основе *анализа* и *диагностики* происходит во время реализации шага *изменения*. На этом этапе персонал обучается новым навыкам, которые могут ему понадобиться, и идеально занимает должностные позиции согласно своим навыкам, талантам и призванию. Следовательно, персонал можно принять на работу в соответствии с рядом условий и профессиональных ожиданий и подобрать ему новую работу в других областях.

Информационные специалисты, поддерживающие научную коммуникацию

Названия новых должностей и потребностей появляются в целях поддержки изменений в научных библиотеках и изменений в практике академического исследования. Такие позиции больше не требуют от информационных специалистов иметь степень в области библиотечного дела и информатики, полученную в организациях под эгидой Американской библиотечной ассоциации (ALA). Исследование, проведенное авторами [14], обнаружило, что библиотекари, поддерживающие службы управления научно-исследовательскими данными, обладают предметным опытом в разных научных дисциплинах, помимо или взамен традиционной степени в области библиотечного дела и информатики. Поэтому такие информационные специалисты могут иметь разные знания и могут иметь традиционную степень в области библиотечного дела и информатики, других областях специализации или и то, и другое.

Смешно, но отмечался рост информационных специалистов в области научной коммуникации в научных библиотеках, но о них кроме пола, мало что известно из списка членов группы SCHOLCOMM в системе Zotero [15]. Согласно данным литературы, в рамках своей работы информационные специалисты в этой сфере ответственны за ряд *подвидов деятельности*. Предполагается, что они готовы помогать преподавателям и другим пользователям в процессе научной коммуникации [16]. Поощрение движения в сторону открытого доступа для библиотечных пользователей также считается проблемой научной коммуникации [17,18].

Одной из первых подобластей научной коммуникации является цифровое кураторство и, кроме того, управление научно-исследовательскими данными. Чтобы поддерживать ученых в исследовательском процессе, информационные специалисты в области научной коммуникации также должны помогать в курировании и использовании данных и цифровых массивов в свете ранее упомянутых технологических инноваций [14]. Действительно, открытый доступ и *открытые исследовательские услуги* присущи развитию поддержки управления данными исследований и курирования. Авторы работы [14] обнаружили, что 85% учреждений, архивирующих данные, участвуют в этом процессе, чтобы поддержать открытый доступ [14, с. 56]. Более того, одна из основных библиотечных служб по управлению научно-исследовательскими данными включает консультации и

образование в сфере публичного доступа ко всем научно-исследовательским результатам, не только данным, но и кодам, анализу, протоколам, рабочим процессам, публикациям и т.п.

Информационные специалисты, работающие в этой сфере, также стремятся понять информационные потребности своих пользователей и предпочтения в цифровой информации [19]. Авторы работы [20] обнаружили, что многие организации развивают службы поддержки ученых в управлении их научными профилями и отслеживании научно-исследовательских результатов. Для поддержки этой службы информационные специалисты развили навыки в анализе данных, гуманитарных областях цифровой среды, управлении данными и понимании разных типов метрик. Короче говоря, по мере развития данной области информационные специалисты в научной коммуникации переходят от поддержки открытой научной сферы к активно предоставляемым услугам, относящимся к передаче и публикации исследований, исследовательским данным и дополнительным аспектам научного исследования.

ДИФфуЗИЯ ИННОВАЦИЙ

Предсказывать будущее трудно, но теория Роджерса «Диффузия инноваций» (DOI) и литература по организации инноваций содержат понятия, касающиеся изучения устойчивости работы по научной коммуникации в информационных организациях Ассоциации научных библиотек (ARL). Ни одна теория не может предсказать будущее; однако DOI является одной из самых цитируемых работ в научной литературе по социальным наукам [21]. Ниже вводится DOI и описывается, как теория помогает предсказать будущее изменение в подобных библиотекам информационных организациях с акцентом на понятиях реструктуризации и формализации организаций. В литературе по социальным наукам DOI (теория) используется, чтобы объяснить, как и почему распространяются инновации. *Инновация* определяется как «любая идея, практика или проект, воспринимаемые отдельным лицом как что-то новое» [6, с. 11-12]. В данной статье определим инновацию как должностные позиции научной коммуникации внутри библиотек ARL. Согласно работе ряда социологов [22, 23], Роджерс достоин уважения за популяризацию DOI путем систематизации 508 исследований, объясняющих, почему индивидуумы и организации принимают инновации [24]. В ходе своей работы Роджерс идентифицировал четыре повседневных фактора: инновация, коммуникационные каналы, время и социальная система [6, с. 11]. Ниже сконцентрируемся на инновации (например, позициях научной коммуникации) и социальных системах (например, библиотеки ARL).

Инновации в организациях

Организации — это «организованные с определенной целью группы людей», они склонны быть комплексами [25, п. 1]. Первоначально DOI появилась как теория, сфокусированная на решениях индивидуумов. Более поздние разработки распространили это описание на инновации в организациях, такие как научные библиотеки, через использование метафор. Например, одной из них является метафора организации процесса принятия решений.

Что касается метафоры организации процесса принятия решений, то один способ, который Роджерс распространил на DOI, был основан на создании шестиступенчатой модели. При разработке Роджерс утверждал, что организации действуют, как если бы они владели агентством и стремились снизить неуверенность относительно расходов и прибылей, связанных с принятием и отказом от инноваций. Шестью этапами решения организационной инновации являются: принятие повестки дня, соответствие, переопределение, реструктуризация, уточнение и установившаяся практика [6, с. 421]. *Принятие повестки дня* касается того, где цели и повестка дня устанавливаются лидерами. *Соответствие* описывает персонал, занятый поиском и нахождением решений, удовлетворяющих повестку дня. *Переопределение* и *реструктуризация* предполагают, что после завершения этапа соответствия повестка дня определена и инновации приняты. Наконец, *уточнение* описывает процесс получения участия от персонала [6, с. 427] и *установившаяся практика* приходит тогда, когда персонал больше не воспринимает инновацию как новую [6, с. 429]. В этой статье мы не фокусируемся на решениях по созданию позиций научной коммуникации, но Роджерс предложил способ обозначения изменений в информационных организациях, которые уже произошли в процессе распространения. Созданием позиций в научной коммуникации утверждается, что это изменение относится к реорганизации библиотек.

Другой основной метафорой в этой статье является понятие организационной структуры. Автор [26] констатирует, что организационные структуры — это «постоянства в действиях, таких как размещение задачи, координация и предвидение» [26, с. 1], а Роджерс утверждает, что существуют шесть особенностей, определяющих, будут ли меняться организации [6]. Такими структурными особенностями, укоренившимися в литературе по организации инноваций, являются: централизация, формализация, сложность, взаимосвязанность, организационная слабость и размер. *Формализация* описывает степень, в которой «организация делает акцент на исполнении ее сотрудниками соответствующих правил и процедур» [6, с. 411], и она интерпретируется нами как создающие должностные позиции информационные организации, расположенные на одной линии с профессионально унифицированными практиками.

Литература, изучающая инновации в организациях с точки зрения структуры, поддерживает аргументы Роджерса — особенно в отношении формализации. Формализованные роли должностей ассоциируются с более низким темпом инноваций. Например, в исследовании социального благосостояния организаций авторы [27] обнаружили, что низкая степень систематизации должностей ассоциируется с высокой скоростью перемен. В 1981 г. Говард проанализировал четыре университетские библиотеки и пришел к выводу, что темп инноваций отрицательный в централизованных, формальных и стратифицированных организациях. Недавно авторы работы [28] провели анализ 260 организаций с помощью опроса и установили, что высокая степень организационной формализации снижает скорость принятия решений и организации инноваций. Формализация является не только предвестницей изменений в организации, но и находится по Роджерсу в одном ряду с фазой реструктуризации процесса принятия решений об инновации [6].

МЕТОДОЛОГИЯ

Оценка распространенности позиций научной коммуникации в библиотеках ARL

Чтобы понять рост числа информационных специалистов в научной коммуникации, принятых на работу в крупные академические библиотеки Северной Америки, проанализированы данные позиций, систематизированные в исследовании зарплат сотрудников Ассоциации научных библиотек (ARL) [29-31]. Хотя организации ARL не представляют все информационные учреждения, поддерживающие исследования, они включают несколько самых крупных и самых престижных академических библиотек в США и Канаде. Раз в год 125 библиотек ARL предоставляют данные персонала для всех позиций. Координаторы опроса выбрали один код должностной позиции, чтобы применить к персоналу, позже эти данные были упорядочены, нормализованы и собраны [32].

В данном исследовании рассматривались данные позиций, систематизированные как библиотекарь сферы научной коммуникации (SCHOLAR) и библиотекарь сферы цифрового кураторства (DIGICUR). Код SCHOLAR определяется в исследовании зарплат как член персонала библиотеки, вовлеченный в научную коммуникацию. Эти индивидуумы работают с удаленным открытым доступом или продвигают его, дают совет в сфере авторского права и т. п. Код DIGICUR определяется как член библиотечного персонала, который создает и курирует цифровые массивы в естественных науках, общественных или гуманитарных областях или который работает с вопросами управления данными в нескольких дисциплинах. Как отмечалось ранее, цифровое кураторство является подобластью научной коммуникации. Обе эти позиции были проанализированы с учетом их поддержки коммуникации и издательства в научной сфере. 2012 г. считается временем первого сбора данных ARL по этим двум категориям.

Оценка навыков и компетенций

Ученые также анализировали рекомендуемые навыки и компетенции, необходимые в *цифровом кураторстве и управлении данными*, области, соединяющей технологический опыт администрации с предоставлением обслуживания при поддержке процесса научной коммуникации. Команда проекта применила многоэтапный подход: 1) идентификация релевантных руководств; 2) сбор компетенций/навыков как способ стимулирования унификации разметки в процессе систематизации; 3) индуктивная разработка широких категорий; 4) систематизация каждой компетенции в категорию.

Для поиска как можно большего числа релевантных навыков цифрового кураторства и отчетов относительно рекомендаций или исследований, применялся следующий подход:

3. Базы данных поиска: Library Literature & Information Science; Library, Information Science & Technology Abstracts; Google Scholar; Google

4. Используемые ключевые слова: «навыки цифрового кураторства», «программы обучения в сфере цифрового кураторства», «обучение цифровому кураторству».

Поиск завершился 418 статьями. Учитывая, что поиск осуществлялся с целью понять компетенции и навыки цифрового кураторства и управления исследовательскими данными, — особый тип работы проводился

информационными специалистами в области научной коммуникации — оставались документы, фокусирующиеся скорее на практике, чем на теории. Статьи, анализирующие заявленные должностные позиции, были в основном ненадежны, так как не было возможности узнать, насколько успешным был поиск или как точно удовлетворена потребность и работа проведена в соответствии с описанием. Имея это в виду, мы разработали и применили следующие критерии включения для сужения результатов поиска:

1. Опубликованы - 2007 г. — по настоящее время;
2. Написаны на английском языке;
3. С особым указанием необходимых или рекомендованных компетенций/навыков для надлежащего кураторства или сопровождения данных;
4. Компетенции и навыки не были разработаны на основе анализа объявлений о работе или обзора литературы.

Анализовались следующие три документа, касающиеся компетенции: *Librarians' competencies profile for research data management* (2016) [33], *Matrix of digital curation knowledge and competencies* (2009) [34] и *Preparing the workforce for digital curation* (2015) [35]. Ниже в хронологическом порядке представим каждый из них на основе их даты публикации.

Matrix of digital curation knowledge and competencies была опубликована в 2009 г. и является одним из самых ранних документов, отразивших требуемые для работы в цифровом кураторстве навыки. Эти навыки были разработаны как часть проекта DigCCur под руководством Хелен Тиббо и Кал Ли из Университета шт. Северная Каролина в рамках проекта Института обслуживания музеев и библиотек (IMLS). Этот финансируемый данным институтом проект разработал анкеты для библиотечных школ и провел ряд обучающих мероприятий с целью поддержки диффузии деятельности по цифровому кураторству. Матрица проекта включала шесть направлений и была разработана как метод идентификации и организации материала по охвату анкет в цифровом кураторстве. Этими направлениями являются: 1) мандаты, ценности и принципы; 2) функции и навыки; 3) профессиональный, дисциплинарный, институциональный и организационный или культурный контекст; 4) тип ресурса; 5) необходимые знания; 6) точка перехода в информационный континуум. Направление «функции и навыки» было источником этого исследования, поскольку определяется как «ноу хау» в противоположность концептуальному, относительному или декларативному знанию.

В 2015 г. Национальный научно-исследовательский совет Национальной академии наук США выпустил отчет *Preparing the workforce for digital curation* [35]. Этот отчет вышел под авторством комитета по «Будущим возможностям карьеры и требованиям образования в сфере цифрового кураторства» и был создан для изучения вопросов, связанных с рабочей силой, в информационных организациях с акцентом на будущее экономическое развитие. Совет состоял из экспертов и представителей отрасли библиотекосведения и информатики, экономики труда и конкретных научных областей. Недавно, в 2016 г. Совместной комиссией по изучению компетенций библиотекарей в поддержку исследования в электронной сфере и научной коммуникации, собранной ARL, был опубликован *Librarians' competencies profile for research data management* [33]. Цель этой совместной комиссии — выделить компетенции, необходимые для исследования в

электронной сфере, управлении хранилищем и научной коммуникации. Совместная комиссия состояла из представителей организаций ARL, канадской ARL, конфедерации хранилищ открытого доступа и ассоциации европейских научных библиотек.

Сбор документов о компетенциях

Руководства относительно трех компетенций варьировались по длине и деталям содержания; два были документами, а один — матрицей навыков. Форматы также варьировались, один — документ формата HTML, а другой — PDF. Для нормализации этих документов в PDF, первый файл был переведен из HTML-документа в PDF. Далее, используя Adobe Acrobat Pro, документировались коды путем выделения релевантного текста и включения комментариев для каждой компетенции. Уровень деталей каждого документа по компетенции также варьировался. *Matrix of digital curation knowledge and competencies* [34] включала категории функции, такие как «доступ» и «администрирование», и функции подуровня, такие как «генерация пакета распространения информации», которые входят в большие детали о каждой полной категории. Для обеспечения надежности гарантии сравнения документов компетенции, использовались высокоуровневые описания и текст в разделе определения и объяснения каждого документа, чтобы провести качественный процесс систематизации.

Индуктивное кодирование

Контент-анализ является подходящей методологией для понимания и сокращения информации из существующих источников данных, таких как документ о компетенции. Поскольку нет ни одной таксономии качественного руководства компетенцией для цифрового кураторства, использовался индуктивный структурный подход кодирования, включающий идентификацию понятий и навыков, которые могут применяться в крупных сегментах текста и способствовать сравнительно частоты подсчета случаев. Индуктивное кодирование подразумевает, что кодированные навыки скорее отражают контент, найденный в документах-источниках, чем категории навыков, предопределенных исследователями или основанных на существующих таксономиях. Каждый документ был прочитан членом команды, и разработан список навыков. Для допуска сравнения и последующего шага кодирования документов были созданы широкие категории или массивы навыков. Например, категория исследовательских рабочих процессов включает такие навыки, как способность куратора понимать исследовательские практики, рабочие процессы и/или способность куратора понимать дисциплинарные нормы и стандарты. После разработки категорий документы компетенции затем повторно кодировались для каждой категории. Поскольку были упомянуты компетенции или навыки, исследователи повторно кодировали примеры для каждой категории.

Устойчивость компетенций как точек данных

Говоря об управлении, Джордан и Ллойд утверждают, что планирование человеческих ресурсов в библиотеках требует тщательного отбора и найма персонала. Одним из способов для этого служит формализация должностных позиций путем определения сфер рабочей ответственности, чтобы сделать работу предсказуе-

мой и регулярной. В случае с научной коммуникацией существование дисциплинарного консенсуса в отношении навыков, которыми должны владеть индивидуумы, предполагает наличие формализованных ролей, привносимых реструктуризацией. Организации не устанавливают профессии; тем не менее, они нанимают обладающих ими индивидуумов. Как уже отмечалось, организационная формализация ассоциируется с снижением темпа изменений, таким образом, в случае если библиотеки ARL создают должностные позиции в научной коммуникации, вероятнее всего, что работа по научной коммуникации будет поддерживаться в будущем.

РЕЗУЛЬТАТЫ

Число библиотекарей в сфере научной коммуникации в библиотеках Ассоциации научных библиотек (ARL)

Число библиотекарей в сфере научной коммуникации постоянно увеличивалось с 2012 г., когда впервые были собраны данные (см. рис.). Между 2012 - 2016 гг. позиции научной коммуникации с наименованиями должностей в *научной коммуникации* (SCHOLAR, т.е. акцент на открытом доступе и т.д.) и *цифровом кураторстве* (DIGICUR, т.е. акцент на управлении исследовательскими данными и т.д.) ежегодно росли в библиотеках ARL. Как уже указывалось, 2012 г. стал первым годом, когда ARL собрала данные по этим позициям; в 2012 г. данные SCHOLAR были объединены с другими кодами, учитывая небольшое число позиций, и, следовательно, не представлены на рис. Интересно, что, как показано ниже, рост в позициях цифрового кураторства составил в целом 27%. Рост научной коммуникации был более быстрым с увеличением в 49%. Делается вывод, что число библиотекарей в сфере научной коммуникации в библиотеках ARL возросло после первого сбора данных и что всего объединение обеих сфер — научная коммуникация и цифровое кураторство (по состоянию на 2016 г.) дало около 225 индивидуумов.

Идентификация компетенций и согласие внутри документов

Систематизация компетенций, рассматриваемая в этой статье, идентифицировала 23 категории для информационных специалистов, чтобы выполнять цифровое кураторство и/или работу по управлению исследовательскими данными (см. табл.). Из этих категорий 12 были найдены во всех трех публикациях о компетенциях. Этими двенадцатью навыками являются: поддержка и программа помощи; инструкция, сохранение данных, управление данными, отбор данных, платформы хранилища данных, связывание данных, аудит данных, лучшие практики по типу данных, лицензирование данных, организация данных и обнаружение информации. Семь дополнительных навыков обнаружены в двух документах о компетенции. Эти навыки — некоторый тип дисциплинарного знания; навыки анализа данных; понимание потребностей спонсора, знание исследовательских рабочих процессов, кураторство данных, обмен данными и навыки визуализации данных. Наконец, четыре навыка ограничены только одним документом о компетенции. Ими являются: безопасность данных, программирование и создание сценариев, базовое знание в науке, технике, инженерии или медицине и понимание потребностей издателя.

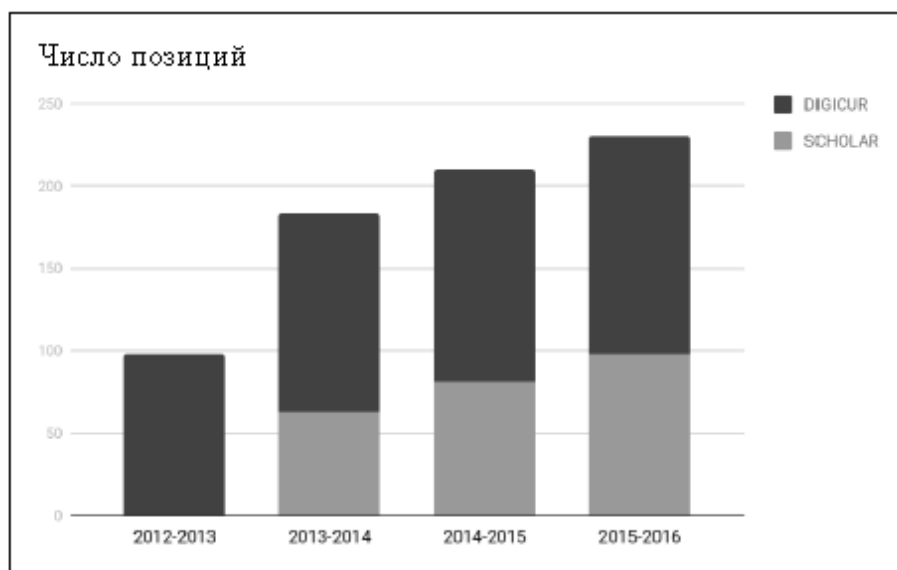


Рис. Позиции в научной коммуникации (SCHOLAR) и цифровом кураторстве (DIGICUR) в библиотеках ARL, 2012-2016 гг.

Таблица

Навыки, требуемые для цифрового кураторства по отобранным документам-руководствам

Умение или навык	Matrix of digital curation knowledge (2009)	Preparing the work-force for digital curation (2015)	Librarians' competencies profile for research data management (2016)
Поддержка и программа помощи	1	1	1
Аудит данных	1	1	1
Связывание данных	1	1	1
Управление данными	1	1	1
Организация данных	1	1	1
Сохранение данных	1	1	1
Платформы хранения данных	1	1	1
Отбор данных	1	1	1
Типы данных	1	1	1
Обнаружение информации	1	1	1
Инструкция	1	1	1
Лицензирование данных	1	1	1
Дисциплинарное знание	0	1	1
Анализ данных	0	1	1
Потребности спонсора	0	1	1
Исследовательские рабочие процессы	0	1	1
Цитирование данных	0	1	1
Обмен данными	0	1	1
Визуализация данных	1	1	0
Безопасность данных	0	1	0
Программирование и создание сценариев	0	1	0
Образование в науке, технике, инженерии и медицине	0	1	0
Потребности издателя	0	0	1

Как упоминалось выше, документы о компетенциях обладали высокой степенью согласия в выражении категорий навыков. Такое согласие заметно, так как все три документа были написаны за последнее десятилетие отдельными группами с отличной дисциплинарной квалификацией. Что касается согласия между документами, наиболее схожими являются *Librarians' competencies profile for research data management* [34] и *Preparing the workforce for digital curation* [35] с долей согласия свыше 78% (18 из 23 навыков). Документы *Matrix of digital curation* и *Preparing the workforce for digital curation* имели следующую высокую степень согласия с долей в 61% или 14 из 23 точек совпадения.

ОБСУЖДЕНИЕ

Чтобы изучить вопрос относительно устойчивости, рассмотрим наши результаты в целом. На основе собранных ARL данных видно, что информационные специалисты в области научной коммуникации, особенно специалисты широкого профиля, названные библиотекарями сферы научной коммуникации, начали действовать в академических библиотеках в 2010-х гг. Восходящая тенденция роста этой должностной позиции заставила нас поверить в то, что этот вид информационных профессий востребован, а также и ведущаяся работа – и то, и другое необходимо и ценно. Консенсус в компетенциях предполагает, что информационные профессии находятся в согласии относительно того, что должно быть сделано в поддержку исследователей, так как они обращаются за помощью в сферы цифрового кураторства и управления исследовательскими данными. Это подразумевает, что роль информационного специалиста в области научной коммуникации, по крайней мере в рамках его работы, становится формализованной. Мы считаем формализацию работы в сфере цифрового кураторства показателем зрелости области, поскольку она фокусируется вокруг установленных компетенций.

На основе обзора литературы и результатов изучения периода между 2009 - 2015 гг. считается экспериментальным в формировании и становлении работы в научной коммуникации благодаря усилиям информационных специалистов. В 2009 г. семинар ACRL под названием the ACRL Scholarly Communication Roadshow начал обучать библиотекарей и других информационных специалистов научной коммуникации во всех ее аспектах. К 2012 г. ARL начала сбор данных о позициях с названием библиотекарь сферы научная коммуникация, но обнаружила, что у них недостаточно данных, чтобы дать право сообщать о них отдельно от других позиций. Было разрешено сформировать расхождение в классификациях массива данных между научной коммуникацией и цифровым кураторством, несмотря на охват в проведенной работе. В 2012 г. респонденты опроса COPPUL подтвердили данные, указав на научную коммуникацию как на область специализации, которая еще не объединена [8]. Однако к 2013 г. ARL собрала данные о библиотекарях сферы научной коммуникации по более, чем 50 позициям в 125 библиотеках ARL. В 2015-2016 гг. были опубликованы два отдельных документа, они показали 78% согласия в условиях необходимых навыков и компетенций для подобласти цифрового кураторства [33, 35]. К 2017 г. организаторы семинара ACRL Roadshow обнаружили, что знания аудитории улучшились относительно того, где

необходимо преобразование их резюме [5], подразумевая, что уровень формализации произошел в информационных профессиях и информационных организациях. На основе этих результатов ответом на вопрос RQ2 является появление цифрового кураторства как связующей подобласти практики.

Дальнейшее свидетельство появляющейся связанности практики в подобласти цифрового кураторства (и управления исследовательскими данными) рождается из результатов, связанных с тремя рассматриваемыми нами документами о компетенциях. Эти документы опубликованы с разных сторон того, что в литературе указывается как формативный период области научной коммуникации внутри информационных организаций; они также были опубликованы до и после учреждения области как формального названия, полученного от ARL. Документ *Matrix of digital curation knowledge* [34] является самым отличающимся из трех изучаемых документов о компетенции, подразумевая, что после его публикации подобласть цифрового кураторства объединилась. Компетенции, установленные в 2015 г. (документ [35]) и 2016 г. (документ [33]), отражают зрелость понимания этой области информационной работы в научной коммуникации. Все, кроме одной компетенции в документе 2009 г., включены в более поздние документы. Более того, 6 из 23 компетенций, фигурирующих в документах 2015 г. и 2016 г., являются уникальными для более новых публикаций, идентифицированных и добавленных после 2009 г. Это свидетельство интерпретируется нами как доказательство зрелости этой подобласти практики внутри информационных организаций ARL и, как де факто, формализации соответствующей работы.

Диффузия инноваций (DOI) и будущее информационных специалистов в научной коммуникации

На основе результатов и понятий реструктуризации и формализации из DOI и литературы по организации инноваций предсказывается рост формализации внутри информационных организаций, способствующий замедлению изменений для информационных специалистов в научной коммуникации. Следовательно, для ответа на вопрос RQ1 исследования прогнозируем сохранение устойчивости работы в научных коммуникациях в будущем. Персонал будет вводиться в работу в библиотеках ARL, и все большее увеличение работы, по крайней мере, в цифровом кураторстве и управлении исследовательскими данными, постепенно будет систематизироваться в терминах компетенций и навыков, необходимых практикам. Работа в областях научной коммуникации все больше становится предсказуемой и регулярной для поддержки планирования людских ресурсов, как отражает консенсус о профессиональных навыках и компетенциях. Предположительно эти позиции рассеивались и рассеиваются по информационным профессиям, как подразумевает успех семинара ACRL Roadshow. Связанность документов о компетенциях и числа должностных позиций внутри информационных организаций ARL также предполагает, что большинство информационных организаций ARL находятся или вышли за рамки этапа реструктуризации процесса принятия решения об инновации Роджерса [6].

Исследование ограничений и будущая работа

Будущее информационных организаций ARL, предоставляющих ученым поддержку в научной коммуникации, кажется, должно быть стабильным, но для этого исследования имеются ограничения. Собранные нами данные позволили рассмотреть только два понятия из DOI и литературы по организации инноваций. Показатели, такие как данные ARL и связанность документов о профессиональных компетенциях, предполагают, что работа научной коммуникации вступает в последние этапы процесса принятия решения об инновации Роджерса, но отдельные решения о принятии внутри организаций нами не изучаются.

Говоря о процессе, с помощью которого организации «выбирают» — принять или отклонить инновацию, Роджерс [6] отмечает, насколько этот процесс не линеен; в любой точке этапы этой модели могут быть пропущены, остаться незаконченными или начаться снова, но позже [6, с.421]. Без данных для проверки модели Роджерса нам неизвестно, кто принимает решение о приеме на работу информационных специалистов по научной коммуникации в информационные организации ARL при условии, что должностные позиции подобны описанным в документах о компетенции и что персонал общается с заинтересованными лицами с целью продемонстрировать свою значимость. Создание новых позиций подразумевает, что работа в научной коммуникации устойчива в той степени, в какой она формализована, но ограничения, связанные с нашими данными, допускают, что долговременная перспектива неопределенна. Для изучения неопределенности будущего исследование должно рассмотреть, как создаются должностные позиции в научной коммуникации, и оценить их важность, выражаемую в устойчивости работы, проводимой информационными специалистами в научной коммуникации. Одним из способов сбора данных может быть интервью с лицами, отвечающими за принятие решений по созданию позиций в библиотеках ARL, а также с учеными, пользующимися услугами научной коммуникации.

Другое ограничение, связанное с этой статьей, распространяется на особенности высокого уровня, касающиеся организации инноваций. В данной статье много сказано о роли формализации как предсказывающего фактора, но работа, реально проводимая библиотеками внутри ARL, не анализировалась. Поэтому возможно, что должностные позиции были менее формальными, чем предполагают документы о компетенциях.

Кратко ранее обсуждалось, что переменные, которые в этом исследовании нами не изучались, включают: централизацию, сложность, взаимосвязанность, слабость организации и размер. Так как другие информационные организации стремятся создавать должностные позиции, то может потребоваться централизованное принятие решений от университетских администраторов. Нехватка доступных ресурсов также способна помешать заполнению позиций. Финансовая сторона многих государственных университетов нестабильна, и сегодня 46 штатов США тратят на одного студента меньше, чем десять лет назад [36]. Размер организации также соотносится с инновацией, и поскольку должностные позиции в сфере научной коммуникации рассенваются по информационным профессиям, вероятно, что небольшие или средние университеты не будут иметь возможность

обеспечивать такой же уровень поддержки, что и более крупные университеты. Автор [37] сообщает, что в 2015 г. типичная библиотека ARL наняла 84 специалиста, что гораздо больше, чем нанимали библиотеки колледжа свободных искусств и педагогических университетов. Чтобы учитывать такие влияния, как эти, будущая работа должна сосредоточиться на сборе данных, обеспечивающих более широкую точку зрения на стабильность должностной позиции в научной коммуникации. В дополнение к этому, поскольку размер организации является заслуживающим внимания фактором, то дальнейшая работа должна смотреть на иные, чем библиотеки ARL, информационные организации, так как эти библиотеки имеют тенденцию к укрупнению.

ЗАКЛЮЧЕНИЕ

Наши результаты показали, что работа информационных специалистов в сфере научной коммуникации формализовалась с 2009 г. Используя два понятия из DOI и литературы по организации инноваций, несмотря на ограничения этого исследования, ссылаемся на то, что имеется ценность в работе, выполняемой этими специалистами, и приходим к выводу, что данная область появляется в период своего уточнения (пользуясь терминологией Роджерса [6]), что говорит о ее устойчивости в библиотеках ARL. Учитывая связанность практики в подобласти цифрового кураторства, научная коммуникация будет вероятнее всего оставаться устойчивой сферой практики в информационных профессиях. Тем не менее, остается много работы. Собранных данных в этом исследовании недостаточно, чтобы проверить другие понятия из DOI, это вынуждает нас сделать вывод, что хотя настоящее стабильно, долговременная перспектива гораздо менее предсказуема. Дальнейшее исследование должно стремиться искать способы выхода за рамки этой попытки и изучать появление работы в сфере научных коммуникаций внутри отдельных организаций, это важно потому, что такая работа заставит ученых и их потенциал создавать устойчивую и долговременную возможность роста для всех типов информационных организаций.

Благодарность. Авторы признательны Ассоциации научных библиотек (ARL) за предоставление доступа к своим данным.

ЛИТЕРАТУРА

1. Regazzi J. J. Scholarly communications: A history from content as king to content as kingmaker. — New York: Rowman & Littlefield, 2015.
2. Borgman C. L. Scholarly communication and bibliometrics. — Newbury Park: Sage Publications, 1990.
3. Kowalski M. Breaking down silo walls: Successful collaboration across library departments// Library Leadership & Management. — 2017. — Vol. 31, No. 2. — P. 1-15.
4. Ketchum A. M. The research life cycle and the health sciences librarian: Responding to change in scholarly communication// Journal of the Medical Library Association. — 2017. — Vol. 105, No. 1. — P. 80-83. — doi:10.5195/jmla.2017.110
5. Cross W., Olen J., Perry A. Jump start your scholarly communication initiatives: Lessons learned from redesigning the scholarly communications roadshow for a new generation of librarians// At the Helm: Leading Transformation, the Proceedings of the ACRL 2017 Conference. 361-371. —

Chicago: Association of College and Research Libraries, a division of the American Library Association, 2017. — <http://www.ala.org/acrl/sites/ala.org/acrl/files/content/conferences/confsandpreconfs/2017/JumpStartYourScholarlyCommunicationInitiatives.pdf>

6. Rogers E. M. *Diffusion of innovations* (5th ed.). — New York: Free Press, 2003.

7. Brundtland G. H. Report of the World Commission on environment and development: Our common future. — United Nations, 1987.

8. Swanepoel M., Kehoe L., Hohner M., Shepstone C., Vanderjagt L., Wakaruk A., [...] Winter C. Developing a community of practice: Report on a survey to determine the scholarly communication landscape in western Canada. — Unpublished manuscript, 2015. — <http://hdl.handle.net/10133/4784b>

9. Brantley S., Bruns T. A., Duffin K. I. Librarians in transition: Scholarly communication support as a developing core competency// *Journal of Electronic Resources Librarianship*. — 2017. — Vol. 29, No. 3. — P. 137-150. — doi:10.1080/1941126X.2017.1340718

10. Hey T., Tansley S., Tolle K. The fourth paradigm: Data-intensive scientific discovery. — Redmond, Washington: Microsoft Research, 2009.

11. Borgman C. L. Digital libraries and the continuum of scholarly communication// *Journal of Documentation*. — 2000. — Vol. 56, No. 4. — P. 412-430.

12. Bossaller J. S., Moulaison Sandy H. Documenting the conversation: A systematic review of library discovery layers// *College & Research Libraries*. — 2017. — Vol. 78, No. 5. — P. 602-619. — doi:10.5860/crl.78.5.602

13. Novak J., Day A. The libraries they are a-changin': How libraries reorganize// *College & Undergraduate Libraries*. — 2015. — Vol. 22, No. (3-4). — P. 358-373. — doi: 10.1080/10691316.2015.1067663

14. Fearon D., Gunia B., Lake S., Pralle B., Sallans A. Research in Data Management Services (SPEC KIT 334). — Association of Research Libraries: Washington, D.C., 2013. — <https://doi.org/10.29242/spec.334>

15. Hayes C., Kelly H. E. Who's talking about scholarly communication? An examination of gender and behavior on the SCHOLCOMM Listserv// *Journal of Librarianship and Scholarly Communication*. — 2017. — Vol. 5, No. 1. — doi:10.7710/2162-3309.201700

16. Klain-Gabbay L., Shobamb S. Scholarly communication and academic librarians// *Library & Information Science Research*. — 2016. — Vol. 28, No. 2. — P. 170-179.

17. Potvin S. The principle and the pragmatist: On conflict and coalescence for librarian engagement with open access initiatives// *The Journal of Academic Librarianship*. — 2013. — Vol. 39, No. 1. — P. 67-75.

18. Rodriguez J. E. Scholarly communications competencies: Open access training for librarians// *New Library World*. — 2015. — Vol. 116, No. (7/8). — P. 397-405. — doi:10.1108/NLW-12-2014-0140

19. Maron N. L., Smith K. K. Current models of digital scholarly communication: Results of an investigation conducted by Ithaka for the Association of Research Libraries// *The Journal of Electronic Publishing*. — 2009. — Vol. 12, No. 1. — doi:10.3998/3336451.0012.105

20. Lewis R., Sarli C., Suiter A. Scholarly output assessment activities (SPEC Kit 346). — Association of Research Libraries, Washington, D.C., 2015. — <https://doi.org/10.29242/spec.346>

21. Green E. What are the most-cited publications in the social sciences (according to Google Scholar)? — [blog post, 2016, May 12]. — <http://blogs.lse.ac.uk/impactofsocialsciences/2016/05/12/what-are-the-most-cited-publications-in-the-social-sciences-according-to-google-scholar/>

22. Ryan B., Gross N. C. The diffusion of hybrid seed corn in two Iowa communities// *Rural Sociology*. — 1943. — Vol. 8, No. 1. — P. 15-24.

23. Rogers E. M. *Diffusion of innovations*. — New York: Free Press of Glencoe, 1962.

24. Valente T. W., Rogers E. M. The origins and development of the diffusion of innovations paradigm as an example of scientific growth// *Science Communication*. — 1995. — Vol. 16, No. 3. — P. 242-273. — <http://doi.org/10.1177/1075547095016003002>

25. Oxford. Organization. — (n.d.). — <https://en.oxford-dictionaries.com/definition/us/organization>

26. Pugh D. S. *Organization theory: Selected readings*. — Harmondsworth: Penguin, 1990.

27. Hage J., Aiken M. Program change and organizational properties a comparative analysis// *American Journal of Sociology*. — 1967. — Vol. 72, No. 5. — P. 503-519.

28. Chen S. T., Chang B. G. The effects of absorptive capacity and decision speed on organizational innovation: A study of organizational structure as an antecedent variable// *Contemporary Management Research*. — 2012. — Vol. 8, No.1. — P. 27-50.

29. Kyrillidou M., Morris S., eds. ARL Annual Salary Survey 2012-2013. — Washington, DC: Association of Research Libraries, 2013. — <https://doi.org/10.29242/salary.2012-2013>

30. Kyrillidou M., Morris S., eds. ARL Annual Salary Survey 2013-2014. — Washington, DC: Association of Research Libraries, 2014. — <https://doi.org/10.29242/salary.2013-2014>.

31. Kyrillidou M., Morris S., eds. ARL Annual Salary Survey 2014-2015. — Washington, DC: Association of Research Libraries, 2015. — <https://doi.org/10.29242/salary.2014-2015>

32. Morris S. ARL Annual Salary Survey 2015-2016. Washington, DC: Association of Research Libraries, 2017. — <https://doi.org/10.29242/salary.2015-2016>

33. Calarco P., Shearer K., Schmidt B., Tate D. Librarians' competencies profile for scholarly communication and open access. — 2016. — [Confederation of Open Access Repositories website]. — <https://www.coar-repositories.org/files/Competencies-for-ScholComm-Open-Access-June-2016.pdf>

34. Lee C. Matrix of digital curation knowledge and competencies. — 2009. — <https://ils.unc.edu/digccurr/digccurr-matrix.html>

35. National Research Council. Preparing the workforce for digital curation. — Washington, DC: The National Academies Press, 2015. — <https://doi.org/10.17226/18590>

36. Mitchell M., Leachman M., Maserson K. Funding down, tuition up state cuts to higher education threaten quality and affordability at public colleges. — Washington, D.C.: Center on Budget and Policy Priorities, 2016. — <https://www.cbpp.org/sites/default/files/atoms/files/5-19-16sfpp.pdf>

37. Wilder S. Hiring and staffing trends in ARL libraries. — 2017. — <http://www.arl.org/storage/documents/publications/rli-2017stanley-wilder-article2.pdf>

38. Howard H. A. Organizational structure and innovation in academic libraries// *College & Research Libraries*. — 1981. — Vol. 42, No. 5. — P. 425-34.

39. Montana P. J., Charnov B. H. *Management*. — Hauppauge, N.Y: Barron's, 1993.

Аналитический подход к информационным стандартам на основе ценностей*

Брайан ДОБРЕСКИ
(Brian DOBRESKI)

Школа информационных наук,
Университет г. Сиракузы, г. Сиракузы,
США

Хотя разработки в сфере информационных технологий включают удивительные новые возможности, они (технологии) в процессе своего развития могут ненамеренно нарушить такие важные ценности, как автономия или конфиденциальность. Соответствующие этические подходы к технологиям гарантируют использование важных перспектив в оценке наших технологических практик и артефактов. Основными в такого рода артефактах являются информационные стандарты, влиятельные документы, представляющие и сохраняющие соглашения об идеальной практике между сообществами. При тщательном изучении стандартов ценности олицетворяют многообещающие концептуальные линзы. В статье выдвигается идея использования анализа ценностей в информационных стандартах, в частности тех, которые относятся к организации знания. Анализ ценностей является методологическим подходом, регулирующим извлечение ценностей из артефактов. Два ситуационных исследования демонстрируют применение анализа ценностей к стандартам в организации знания и представляют свои результаты, отражающие то, как ценности и их применение можно увидеть в этих информационных артефактах. Затем анализ ценностей можно распространить за рамки стандартов в качестве документов в целях изучения более широкой экосистемы, внутри которой располагаются и работают информационные стандарты, чтобы полностью понять их применение, поддержать важные ценности сообществ и сохранить этические информационные практики.

ВВЕДЕНИЕ

Хотя разработки в сфере информационных технологий содержат вызывающие интерес новые возможности, они могут в ходе своего развития нечаянно нарушить наши ценности. Создание, сбор и использование данных постоянно осложняется, становится трудным во всех сферах деятельности и представляет усиливающую угрозу автономии и конфиденциальности пользователя. В целях поддержки соответствующих этических подходов к информационным технологиям оправданным служат новые важные исследования ряда техно-

логических практик и артефактов. Основными среди таких артефактов являются информационные стандарты, документы, представляющие и хранящие соглашения об идеальной практике между сообществами и управляющие созданием и представлением данных. Стандарты являются, по-видимому, нейтральными элементами инфраструктуры, но играют ключевую роль в объединении и сохранении перспектив сообщества [1]. В последующих исследованиях роли стандартов в поддержке или разрушении этической информационной практики анализ ценностей представляет многообещающую оценочную призму.

Теория ценностей концептуализирует их как глубоко поддерживаемые убеждения об определенных способах бытия или действия [2]. Проще говоря, ценностями являются любые вещи, которые люди или организации

* Перевод Dobreski B. Toward a value-analytic approach to information standards// Proceedings of 81 st Annual Meeting ASIS&T. — 2018. — P. 114-122 —<https://www.asist.org/wp-content/uploads/Final-81st-Annual-Meeting-Proceedings-1.pdf>

считают важными [3], хотя они, как правило, объясняют абстрактные понятия, такие как эффективность или уважение. Теория ценностей и исследование этой сферы давно применялись к дисциплинам, особенно к общественным наукам. В рамках информатики как науки об информации ценности являются неотъемлемыми частями нескольких направлений исследований, включая профессиональную этику [4, 5] и дизайн [6, 7]. Хотя опросы и интервью являются мощными методологическими средствами в изучении ценностей во всех структурах, извлекаемые ценности из письменных документов и иных коммуникативных артефактов дали толчок появлению другой методологической традиции, известной как анализ ценностей. В информатике анализ ценностей применялся к таким артефактам, как алгоритмы [8] и твиты [9], хотя до сих пор не применялся к информационным стандартам.

Стандарты — это документы, кодирующие и излагающие спецификации или протоколы и согласованные и распространяемые внутри сообщества [10]. Они призваны поощрять коллективную деятельность людей, но часто принимаются как должное в процессе, приводящем к сложностям в распознавании их роли и действий [11, 12]. Тем не менее, подобно всем социальным продуктам, стандарты формируются под воздействием точек зрения тех, кто их создает. Информационные стандарты следят за созданием и управлением данными и информационными системами, облегчая доступ и обмен информацией. В установившейся идеальной практике информационные стандарты должны поддерживать определенные точки зрения за счет других [13] и могут таким образом считаться артефактами, основанными на ценностях. Какие ценности выражают эти стандарты и поддерживают ли они или не оправдывают существующие в сообществе ценности, остается открытым вопросом.

В данной статье выдвигается основанный на анализе ценностей подход к исследованию информационных стандартов как основных технологических артефактов. После представления соответствующей проблемы, касающейся ценностей в информатике и изучения стандартов, будут изложены два ситуационных исследования. В этих исследованиях применяется анализ ценностей, чтобы продемонстрировать, как можно считывать ценности из информационных стандартов и использовать для этого данные, и как стандарты и данные объединяют и усиливают систему ценностей. Затем обсуждается наследуемая, основанная на ценностях и этическая природа стандартов, а также применение их как несущих ценности артефактов, проявляющихся через практику. Делается вывод, что ценности представляют полезный механизм оценки для анализа и разработки информационных стандартов, чтобы поддерживать этическую технологическую практику и наметить план дальнейшего исследования в этой области.

ПРОБЛЕМА И ПОДХОД

Исследование ценностей берет свое начало в философии, где оно известно как теория ценностей, относясь скорее не к одной формальной теории, а к дисциплине исследования [2]. В соответствии с теорией ценности могут принимать множество форм, но, как правило, изображают абстрактные понятия, такие как счастье или эффективность; в частности, ценности справедливости

и доброты более тесно связаны с этикой [14]. С тех пор исследование ценностей распространилось в другие сферы, особенно на общественные науки, в которых оно использовалось в целях изучения мотивации [15], поведения [16] и социальной организации [17]. Ценности в течение длительного периода времени были ключевым понятием информатики, с более пристальным вниманием к базовым ценностям, связанным с данной областью и ее участниками, например, доступу, интеллектуальной свободе и конфиденциальности [18; 4, 5]. Вне этого рефлексивного акцента на ценностях исследование в информатике также изучало взаимодействие между ценностями и дизайном технологий. Одним из подходов понимания этой взаимосвязи является ценностный дизайн, первоначально усовершенствованный Фридманом [19] как метод оценки того, как ряд эвристических ценностей, таких как применимость, автономия, свобода из предубеждения вовлечены в дизайн систем. В отдельной, но близкой области исследования, известной как ценности в дизайне, фокус смещен на более широкий ряд ценностей, кроме эвристических, моральных ценностей ценностного дизайна, и стремится отойти от такого рода ретроспективного анализа [20]. Недавно ценности и дизайн закрепились как общий термин, относящийся к исследованию как ценностного дизайна, так и ценностей в дизайне [7].

Во всех областях традиция исследования ценностей на основе контент-анализа ассоциируется с артефактами и относится к анализу ценностей [21]. Ранние сторонники анализа ценностей пользовались предопределенным подходом к ценностям при кодировании текстовых данных с тем, чтобы выявить важные, внедренные ценности; подобные подходы были, как правило, качественными [21,22]. Поздние подходы к анализу ценностей больше объединяли индуктивные и количественные методы и выходили за рамки текстовых документов, чтобы изучать другие артефакты [23]. В рамках общественных наук анализ ценностей применялся к разнообразию материалов, включая учебники [24], новеллы [25] и комиксы [26]. Безотносительно к определенному анализируемому типу контента анализ ценностей направлен на выявление предпочтений и преимуществ, внедренных в ряд материалов, и разграничение их по системе ценностей. В информатике анализ ценностей был востребован в рамках анализа организационных кодов этики, включая коды библиотечных ассоциаций [27] и архивных организаций [28]. В последнее время автоматизированные и краудсорсинговые подходы к анализу ценностей применялись к социальным медиа в целях раскрытия ценностей пользователя [9]. Хотя исследование ценностей и дизайна применяло анализ ценностей к артефактам технологий, таким как алгоритмы [8], исследование в информатике не использовало данный подход к базовому типу артефактов технологий в этой сфере — стандартам.

Информационные стандарты служат технологической инфраструктурой, регламентирующей процедуры и продукты и стимулирующей совместную работу. Особо сильный акцент имелся на стандартизации в рамках информационной сферы [29], затрагивая, в частности, организацию знания, представление знания или информации в различных формах организации систем [30]. Эти стандарты чрезвычайно востребованы и регламентируют создание и применение таких технологий,

как классификации, онтологии, базы данных, а также связанные с ними данные. Хотя широко принятые информационные стандарты стремятся сохранять нейтралитет [31], они фактически оказываются внедренными в определенные перспективы за счет других [1;13]. Критический анализ информационных стандартов по терминологии, в частности, вскрыл ряд встроенных предубеждений [32-34]. Хотя данный ряд исследований придерживается основных применений ценностей в информационных стандартах, ценности и анализ ценностей относительно этих артефактов формально еще не рассматривались [35]. Внимание к ценностям во всех информационных стандартах важно, но стандарты по организации знания представляют особо интересную часть исследования из-за способности генерировать данные, которые потом могут сохранять внедренные ценности.

Ниже приводятся два ситуационных анализа в целях иллюстрации возможностей анализа ценностей как подхода оценки к стандартам организации знания и данных. Первый ситуационный анализ рассматривает текст некоторых стандартов культурного наследия, фокусируясь на фрагментах, затрагивающих варианты имен людей, и затем прослеживая этот элемент данных на представлениях в сети. Второй ситуационный анализ относится к изучению библиографической записи, созданной по библиотечному стандарту *Описания ресурса и доступа*, помещая ее в контекст утвержденных целей этого стандарта и сопоставляя с данными Google Books. Оба случая демонстрируют пользу анализа ценностей в изучении важных политик, внедренных в стандарты организации знания и данных. Дальнейшие применения анализируются на основе данных примеров.

ВАРИАНТЫ ИМЕН ЛЮДЕЙ

Культурное наследие представляет собой исследование, сохранение и курирование культурных практик и артефактов, передаваемых через социальные группы [36]. Учитывая сложный и распространенный характер работы, связанной с культурным наследием, стандарты играют важную роль в координации деятельности в этой сфере, особый акцент делается на стандартах организации знания. Библиотеки, архивы и музеи тратят массу усилий на создание, применение и поддержание систем организации знания. Основным кругом интересов стандартов в культурном наследии является описание ресурса, процесс создания презентаций из артефактов и средств [37]. Обычно библиотеки, архивы и музеи поддерживают свои собственные определенные стандарты описания ресурсов. В архивном деле – Описание архивных материалов: стандарт содержания (DACS - Describing Archives: A Content Standard) [38] служит де-факто стандартом описания, выделяющим структурные и относительные особенности архивных материалов и их создателей. В музеях каталогизация культурных объектов (CCO - Cataloging Cultural Objects) [39] является основным стандартом содержания по описанию произведений искусства и иных артефактов и часто применяется вместе с структурными стандартами, включая Категории для описания произведений искусства (CDWA - Categories for the Description of Works of Art) и базовые категории Ассоциации визуальных ресурсов (VRA core). В библиотеках стандарты описания, в целом относящиеся к библиографическим стандартам, регулируют создание каталожных записей библиотечных ресурсов и

связанных с ними людей и организаций. Последние 40 лет значительное число библиотек, применяющих английский язык, пользуются американскими правилами каталогизации (AACR 2 - Anglo-American Cataloging Rules, 2 nd edition) в качестве стандарта описания. Хотя многие библиотеки работают по этим правилам, 2010 г. стал свидетелем появления публикации последующей редакции AACR 2, сфокусированной в международном плане на описании ресурса и доступа (RDA - Resource Description and Access), которое сегодня все больше используется библиотеками по всему миру.

Несмотря на разнообразие регламента установок и применений, стандарты описания в сфере культурного наследия предписывают и определяют ряд схожих элементов данных, в частности для людей [40]. Например, личные имена, названия и даты являются общими для всех стандартов культурного наследия. Каждый из этих стандартов также отличается предпочтениями личных имен и их вариантами. Хотя предпочтительные имена относятся к «официальным», используемым в обращениях за помощью или поиске каталожной записи, варианты имен представляют любые другие формы значимости имен. Например, хотя «Джордж Элиот» может быть предпочтительным именем, «Мэриам Эванс Кросс» может быть записано как основной вариант имени этого автора.

С учетом сходства в концепции и возможностях инструкций по вариантам имен в AACR2, CCO, DACS и RDA предоставляют перспективу опробовать различия в соответствующих системах ценностей. Хотя эти четыре стандарта отображают аналогичные понятия вариантов имен, отображают ли они по-прежнему их при описании одних и тех же ценностей? Краткий анализ ценностей первоначальных инструкций соответствующих фрагментов (табл. 1) проливает на это свет.

На первый взгляд каждый из стандартов, судя по всему, придает значение идеи использования: использование делает значимым конкретный вариант имени и оправдывает его включение в запись, касающуюся человека. Использование можно рассматривать как общую ценность. Однако более тщательная проверка этих инструкций выявляет важные изменения, особенно при определении того, чье использование имеет ценность. Для CCO использование ограничено официально опубликованными источниками. AACR2 относит нечто подобное к использованию в библиографических источниках, но также включает использование описываемым лицом. DACS переводит фокус на использование подразумеваемым пользователем, при этом варианты имен фиксируют любые разумные альтернативы, в которых они могут исходить. Наконец, RDA придает значение двум конкретным пользовательским задачам: идентификации человека и доступу к данным о человеке. Варианты имен, которые, по мнению каталогизатора, поддерживают эти случаи использования, должны быть записаны.

Можно увидеть, что хотя идея вариантов имени является одинаковой, связанные с ним ценности в четырех стандартах различны. Разные виды использования представляют ценность и таким образом оправдывают или нет значимость имени. Ценность присваивается официально опубликованным источникам, потенциальным пользователям или описываемому лицу, таким образом зафиксированные варианты имен лица могут существенно отличаться в четырех стандартах. Придание ценности или обеспечение собственного использования имени лица имеет дальнейшие применения в автономии, а также самиден-

тификации. Даже в рамках относительно небольшой области описания ресурсов культурного наследия подобные идеи несут разные ценности в разных стандартах с непосредственным применением к данным.

Выходя за рамки области культурного наследия в более широкую плоскость общей сети можно заметить даже большие варианты в практиках стандартизации имен. Например, руководство относительно стиля (Wikipedia Manual of Style – MoS) Википедии дает руководящие указания к библиографическим статьям с целью поддержки логичности и простоты использования,

что можно интерпретировать как официальный ряд ценностей [41]. Функционально MoS устанавливает стандарт для информации, включаемой в статьи, представляющие людей. Особые правила устанавливаются относительно личных имен, с более известной формой имени, служащей названием статьи. Помимо этого, регламентируются типы вариантов имен и порядок их записи, включая самую полную форму имени, изменения имени, если это имело место, другие предшествующие имена, а также любые псевдонимы, уменьшительные имена или прозвища.

Таблица 1

Соответствующие фрагменты, касающиеся вариантов имен

Стандарт	Фрагмент	Инструкция
AACR2	26.2A1	«Ссылка на имя, используемое лицом или найденное в библиографических источниках, отличное от имени, используемого в заголовке для данного лица»
CCO	A1.2.1.2.2	«Включение альтернативных имен и вариантов имен, отображаемых в опубликованных источниках и представляющих существенные различия в форме или орфографии»
DACS	10.3	«Создание ссылки на форму имени частного или корпоративного лица или названия произведения, которая может обоснованно запрашиваться, на форму, которая была выбрана в качестве имени или единообразного названия или в качестве описания под заголовком»
RDA	9.2.3.3	«Запись альтернативного имени для лица, считающегося важной для идентификации или доступа, с применением общих руководящих принципов в отношении регистрации имен в пункте 8.5»



Рис. 1. Карточка Википедии для Джейка Зайруса

LC control no.:	no2009052666
LCCN Permalink:	https://lccn.loc.gov/no2009052666
Personal name heading:	Charice, 1992-
Variant(s):	Pempengco, Charice, 1992- Relucio Pempengco, Charmaine Clarice, 1992-
Found in:	Hit man [SR] p2008: container (Charice) Charice Pempengco WWW site, Apr. 3, 2009 (Charice Pempengco; Charice; b. Charmaine Clarice Relucio Pempengco, May 10, 1992, San Pedro, Laguna province, Philippines)

Рис. 2. Запись RDA для Чариса

В соответствии с этими правилами, логичность явно является важной ценностью и может служить инструментом поддержки простоты использования. Варианты имен подразделяются на общие типы и предписанные общими правилами обработки во всех статьях, касающихся данных лиц. Использование, особенно наиболее общее, представляется важным в определении названия статьи по аналогии с предпочтительным именем в описании культурного наследия. Вне этого варианты имен применяются через установление использования, хотя снова вопрос о том, чье использование важнее, не теряется и здесь актуальности.

Статья Википедии, посвященная Джейку Зайрусу, приводит список данных при рождении имен и имен более ранних этапов их смены певцом до его изменения половой принадлежности (рис. 1). Ссылки на варианты имен в этой статье относятся к недавним онлайн-статьям. В оправдании вариантов имен в Википедии все источники в сети содержат область использования. Однако, консультируясь с файлом имен Библиотеки конгресса, можно увидеть, что в записи полномочий RDA все еще указан певец с предпочтительным именем Чарис и варианты имен, данные ему при рождении, без упоминания Джейка Зайруса (рис. 2). Хотя такая схема может препятствовать идентификации и доступу, ценными RDA, она также отражает отличающуюся концепцию использования. Использование здесь отражает библиотечный домен, и поскольку певец выпустил только опубликованные записи до принятия имени Zyugus, это имя в настоящее время выходит за рамки круга материалов, имеющих отношение к использованию. Помимо этого, представления Zyugus в Википедии и RDA потенциально могут рассматриваться как нарушающие основные общечеловеческие ценности, такие как конфиденциальность и самоидентификация.

Хотя и культурное наследие, и более широкая сетевая среда предписывают варианты использования имен, разные концепции, чье использование важнее, могут привести к коренным различиям в способах именования людей. Эти отличия могут содержать иные применения ценностей и могут в значительной степени воздействовать на понимание и доступ пользователей.

БИБЛИОГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ

Подобно всем стандартам стандарты организации знания служат приведению практики к согласованности [42]. В рамках библиотек библиографические стандарты призваны регулировать практики, поддерживающие создание каталожных данных. Вышеупомянутые стандарты AACR2 и RDA являются новейшими в явно линейной последовательности библиографических стандартов, которые ускорили универсальность каталожных записей за последний век. Как имеющийся *де факто* стандарт описания для библиотек RDA регулирует создание данных в библиотечных каталогах, эти данные также все больше и больше распределяются по миру, за пределами традиционного каталога. RDA устанавливает современный стандарт для "достойного" библиографического представления.

Несмотря на то, что ценности стандартов конкретных организаций, занимающихся вопросами знаний, еще предстоит в полной мере изучить, RDA свидетельствует о своих заявленных ценностях путем включения четкого положения о целях и принципах. Основные цели включают удовлетворение потребностей пользователей, экономическую эффективность и гибкость, в то

время как принципы подчеркивают дифференциацию, достаточность, точное представление, единообразие и взаимоотношения [43]. В совокупности эти цели и принципы могут служить в качестве подтвержденной системы ценностей, хотя до сих пор не проведена оценка того, насколько хорошо текст RDA поддерживает и отражает эти ценности. Однако еще одним критическим доказательством здесь являются получившиеся библиографические записи RDA. Каждая запись компилирует данные, предписанные RDA для конкретного ресурса, и может рассматриваться как принятие стандарта. Анализ ценностей библиографических записей может использоваться для определения того, отражают ли продукты стандартов организации знаний ценности, и если да, то как эти ценности соотносятся с ценностями руководящего стандарта.

Опять, краткий анализ ценностей конкретного случая может пролить свет на это. Рис. 3 иллюстрирует библиографическую запись эпистолярной новеллы 2016 г. *The Pharos Gate: Griffin & Sabine's lost correspondence* от Ника Бантока [44]. Эта новелла является завершающей в серии работ, документирующих переписку двух вымышленных влюбленных через различные стили искусства, открыток и вложенных писем. Библиографическая запись для этого источника была создана каталогизаторами Библиотеки конгресса с помощью стандарта RDA; впоследствии она была пересмотрена и обновлена другими организациями. Тематические заголовки источника, которые продиктованы отдельными стандартами организации знания, удалены из этого рисунка.

Для библиографического представления типично, что RDA предлагает много ожидаемых элементов описания, включая автора, название, издателя и дату издания. Тем не менее, отдельные особенности этой записи стоит отметить. Во-первых, имя автора («Личное имя») количественно определяется двумя обозначениями связи, фиксируя его роль как автора и иллюстратора этой работы. Включение этих данных в запись отражает акцент RDA на взаимосвязях, как и подтверждает ясность от лица пользователя. Физическое описание представляет книгу с отсутствием номеров страниц и с содержанием пяти писем. Хотя информацию о соответствующих письмах можно считать полезной для пользователей в определении и понимании ресурса, выбор отразить работу без нумерации страниц можно считать любопытным. Несмотря на то, что каталогизатор RDA мог подсчитать количество страниц, запись этого источника как бесстраничного точнее отражает то, как источник сам себя представляет. Это поддерживает точность представления, но может смутить пользователя относительно масштаба ресурса, и таким образом окажется неспособным удовлетворить пользователя. Наконец, резюме предлагает непрямую ссылку на членство этой работы в серии подобных работ о Griffin and Sabine. Более точная информация о взаимосвязях с другими работами не представлена, демонстрируя, что надежность связей не была полностью осознана.

Хотя библиотеки можно рассматривать главными поставщиками библиографической записи, каталоги также выполняют другие задачи. Снова можно обратиться к общей сетевой среде для выявления контраста. Google Books (Книги) является онлайн-каталогом книг с полнотекстовыми (по возможности) и основными метаданными, разработанными для поддержки сетевого поиска, просмотра и получения [45]. На рис. 4 показано библиографическое представление Google Books книги *The Pharos Gate*.

Personal name:	Bantock, Nick, author, illustrator.
Main title:	The pharos gate : Griffin & Sabine's lost correspondence / written and illustrated by Nick Bantock.
Published/Produced:	San Francisco : Chronicle Books, [2016]
Copyright:	©2016
LCCN:	2015020211
ISBN:	9781452151250 (alk. paper)
Description:	1 volume (unpaged) ; 21 cm + 5 folded letters
Type of material:	Book
Content type:	text
Media type:	unmediated
Carrier type:	volume
Summary:	Shares the fate of lovers Griffin and Sabine, in a volume published simultaneously with the twenty-fifth-anniversary edition of "Griffin & Sabine."

Рис. 3. Библиографическая запись RDA для *The Pharos Gate*.

Title:	The Pharos Gate: Griffin & Sabine's Lost Correspondence
Author:	Nick Bantock
Edition:	illustrated
Publisher:	Chronicle Books, 2016
ISBN:	1452151253, 9781452151250
Length:	60 pages
Subjects:	Fiction > General

Рис. 4. Запись Google Books для *The Pharos Gate*

Библиографическая запись Google Books явно краткая, хотя более длинные описания и обзоры этого ресурса включены в систему поиска и показываются, чтобы подтверждать важные применения поиска, просмотра и получения. Хотя эта запись кажется менее завершенной, чем RDA, отметим, что она содержит 60 страниц подсчета. Пользователь может быть озадачен самой причиной того, почему относительно короткая запись содержит подсчет страниц, тогда как более пристальный взгляд на запись, кажется, пропустил эту деталь. Это, однако, не является причиной завершенности, но, может быть объяснено различием ценностей. Важность точного представления RDA приводит к полностью отличающейся концепции объема этой книги и гораздо более дружелюбному по отношению к пользователю подходу в отличие от более общего смысла Google Books.

Данный краткий анализ показывает, что запись RDA для книги *The Pharos Gate* относит ценность к связям и потребностям пользователей, несмотря на переменный успех. Точность представления также важна, хотя может достигаться за счет ясности. Следуя точному описанию, это библиографическое представление может потерпеть неудачу в сохранении потребностей пользователей. Хотя текст RDA признает здесь баланс между отдельными целями и принципами [43], это все еще может считаться конфликтом между утвержденными и функциональными ценностями. Ожидаемые ценности, такие как взаимосвязи и потребности пользователей, могут быть нарушены или даже отсутствовать при введении стандар-

тов. В целом, этот случай показывает, что даже на вид нейтральные информационные артефакты включают ценности, тесно связанные со стандартами, используемыми для их генерации, и могут быть местами конфликтов и упущений важных ценностей.

ОБСУЖДЕНИЕ

Предыдущие примеры показывают, как анализ ценности можно применить к информационным стандартам и порождаемым ими данным, также они показывают, что стандарты и данные выражают описываемые ценности, и что эти ценности различаются даже в стандартах одной и той же сферы. Даже с точки зрения таких относительно небольших случайных примеров имеются важные применения, связанные с ценностями и информацией. В случае использования вариантов имен это имеет весьма важную ценность, но чье использование важно, отличается между стандартами, порождая различия, в которых имена считаются релевантными. В то же время важные общечеловеческие ценности, такие как автономия и самоидентификация, представляют меньшее значение или вовсе не рассматриваются. В случае с Джейком Зайрусом различия в том, как использование и самоидентификация важны в зависимости от стандарта, могут привести к совершенно разным именам и описаниям одного и того же лица в разных системах организации знания, с противоречивыми и смущающими результатами для пользователей. Хотя пол Зайруса не устанавливается через действующую библиотечную запись, в соответствии с RDA пол и его история также

могут быть зафиксированы. Включение пола в данные RDA подверглось критике за отказ в индивидуальном выборе обнародования, подверженность риску автономии, нарушение конфиденциальности [46], проявление возможности дальнейших нарушений ценности.

В случае *The Pharos Gate* различия в ценностях уже снова играют роль описания одного и того же источника. Гораздо важнее то, что действующая запись RDA для данной книги была показана с целью сокращения поддержки утвержденной ценности взаимосвязей. Взаимосвязь настоящего тома с предыдущими работами в сериях, посвященных тем же самым героям (персонажам), подпадает под область интереса библиографических взаимосвязей к RDA и ее соответствующей концептуальной модели *Функциональные требования к библиографическим записям*. Тогда почему данная взаимосвязь дальше не освещается в действующей библиографической записи? Ответ может лежать в современной технологической среде, в которой производятся данные RDA. RDA описывает идеализированное библиографическое представление, основанное на модели объект-взаимосвязь, хотя имеющиеся каталожные системы полагаются на кодированный формат MARC, подверженный критике за его неспособность легко и точно выражать взаимосвязи [47]. Реализация ценных взаимосвязей RDA сдерживается современными технологическими ограничениями, демонстрируя таким образом, что ценности в стандартах могут быть подвержены угрозе со стороны других аспектов технологической инфраструктуры. Это предполагает, что стандарты как выражения обязательств сообщества перед ценностями должны учитываться в исследовании ценностей и дизайна.

Как жанр документа, стандарты призваны передаваться особыми способами, что ведет к их способности сохранять потенциальных пользователей [48,49]. Точно так же цели и формы стандартов могут способствовать их возможности нести ценности. В узком смысле стандарты призваны передавать и регулировать идеальную или оптимальную реальность и в ходе этого устанавливать закладки для правильных практик и продуктов. Стандарты стремятся определять, что правильно или предпочтительно для данного сообщества, и должны оценивать некоторые вещи за счет других. Таким образом эти документы наследуют подобные соответствующие ценности и более того несут этические коннотации. Этика считается близко связанной с определенными ценностями добра и справедливости [14]. Призванные определять «что такое хорошо», все стандарты имеют также этическое применение. Изучения технологий и этики должны рассматривать стандарты и их роль в установлении того, каковы ценности сообщества и что считать правильным.

Стандарты как письменные документы — это одно, а стандарты на практике — другое. Стандарты предписывают идеал, но они должны вводиться в действие лицами в реальной рабочей среде [50]. Информационные стандарты являются технологическими артефактами, но они также являются и технологическими представлениями. По стандарту, такому как RDA, участники могут руководствоваться ценностями, воспринимаемыми ими из этих документов, но они должны согласовывать их с ценностями своего профессионального сообщества, их отдельными организациями и даже своими собственными ценностями. Как видно в случае *The Pharos Gate*, они также могут нуждаться в компромиссе ради соответствия

важных ценностей технологическим реалиям. Хотя итоговые данные предоставляют нам свидетельство относительно введения в действие информационного стандарта, чтобы полностью понять роль ценностей в использовании стандартов, нам следует изучить перспективы и действия участников. Понимание решений тех, кто интерпретирует и вводит стандарт, является важным для более полного раскрытия применения их ценности.

Анализ ценности предлагает удобную линзу для оценки технологических артефактов и практик и может эффективно служить в раскрытии важных обязательств сообщества относительно ценностей и этических применений. Как правило, анализ ценностей, примененный к текстовому контенту стандартов как документов, гарантирует дальнейшее понимание и сравнение определенных стандартов. Тем не менее, наша концепция анализа ценностей должна распространиться за рамки текста, не только для реализации участниками, но и на более широкую экосистему, в которой располагаются стандарты. Любую область можно рассматривать коллизией многих систем ценностей с ценностями индивидуумов, организаций и артефактов, так или иначе взаимодействующих. Итоговые компромиссы и конфликты имеют важное влияние на истинную роль стандартов, а также либо подтверждают, либо разрушают внутренние ценности. Таким образом, применения анализа ценностей к текстам, участникам, организациям, сообществам, областям, системам и данным — все это поддерживается. Хотя все информационные стандарты представляют здесь интерес, на практике будет особенно полезно сфокусироваться на стандартах организации знания, которые управляют описаниями информации и генерируемых данных, которые широко используются и перерабатываются. RDA, в частности, видит увеличивающееся международное применение и устанавливает влиятельный стандарт библиографических данных в каталогах и более широкой сетевой среде [51]. Анализ ценностей стандартов культурного наследия, таких как RDA, также склонен извлекать выгоду от присутствия четко утвержденных сообществом ценностей, которые уже активно изучаются [52].

ЗАКЛЮЧЕНИЕ

Информационные стандарты не являются нейтральными, не являются они и порожаемыми ими данными. Подобно всем информационным артефактам они способны как сохранять, так и разрушать важные ценности сообщества. Анализ ценностей представляет эффективный подход к раскрытию основных обязательств перед ценностями. Как методологическая традиция он заслуживает место в репертуаре важных подходов к технологиям и технологическим практикам. Информационные стандарты, в частности, представляют значимую область применения для анализа ценностей, поскольку эти документы формализуют важные идеалы сообщества. Как показывают предыдущие ситуационные анализы, ценности, найденные в стандартах и управляемых стандартами данных, имеют применения для потенциально большой базы пользователей в сети. Хотя внимание к этим артефактам крайне необходимо, также важны применения их описаний участников в реальной рабочей среде. За рамками этого, внимание к более широкому ряду информационных артефактов, лиц, групп и практик может помещать стандарты в контекст расширенной

экосистемы ценностей и предложить более полную картину их ролей и действий. С точки зрения информатики, анализ ценностей является полезным подходом к пониманию и сохранению важных ценностей, таких как автономия и конфиденциальность, и поддержанию этических информационных практик.

ЛИТЕРАТУРА

1. *Bowker G. C., Star S. L.* Sorting things out: Classification and its consequences. — Cambridge, Mass.: MIT Press, 2000.
2. *Orsi F.* Value theory. — New York: Bloomsbury Publishing, 2015.
3. *Cheng A.-S., Fleischmann, K. R.* Developing a meta-inventory of human values// Proceedings of the American Society for Information Science and Technology. — 2010. — Vol. 47, No. 1. — P. 1–10.
4. *Gorman M.* Our enduring values revisited: Librarianship in an ever-changing world. — Chicago, Ill.: American Library Association, 2015.
5. *Koehler W.* Ethics and values in librarianship. — Lanham, Md.: Rowman & Littlefield, 2015.
6. *Friedman B., Kahn P., Borning A.* Value sensitive design: Theory and methods (Technical Report 02-12-01). — Seattle, WA: University of Washington, 2002.
7. *Shilton K., Koepfler J. A., Fleischmann K. R.* Charting sociotechnical dimensions of values for design research// The Information Society. — 2013. — Vol. 29, No. 5. — P. 259–271.
8. *Friedman B., Nissenbaum H.* Bias in computer systems// ACM Transactions on Information Systems (TOIS). — 1996. — Vol. 14, No. 3. — P. 330–347.
9. *Fleischmann K. R., Oard D. W., Cheng A. S., Boyd-Graber J., Templeton T. C., Ishita E., Wallace W. A.* Content analysis for values elicitation// Proceeding of the ACM SIGCHI 2012 Conference on Human Factors in Computing Systems, Workshop on Methods for Accounting for Values in Human-Centered Computing. — New York: ACM, 2012.
10. *IEEE Standards Association.* The IEEE Standards Association - Home. — 2017. — <http://standards.ieee.org/>
11. *Busch L.* The moral economy of grades and standards// Journal of Rural Studies. — 2000. — Vol. 16, No. 3. — P. 273–283.
12. *Bowker G. C., Baker K., Millerand F., Ribes D.* Toward information infrastructure studies: Ways of knowing in a networked environment/ International handbook of internet research, (pp. 97–117). — Netherlands: Springer, 2009.
13. *Lampland M., Star S. L.* Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life. — Ithaca, N.Y.: Cornell University Press, 2009.
14. *Rescher N.* Introduction to value theory. — Englewood Cliffs, N.J.: Prentice-Hall, 1969.
15. *McGuire W. J.* Psychological motives and communication gratification// J. G. Blumler & E. Katz (Eds.), The uses of mass communications: Current perspectives on gratifications research (pp. 167–196). — Beverly Hills, Calif.: Sage Publications, 1974.
16. *Rokeach M.* The nature of human values.— New York: The Free Press, 1973.
17. *Hofstede G.* Culture's consequences: Comparing values, behaviors, institutions and organizations across nations (2nd ed.).— Thousand Oaks, Calif.: Sage Publications, 2003.
18. *Bates M. J.* The invisible substrate of information science// Journal of the Association for Information Science and Technology. — 1999. — Vol. 50, No. 12. — P. 1043.
19. *Friedman B.* Value-sensitive design// Interactions. — 1996. — Vol. 3, No. 6. — P. 16–23.
20. *Knobel C., Bowker G. C.* Values in design// Communications of the ACM. — 2011. — Vol. 54, No. 7. — P. 26–28.
21. *White R. K.* Value-analysis: The nature and use of the method. — Glen Gardner, N.J.: Libertarian Press, 1951.
22. *Rokeach M.* Beliefs, attitudes and values: A theory of organization and change.— San Francisco, Calif.: Jossey-Bass Inc, 2001.
23. *Clawson C. J., Vinson D. E.* Human values: A historical and interdisciplinary analysis/ K. Hunt (Ed.), NA – advances in consumer research: Vol. 5 (pp. 396–402). — Ann Arbor, MI: Association for Consumer Research, 1978.
24. *Dhand H.* A value analysis of Saskatchewan social studies textbooks (Doctoral dissertation). — University of Montana, 1967.
25. *Lester N. B.* A system for analyzing characters' values in literary texts// Research in the Teaching of English. — 1982. — Vol. 16, No. 4. — P. 321–338.
26. *Spiggle S.* Measuring social values: A content analysis of Sunday comics and underground commix// Journal of Consumer Research. — 1986. — Vol. 13, No. 1. — P. 100–113.
27. *Shachaf P.* A global perspective on library association codes of ethics// Library & Information Science Research. — 2005. — Vol. 27, No. 4. — P. 513–533.
28. *da Silva A. P., Chaves Guimarães J. A., Bolfarini Tognoli N.* Ethical values in archival arrangement and description: An analysis of professional codes of ethics// Knowledge Organization.— 2015.— Vol. 42, No. 5. — P. 346–352.
29. *Delsey T.* Standards for descriptive cataloging: Two perspectives on the past twenty years/ E. Svenonius (Ed.), The conceptual foundations of descriptive cataloging (pp. 51–60). — San Diego, Calif.: Academic Press, 1989.
30. *Andersen J., Skovvig L.* Knowledge organization: A sociohistorical analysis and critique// The Library. — 2009. — Vol. 76, No. 3. — P. 300–322.
31. *Olson H. A.* The power to name: Representation in library catalogs// Signs. — 2001. — Vol. 26, No. 3.— P. 639–668.
32. *Knowlton S. A.* Three decades since prejudices and antipathies: A study of changes in the Library of Congress Subject Headings// Cataloging & Classification Quarterly. — 2005. — Vol. 40, No. 2. — P. 123–145.
33. *Diao Cao H. J.* Chronology in cataloging Chinese archaeological reports: An investigation of cultural bias in the Library of Congress Classification// Cataloging & Classification Quarterly. — 2016. — Vol. 54, No. 4. — P. 244–262.
34. *Adler M., Huber J. T., Nix A. T.* Stigmatizing disability: Library classifications and the marking and marginalization of books about people with disabilities// The Library Quarterly. — 2017.— Vol. 87, No. 2.— P. 117–135.
35. *Dobreski B.* Authority and universalism: Conventional values in descriptive catalog codes. In Proceedings of the 2017 North American Symposium on Knowledge Organization.— Seattle, WA: University of Washington, 2017.
36. *UNESCO.* Protecting Our heritage and fostering creativity.— 2017.— <http://en.unesco.org/themes/protecting-our-heritage-and-fostering-creativity>

37. *Hider P.* Information resource description: Creating and managing metadata. — Chicago, Ill.: American Library Association, 2012.
38. *Society of American Archivists.* Describing archives: A content standard (2nd ed.). — Chicago, Ill.: Society of American Archivists, 2013.
39. *Baca M.* Visual Resources Association. Cataloging cultural objects: A guide to describing cultural works and their images. — Chicago, Ill.: ALA Editions of the American Library Association, 2006.
40. *Dobreski B., Kwaśnik B.* Changing depictions of persons in library practice: Spirits, pseudonyms, and human books// Knowledge Organization. —2017.— Vol. 44, No. 8.— P. 656-667.
41. *Wikipedia.* Manual of Style/Biographies/Wikipedia. — 2018.— https://en.wikipedia.org/w/index.php?title=Wikipedia:Manual_of_Style/Biographies&oldid=831036517
42. *Svenonius E.* The intellectual foundation of information organization. — Cambridge, Mass.: MIT Press, 2000.
43. *Canadian Library Association.* Chartered Institute of Library and Information Professionals (Great Britain), Joint Steering Committee for Development of RDA, & American Library Association. RDA toolkit: Resource description & access. — 2010. —<http://www.rdatoolkit.org/>
44. *Bantock N.* The Pharos Gate. — San Francisco: Chronicle Books, 2016.
45. *About Google Books.* — Google Books. — n.d. — <https://www.google.com/googlebooks/about/>
46. *Billey A., Drabinski E., Roberto K. R.* What's gender got to do with it? A critique of RDA 9.7//Cataloging & Classification Quarterly. — 2014.— Vol. 52, No. 4.— P. 412–421.
47. *Boehr D., Reynolds R. R., Shrader T.* The US RDA test process // The Serials Librarian. —2015. — Vol. 62, No. 1–4. — P. 125–139.
48. *Young J. J.* Constructing, persuading and silencing: The rhetoric of accounting standards// Accounting, Organizations and Society. — 2003. — Vol. 28, No.6. — P.621–638.
49. *Feinberg M.* Information system design for communication: The use of genre as a design element// Proceedings of the Association for Information Science and Technology. — 2009. — Vol. 46, No. 1. — P. 1–18.
50. *Palme J., Pargman D.* ASCII Imperialism// M. Lampland & S.L. Star (Eds.), Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life (pp. 177–199). — Ithaca, N.Y.: Cornell University Press, 2009.
51. *Marden J., Li-Madeo C., Whysel N., Edelstein J.* Linked open data for cultural heritage: Evolution of an information technology//Proceedings of the 31st ACM International Conference on Design of Communication (pp. 107–112). — New York: ACM, 2013.
52. *Vecco M.* A definition of cultural heritage: From the tangible to the intangible// Journal of Cultural Heritage.— 2010.— Vol. 11, No. 3.— P. 321–324.

Связывание научных ресурсов: построение жизненного цикла их связи*

Мэтью МАЙЕРНИК
(Matthew MAYERNIK)

Национальный центр атмосферных исследований, Университетская корпорация атмосферных исследований, США

Научные ресурсы, включая публикации, программное обеспечение, массивы данных и инструменты, создаются итеративно и взаимосвязано. Управление взаимосвязями, существующими между такими ресурсами и среди них, является основным требованием для информационных систем. Однако практически многие научные ресурсы существуют в режиме онлайн в виде дискретных связей, отделенных от других ресурсов, с которыми они тесно связаны. Надежная система для объединения научных ресурсов на широкой и устойчивой основе должна будет ориентироваться на комплекс сложных и взаимосвязанных потребностей. В статье представлены результаты и выводы трех различных проектов, которые направлены на поддержку более прочных связей между научными ресурсами. В ходе обсуждения подробно рассматриваются ключевые технические и институциональные проблемы, которые во времени могут рассматриваться как «жизненный цикл отношений»: идентификация, подтверждение, характеристика и сохранение отношений. Цель статьи — помочь в руководстве новыми исследовательскими инициативами и оперативными службами, нацеленными на интеграцию информации о взаимоотношениях в научные исследования.

ВВЕДЕНИЕ

Многие научные ресурсы сегодня доступны онлайн. Журнальные статьи публикуются через платформы в режиме онлайн, отчеты и другая серая литература предоставляются через институциональные и общеперсональные хранилища, массивы данных все больше архивируются в доступных сетевых хранилищах данных, а пакеты программного обеспечения широко распространяются через GitHub и иные средства обмена кодами. Безусловно, не все ресурсы предоставляются онлайн. Многочисленные хорошо известные социальные, культурные и технические факторы сдерживают распространение научно-исследовательских продуктов, а ученые в некоторых

случаях обеспечивают конкурентное преимущество за счет максимизации их уникального доступа к новым данным, средствам или знанию [1-3]. Конечно, это также относится к случаю, когда предоставление онлайн доступа к чему-либо автоматически не гарантирует его пользу или завоевание понимания со стороны широкой аудитории [4]. Тем не менее, тенденция ясна и предполагает, что научные ресурсы уже сейчас доступны онлайн или будут все больше становиться доступными онлайн. Политическое давление, технические возможности и социальные нормы — все это подталкивает к такой направленности [5-7]. В рамках научной работы наличие доступных онлайн ресурсов дает широкие возможности. Научные сообщества извлекают пользу из повсеместного доступа к огромному числу ресурсов, а отдельные ученые получают выгоду от появляющихся преимуществ библиографических ссылок и чтения [8, 9].

Если ресурсы доступны онлайн, то, естественно, возникает вопрос: можем ли мы связать ресурсы вместе? В конце концов, связи (ссылки) — это то, из чего состоит весь Интернет. Публикации, программное обеспечение

* Перевод Mayernik M. Scholarly resource linking: Building out a “relationship life cycle”//Proceedings of 81 st Annual Meeting ASIS&T. — <http://www.asist.org/wp-content/uploads/Final-81st-Annual-Meeting-Proceedings-1.pdf>

ние, массивы данных и другие научные ресурсы создаются итеративными и взаимосвязанными. Взаимосвязи между научными ресурсами можно охарактеризовать как образующие важную цепь, в которой они обеспечивают существенное значение для подходов к обнаружению, использованию, управлению и сохранению ресурса [10, 11]. Управление такими взаимосвязями в принципе является ключевой составляющей (основной компонентой) информационных систем [12]. В самом деле многие системы информации и данных управляют и усиливают различные взаимосвязи, в частности, взаимосвязи между словарными терминами и структурами содержания [13]. Однако на практике большинство научных ресурсов существуют онлайн как дискретные связи. Исследование реестра хранилищ научно-исследовательских данных (registry of research data repositories, <https://www.re3data.org>, исследование проводилось 28 марта 2018 г.) обнаружило только 48 хранилищ из 2040 идентифицированных как принимающих «код источника» и 26 хранилищ как поддерживающих «применение программного обеспечения», некоторые хранилища имеют и то, и другое. Эти цифры, скорее всего, не являются точными, но отражают, насколько услуги по хранению данных и сопровождению программного обеспечения в хранилищах разъединены. Наравне с этим, научные статьи принадлежат системам и хранилищам, которые, как правило, не собирают массивы данных или программного обеспечения.

Таким образом, организация связи научных ресурсов требует подхода, который осуществляет навигацию между многими научными учреждениями и техническими системами. Установление и управление отношениями между информационными ресурсами стало общей темой в информатике и изучении технологий. Данная статья не стремится проанализировать всю релевантную литературу и инициативы в этой сфере. Полезные обзоры можно найти в других источниках [14-17]. Наоборот, статья представляет результаты и выводы трех различных проектов, направленных на организацию связи между данными и литературой. В ней подчеркиваются основные проблемы идентификации, подтверждения, характеристики и сохранения отношений между научными ресурсами и обсуждается, как эти проблемы различаются при взгляде во времени в будущее или прошлое. Целью статьи является помощь в руководстве новым исследованием и оперативными службами, нацеленными на интеграцию информации об отношениях в более полном объеме в научные исследования.

ИССЛЕДОВАНИЯ В СФЕРЕ СВЯЗЫВАНИЯ

В этом разделе представлены три отдельных проекта, реализованных в последние четыре года. Каждый проект отличался своей целью, масштабом и партнерами, но их общей темой было исследование подходов к перекрестному связыванию схожих научных ресурсов. Как отмечают авторы работы [16] в своей недавней ретроспективе, относящейся к серии научных ресурсов и проектов по интероперабельности инфраструктуры, из-за разнообразия проблем и вовлеченных участников «едва ли можно точно знать, как следует начать работу в направлении растущей интероперабельности». Указанные проекты следует рассматривать, как попытку дать старт множеству параллельных начинаний в поисках понимания и рассмотрения вероятного исследования проблем перекрестного связывания научных ресурсов.

Ниже отдельно обсуждается каждый проект. Каждый подраздел описывает соответствующие цели, подходы и релевантные действия проекта. Также определяются некоторые важные уроки, усвоенные через работу в каждом проекте. Следуя описаниям этих проектов, приводится междисциплинарное обсуждение, связывающее воедино точки зрения по каждому проекту в целях разработки широко применяемых выводов, сделанных на основе усилий по перекрестному связыванию научных ресурсов.

Перекрестное связывание от хранилища к хранилищу

Первый проект (январь 2015 г. – март 2016 г.) содержал экспериментальную попытку провести обмен связанной информацией между двумя хранилищами – одним хранилищем данных и одним хранилищем литературы. Цель состояла в том, чтобы позволить двум хранилищам напрямую взаимодействовать для обмена информацией о ссылках на ресурсы, которые имели известные отношения, но располагались и управлялись по отдельности, такие как, например, когда массивы данных, расположенные в хранилище А, использовались для производства публикаций, расположенных в хранилище В. Задумкой данного проекта была разработка процесса взаимообмена между двумя хранилищами, позволяющая исследователям, депонирующим ресурсы в одной системе, инициировать депонирование близких источников в другом хранилище. В идеале связи между близкими ресурсами, содержащимися в двух отдельных системах, могут быть обменены и видимы в соответствующих хранилищах. С точки зрения системы, целями были идентификация, охват, проверка и использование особенностей хранилища, позволяющих создавать, обменивать и поддерживать во времени связи между близкими ресурсами при небольших технических препятствиях и минимальных усилиях со стороны куратора хранилища. Полное от начала до конца применение этой идеи не было завершено, как обсуждается далее. Но процесс работы над проектом позволил нам исследовать основные потребности в том, как может быть достигнуто (получено) перекрестное связывание хранилищ. У нас также был семинар в конце проекта, на котором произошла более широкая дискуссия по его темам [17].

Подход к проекту состоял в: 1) установлении случаев использования и участников; 2) разработке технических потребностей для этих случаев использования и участников и 3) разработке системных функциональностей, которые могли бы удовлетворить потребности случаев использования и участников. Рис. 1 отражает четыре сценария организации связи между ресурсами, которые относятся к случаям использования и участникам, для связи документа с соответствующими ему данными. В сценариях #1-3 рис. 1 создатель контента (например, автор публикации и/или данных) может быть заинтересованным в депонировании двух новых ресурсов или депонировании нового ресурса, относящегося к депонированному ранее ресурсу. В сценарии # 4 кураторы хранилища могут быть заинтересованы в идентификации связей между ресурсами, уже являющимися частью существующих у них собраний.

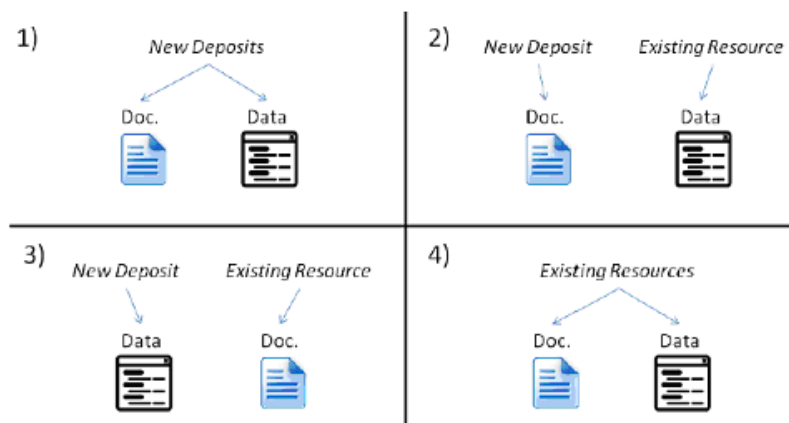


Рис. 1. Сценарии рабочих процессов организации связи между хранилищами

Эти сценарии представили ряд поддерживающих задач, включая: а) провайдеры массива данных могут загружать публикации в одно хранилище и депонировать массивы данных в другое (или наоборот); б) кураторы контента хранилища будут оповещены при попытке инициировать новые представления из внешней системы; в) авторы публикаций или данных будут иметь возможность обмениваться депонированными ресурсами с коллегами и соответствующим образом ссылаться на них; г) потребители результатов научных исследований (люди или машины) могут пользоваться хранилищами для поиска ресурсов и их взаимосвязей. При подразделении этих задач на подзадачи и другие потоки данных были сформированы следующие технические потребности в качестве основы для установления и поддержания надежных связей между хранилищами связанных ресурсов:

- Механизм оповещения для отправки информации о связанных ресурсах между хранилищами;
- Модель данных для обмена метаданными, созданная в целях:
 - выявления наследия метаданных (например, авторов или их мест работы, являющихся общими для разных ресурсов);
 - выявления типов взаимосвязей (например, массив данных соответствует публикации);
 - выявления во времени информации, необходимой для поддержки взаимосвязей между ресурсами в качестве появляющихся обновлений (например, создание нового массива данных или создание новой публикации на основе существующего массива данных).

- Эффективные интерфейсы депонирования ресурса, допускающие простую декларацию перекрестных связей между данными и публикациями в хранилище;
- Эффективный показ интерфейса, делающий перекрестные связи в хранилище прозрачными для пользователя.

Изучался сервис SHARE Notify (<https://share.osf.io>) как средство оповещения, позволяющее хранилищам взаимодействовать и обмениваться между собой информацией о взаимосвязях. К моменту нашего исследования SHARE Notify разрабатывался Центром открытой науки как еще один сервис третьей стороны по созда-

нию оповещений о «событиях» вокруг научных публикаций. Наш интерес в использовании SHARE Notify заключался в отправке оповещений о «событиях» взаимосвязи между нашими хранилищами, при этом данный сервис выступал в качестве посредника между двумя хранилищами. Обращение к стороннему промежуточному сервису, такому как SHARE Notify, имело место по причине того, что он исключал потребность каждого хранилища напрямую связываться с другим. Хотя наш случай использования ставил своей целью связать два хранилища, для нас представляло интерес средство распространения информации о взаимосвязях нескольким потенциальным партнерам. Таким образом, нам нужен был подход, который, вероятно, мог определить масштаб поддержки обмена информацией о взаимосвязях между несколькими хранилищами. Нам хотелось найти третью сторону, которая могла бы обеспечить основной реестр выпуска научных «событий» через обратную связь метаданных. У сервиса SHARE Notify имелся потенциал удовлетворить наши желания, хотя наш случай использования не был для него приоритетным.

Проводились пробные эксперименты по использованию сервиса SHARE Notify, но в конечном счете по ряду причин не были полностью выполнены рабочие процессы во время реализации нашего проекта. Во-первых, сервис SHARE Notify сам был на стадии разработки и итераций. Поэтому он представлял подвижную цель наших потребностей. Модель данных сервиса SHARE Notify и интерфейсы прикладного программирования не были завершены на момент старта нашего проекта. Это основной урок, полученный из данного проекта. Стабильность стороннего сервиса оповещений является внешним фактором, который может воздействовать на надежность локальных рабочих процессов.

Другим, усвоенным из проекта и не связанным с сервисом SHARE Notify, уроком было то, что подобная нашей попытка должна быть агностической относительно идентификации схем, используемых для установления ресурсов. Некоторые из изучаемых нами технических рабочих процессов требовали определения массивов через DOI, в особенности, чтобы использовать преимущество хранилищ метаданных, принятое агентствами регистрации DOI (например, CrossRef и

DataCite). Однако многие ресурсы в двух задействованных в этом проекте хранилищах не были (и предположительно не должны были быть) приписаны к DOI.

Отслеживание связей ресурсов через библиографические ссылки

Описываемый здесь второй проект (с 2014 г. по настоящее время) фокусировался на разработке и средствах оценки автоматического отслеживания научных инфраструктур в научной литературе с помощью постоянных идентификаторов (PID) цитирования. Относительно связи интерес состоял в отслеживании связей между статьями и соответствующими ресурсами, такими как массивы данных, которые использовались для производства статей. Стимулом данного проекта был рост интереса к «цитированию данных», т. е. присвоение данным постоянных идентификаторов, таких как DOI, данным с целью стимуляции их цитируемости и возможности отслеживания наравне с традиционной научной литературой [18, 19]. Однако для нас интерес заключался в более широком изучении пользы постоянной идентификации научных ресурсов, включая научное оборудование (или набор инструментов), пакеты научного программного обеспечения, вычислительные системы и коммуникационные сети. Воспользуемся многозначным термином «научные инфраструктуры», чтобы выделить эту широкую группу. Поскольку присвоение и использование PID научно-исследовательским инфраструктурам является относительно новой разработкой, было проведено очень мало оценок, которые системно изучали бы воздействие таких идентификаторов, присвоенных подобным инфраструктурам. Таким образом, цели нашего проекта состояли из: 1) развития понимания того, как методологически и последовательно анализировать научное влияние исследовательских инфраструктур, и 2) разработки автоматизированных технологий, дающих возможность отслеживать научные инфраструктуры с целью более эффективной работы. Отдельные вопросы исследования касались, например, того, как со временем изменялись ссылки на научно-исследовательские инфраструктуры относительно присвоения PID данным, программному обеспечению и иным компонентам научных инфраструктур.

Первой задачей данного проекта было проведение оценки на основе ситуационного анализа того, насколько часто исследовательские инфраструктуры идентифицируются и отражаются в научной литературе через постоянные идентификаторы (PID - Persistent Identifiers) цитирования. В исследовании [20] цитирования и библиографические ссылки на четыре ресурса – два массива данных, пакет программного обеспечения и чрезвычайная вычислительная легкость – собирались вручную с помощью поиска в сервисе Google Scholar и анализировались, чтобы оценить способы, которыми ссылаются на ресурсы пользующиеся ими ученые, через приведение характеристики того, как часто на данные ресурсы делались ссылки в статьях с помощью полнотекстового описания, упомянутого в благодарности или эксплицитно цитируемого в библиографическом списке. Результаты этого исследования показали, как на самом деле использовались постоянные, присвоенные четырем анализируемым ресурсам, идентификаторы в ссылках в опубликованных статьях на основе роста. Но здесь не было постоянной последовательной модели во

всех четырех ситуационных анализах относительно того, как этот рост управлялся. Наравне с этим анализ обнаружил, что практики библиографических ссылок со временем менялись, а степень этих изменений варьируется в значительной мере от ресурса к ресурсу. Основным выводом из этих результатов является то, что изменение установленных практик библиографических ссылок и признание массива данных будут вероятно более трудными, чем создание новых практик библиографических ссылок для иного рода продуктов, таких как возможности программного обеспечения или вычислений.

Вышеупомянутое исследование проводилось в основном вручную. Второй задачей в рамках данного проекта была разработка вычислительных алгоритмов/методов с целью облегчения оценки подобного рода проекта. Разработка алгоритмов и методов отслеживания придерживалась типового подхода машинного обучения, сконцентрированного на трех направлениях: 1) сбор публикаций-кандидатов для классификации, 2) разработка экспериментальной методологии по классификации и 3) автоматизированный подход к проведению классификации и анализа документа. Данный подход признается многообещающим благодаря первоначальной проверке классификаторов документов на способность правильно определять, действительно ли документ из тестового массива использует интересующую нас вычислительную способность, основанную на характеристиках метаданных и полного текста документа. Ограничение этого исследования, таким образом, намного больше и состоит в том, что основано на относительно небольшом массиве документов. Пробный массив (исследуемый вручную и присваивающий обозначения) включает около 300 документов, а тестовый массив (присваивающий обозначения, но еще не изученный или обработанный) содержит около 120 документов. Эти цифры далеки от идеальных в исследованиях по машинному обучению, которые часто используют тысячи или даже миллионы документов.

Основным выводом этого проекта является то, что трудно предсказать масштаб человеческого опыта, но он (опыт) крайне важен для идентификации и подтверждения взаимосвязей в тех случаях, где нет возможности использовать подходы на основе вычислений. Машинная автоматизация легко масштабируется, но гарантии и точность измерения автоматизированными средствами сбора связей очень непросты.

Оба исследования, предпринятые в ходе этого проекта, столкнулись с одинаковой проблемой, главным образом той, что сбор большого числа документов, на которые опирается анализ любых тенденций на основе литературы, весьма труден вне узких областей, таких как биомедицина (PubMed), физика (arXiv.org) и астрономия (Astronomical Data Service), где масса литературы хранится в публично доступных и машиночитаемых системах. В других областях, включая сферы интереса данного проекта, литература распределена по многим издательским платформам, не допускающим любую всеохватывающую возможность доступности через ЭВМ. Google Scholar, хотя и имеет достаточно большой охват областей, не позволяет какое-либо значительное автоматическое извлечение литературы. Данные методологические трудности служат общей предпосылкой для многих исследований, нацеленных на выполнение и анализ метрик влияния научно-исследовательских инфраструктур [18]. Эти ограничивающие масштаб про-

блемы авторского права, наравне с редакционной политикой журналов, издательскими платформами и различиями в форматах статей, представляют иные неконтролируемые факторы для автоматизации такого вида исследований.

EarthCollab – связывание по семантической сети

Третий проект, представляющий интерес для данной статьи и называемый EarthCollab (2014 – 2018 гг.), фокусировался на использовании семантической сети и технологий связанных данных, чтобы облегчить координацию и организацию сложных научных проектов и их продуктов. С технологической точки зрения цель проекта заключалась в разработке информационных систем, демонстрирующих, как геонауки могут усилить связанные данные для производства более когерентных методов обнаружения информации и данных в крупных междисциплинарных проектах и виртуальных организациях. Мотивацией к этому проекту служило улучшение обнаружения и обмена информацией в целях оказания содействия исследовательскому и научному сотрудничеству, стимуляции ученых к более простому поиску людей, организаций и научно-исследовательских ресурсов, отвечающих их занятости. Пакет программного обеспечения семантической сети VIVO (<http://vivoweb.org/>) был выбран из-за его опоры на модель данных, сконцентрированную вокруг сети и сфокусированную на представлении взаимосвязей между ее объектами [21]. Рис. 2 отражает, как взаимосвязи «многие-ко-многим» (many-to-many) существуют среди (и внутри) научных ресурсов, проектов, людей и организаций.

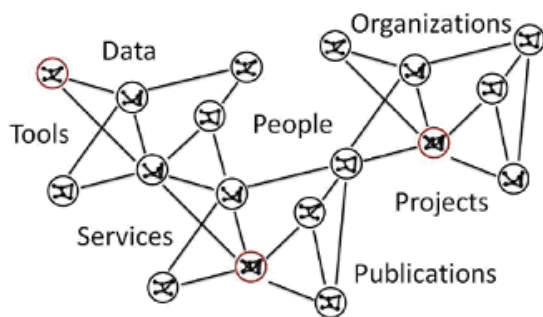


Рис. 2. Сетевая наука

Проект использовал VIVO для обеспечения ученых сетевыми интерфейсами в рамках двух целевых научно-исследовательских областей, чтобы они изучали людей, публикации, платформы и массивы данных внутри их соответствующих сообществ. Эти системы, названные «Connect UNAVCO» (<https://connect.unavco.org>) и «Arctic Data Connects» (<http://vivo.eol.ucar.edu>), более года были открыты для публичного пользования. В иллюстрации того, какого рода информация находилась в этих системах: например, по состоянию на 1 апреля 2018 г. система «Connect UNAVCO» содержала записи почти 6000 научных документов, 4307 массивов данных, 3841 научно-исследовательский сайт, 228 грантов, сведения относительно 803 человек и 381 организации. Система «Arctic Data Connects» с более пристальным вниманием

на ситуационный анализ по выбору арктических научно-исследовательских проектов содержала записи 354 массивов данных, 26 грантов, сведения о 146 ученых и 53 организациях. Информация в этих системах появлялась из сочетания существующих баз данных метаданных и недавно созданных метаданных. Ресурсы представлены через модель данных семантической сети с помощью использования многочисленных онтологий [22]. В семантической сети онтологии применяются для определения типов объектов (классов) и отношений между ними (свойств). Все, что угодно, может быть представлено в качестве объекта первого порядка, как декларируется в соответствующих онтологиях. В результате получается сетевой набор информации, в котором каждые массивы данных, публикации, научные инструменты, исследовательские проекты и сайты представлены в модели данных отдельными объектами, имеющими определенный тип отношений с другим объектом. В частности, программное обеспечение VIVO показывает сетевую страницу каждого объекта, предоставляющую информацию об объекте вместе со связями с другими объектами, выраженными через эксплицитно заявленные отношения. Например, сетевая страница массива данных внутри одной из систем покажет все известные ссылки на соответствующие публикации, организации, гранты, создателей и инструменты.

Другой составляющей этого проекта было осуществление разработки программного обеспечения с целью добавить возможности пакету программного обеспечения VIVO. В течение первых двух лет ведения проекта был разработан прототип подхода «перекрестного связывания» для обмена информацией на примерах VIVO. Мотивацией для данной работы являлось то, что участники многих научно-исследовательских проектов находятся в разных организациях. Некоторые из этих организаций уже пользуются VIVO, чтобы управлять информационными профилями преподавателей и персонала. Новая особенность перекрестного связывания была разработана в целях поощрения обмена информацией об отдельных людях или объектах, общих для разных примеров VIVO. Задача новой особенности – сократить дубликации информации между системами и стимулировать распределения авторитетной информации об отдельном объекте, такой как в случае наличия у данного лица VIVO профилей в разных организациях. Возможность перекрестного связывания была развернута в рамках системы «Connect UNAVCO» и передана базе кода открытого источника ядра VIVO.

Основным уроком, усвоенным из проекта EarthCollab, было представление взаимосвязей об использовании и переработке онтологий между центрами научных ресурсов. Для упрощения интеграции и обмена данными в геонаучном сообществе нашей целью в рамках данного проекта была, насколько это возможно, переработка существующих онтологий в ходе реализации проекта онтологий и сетевых применений. Переработка существующих онтологий оказалась не такой легкой, как предполагалось первоначально. Трудность в работе по использованию онтологий была в концептуальном моделировании, в котором определены и отображены представляющие интерес ключевые объекты и отношения. Только после этого можно делать следующий шаг — главным образом поиск релевантных существующих онтологий, представляющих вашу концептуальную мо-

дель, что не является тривиальной задачей. В нашем проекте ни одна онтология не поддерживала потребности нашего проекта. Таким образом, сочетались компоненты пары разных онтологий и создавались расширения онтологий потребителей в качестве необходимости устранения пробелов [22].

ОБСУЖДЕНИЕ

Все, описанные выше, три проекта фокусируются на проблемах, касающихся связывания вместе научных ресурсов в сетевой среде. Проекты велись параллельно с некоторым перекрытием участников, но без прямого повтора в полученных результатах. Принципы работы проектов, однако, перекрывались. Вместе они обеспечивают более широкий взгляд на компоненты и участников попыток перекрестного связывания научных ресурсов, чем каждый в отдельности. Такого рода подход «параллельных усилий» оказался весьма полезен в ситуациях, когда оптимальный результат (или путь к результату) не ясен с самого начала [23].

В этом разделе представлено всестороннее обсуждение потребностей перекрестного связывания научных ресурсов, основанное на знании и опыте, полученном при одновременной работе над многими релевантными проектами. Сначала категоризируем по типу наиболее заметные потребности и связанные в пространстве этой проблемы задачи, затем опишем, как по-разному выглядят эти потребности и задачи при взгляде в прошлое и будущее во времени.

Потребности перекрестного связывания научных ресурсов

Надежная система связывания научных ресурсов в широкой и устойчивой манере должна будет преодолеть ряд сложных и взаимосвязанных требований. Настолько, насколько научные данные и создание, управление и сохранение программного обеспечения могут быть характеризованы через модель жизненного цикла [24, 25], можно характеризовать потребности инфраструктуры связывания научных ресурсов через «жизненный цикл их связи», включающий потребность в идентификации, подтверждении, характеристике и сохранении информационных отношений и связей. Каждая из этих областей обсуждается ниже.

Идентификация отношений. Отношения между научными ресурсами, столь полезные для обнаружения, понимания и применения ресурса, должны быть декларированы кем-то или каким-то объектом. Откуда могут (или должны) происходить декларации этих отношений? Отношения, не декларируемые явно, иногда могут определяться вычислительными процессами, иногда полагаться на статистические измерения идентификации скрытых отношений между отдельными объектами или словарными терминами, включенными в текст [26]. Но эти вычислительные техники, как правило, более успешны в особых случаях применения и трудны для обобщения. Другой очевидный источник информации об отношениях находится в создателе связанных ресурсов. Банальностью поколения метаданных для научных ресурсов является то, что ученые наилучшим образом позиционированы для предоставления метаданных о своих собственных ресурсах с учетом их непосредственного знания о том, как эти ресурсы были созданы и

как использовались. Но попытки собрать информацию о взаимосвязях между данными, программным обеспечением, публикациями и иными научными ресурсами напрямую от ученых также наталкиваются на известные проблемы создания метаданных, главным образом те, что ученые проявляют мало инициативы по отношению к четкому описанию этих взаимосвязей и что знание отдельных отношений может быстро снижаться и распределяться между командами участников [27, 28]. Сбор информации об отношениях непосредственно от ученых, таким образом, получается трудной задачей, которая плохо поддается масштабированию из-за распределения в сети участников научного сообщества. Другим потенциальным источником информации об отношениях является опубликованная литература. Как обнаружилось в ходе нашего проекта *Tracing identifiers*, цитирования данных и программного обеспечения в самом деле формально все больше приводятся в ссылках, тем самым делая вероятно возможным автоматизированный анализ цитирований. Но наш проект также обнаружил, что временной интервал для укоренения этих новых цитирований будет очень долгим. Таким образом, идентификация отношений является важным ограничителем роста любой попытки по развитию подходов на основе связывания научных ресурсов. Эта задача также имеет бинаправленную составляющую, главным образом такую, что даже если идентификация отношений становится более прямолинейной, то как пропагандируются эти отношения? Иными словами, если я знаю о связи между моим и вашим ресурсами, как вы узнаете об этой связи? Как декларация моей связи распространится на вас? Это условие подогревало наш интерес в обслуживании уведомления о взаимосвязи в рамках нашего проекта перекрестного связывания хранилищ. Сегодня нет всестороннего агрегатора информации об отношениях, который мог бы служить в качестве подобной общей службы уведомления. В отдельном случае связывания данных с литературой, подход Scholix и близкая служба взаимосвязи данных с литературой (DLI), способны обслуживать основные роли централизованного стороннего агрегатора деклараций взаимосвязей [29, 30].

Подтверждение отношений. Подтверждение представляет следующую задачу и условие. Какой бы метод не использовался в идентификации отношений между научными ресурсами, декларации отношений нуждаются в подтверждении через некий процесс, который опять может основываться на человеческом или машинном труде. Основной проблемой процесса подтверждения является простое подтверждение относительно того, какие объекты соотносятся. Цифровые объекты рассеивают границы и могут изменяться со временем. Многие подходы по идентификации взаимосвязи полагаются на использование постоянных идентификаторов, таких как DOI, чтобы гарантировать постоянство и фиксированность. Тем не менее, технически PID действуют как локаторы ресурсов, к которым они приписаны, а не как идентификаторы этих ресурсов [31, 32]. Постоянные идентификаторы (PID) также приписываются на различных уровнях градации разного рода ресурсов, и нет повсеместно принятых правил управления решениями относительно того, как PID должны быть приписаны сложным цифровым объектам [33]. На практике использование PID не всестороннее, как обнаружил наш про-

ект *Tracing identifiers*. Многие библиографические ссылки на данные, программное обеспечение и иные источники все еще происходят через неформальные благодарности или внутритекстовые ссылки. Подтверждение деклараций отношений также зависит от понятия авторитетного источника информации об отношениях. Является ли информация об отношениях, полученная от создателя ресурса или из опубликованных статей, более авторитетной, чем информация из других источников? Служба DLI, отмеченная в предыдущем абзаце, предоставляет информацию об организациях, публикующих декларации об отношениях. Но должны ли все провайдеры рассматриваться авторитетными источниками информации о взаимосвязях? Механизмы обеспечения доверия в декларациях об отношениях являются основой любого подхода подтверждения отношений.

Характеристика отношений. Для многих целей основной акт декларации о существовании взаимосвязи не является очень полезным. Больше информации о типе отношений также можно собрать. Иными словами, использование взаимосвязи часто требует дополнительного описания самой взаимосвязи. Подход семантической сети основан на этой возможности, а именно на том, что отношения должны быть обозначены как особые типы наименований. Многие онтологии и словари создавались, чтобы определять особые типы отношений, которые могут встречаться между научными ресурсами [17], но эти типологии отношений очень разнообразны и непостоянны. В нашем проекте EarthCollab внимание сосредотачивалось на переработке существующих онтологий таким образом, чтобы представлять объекты и их отношения интероперабельно. Это кажется более сложным, чем предполагалось заранее, в силу того факта, что многие онтологии моделировали одинаковые типы ресурсов (например, массивы данных) и отношения (например, отношения ссылок) по-разному [22]. Дополнительной трудностью характеристики отношений является то, что разные применения информации об отношениях могут нуждаться в различных уровнях описания объектов при любом типе отношений. Подход Scholix, например, определяет очень простую модель данных для представления отношений [30]. Что касается других применений, включая наш проект EarthCollab самая простая модель данных не поддерживала цели обнаружения и понимания данных. Общий вопрос, касающийся характеристики отношений, состоит в том, что модели данных и схемы метаданных не всегда включают способы представления информации об отношениях или они имеют разные требования к описанию отношений. Модель данных SHARE Notify, например, не имела явного способа представления отношений в то время, когда изучался потенциал ее использования как службы уведомления об отношениях в рамках нашего проекта перекрестного связывания хранилищ.

Сохранение отношений. «Связывание таблиц объектов» является самой очевидной проблемой сохранения отношений между научными ресурсами на основе сети. Сетевые сайты изменяют URL или неожиданно исчезают, вызывая каскад ошибок любых ссылок, указывающих на сайт, который перестал существовать. Эта проблема представляет важный вопрос для проектов, которые стремятся использовать научную литературу как источник информации об отношениях. Постоянные идентификаторы (PID) снова являются решением этого

вопроса. Все системы PID работают через перенаправляющие серверы. Поддержка перенаправлений представляет институциональную проблему, поскольку является технической. Организации, которые регистрируют DOI, требуются, например, для поддержки резолюции их идентификаторов и связанных языковых страниц. Кроме этого, проблемы сохранения затрагивают уровень описания, необходимый для понимания отношений с течением времени и /или по мере изменения во времени сообществ пользователей ресурсов и их отношений. Процессы, связанные с написанием и задокументированием отчета Американской национальной комиссии по изменению климата, например, межорганизационного отчета—соглашения по борьбе с изменением климата, значительно изменились в последнее время, преследуя цель обеспечить высоко структурированную информацию об отношениях для иллюстрации особых связей между научными потребностями и соответствующими данными и научно-исследовательскими статьями [34]. Авторы работы [35] приводят другой наглядный пример того, как потребности пользователя и потребности документации могут со временем значительно смещать акцент даже в одном и том же научном ресурсе. Следует отметить, что в двух приведенных примерах создание и курирование этой информации входило в ответственность соответствующего персонала. Сохранение понимания отношений является, таким образом, продолжающимся процессом, который может требовать серьезного опыта. Как часть потребности этого сохранения, подходы, подобные Scholix, и системы, такие как служба DLI, должны знать о своих собственных моделях поддержки, чтобы порождать доверие потенциальных пользователей.

Взгляд во времени на прошлое и будущее

Вопросы, связанные с идентификацией отношений, подтверждением, характеристикой и сохранением, представляют разные вызовы и потребности при взгляде во времени в прошлое и будущее. Рабочие процессы проекта перекрестного связывания хранилищ, изображенные на рис. 1, дают об этом четкое представление. Смотри в будущее во времени, вызовы концентрируются на том, как выстроить отношение декларация/идентификация в качестве установившейся и надежной части научного издательства, архивирования данных и программного обеспечения и технологий хранилища. Как отмечают авторы [36], «эффективный охват данных, включая источник и метаданные, легче всего осуществляется на начальном этапе рабочего процесса» [36, с. 17]. В сценарии, в котором ресурсы созданы заново и депонированы в системах хранилищ, представляется возможным сбор информации об отношениях, по мере их производства.

Смотри назад в прошлое, связывание научных ресурсов включает поиск доступной литературы (и другой документации) вместе с потенциальным непосредственным запросом создателей научных ресурсов относительно отношений, существующих между имеющимися ресурсами. Поиск литературы может быть ручной или автоматизированной, как в нашем проекте *Tracing identifiers*, с предполагаемыми достоинствами и недостатками любого подхода. Ограничения авторского права и лицензий явно уменьшают современные усилия по поиску литературы, за исключением академических специальностей, где хорошо организованы модели открытого доступа.

**Техническая и институциональная работа, необходимая для поддержки
связывания научных ресурсов**

	Взгляд во времени в прошлое	Взгляд во времени в будущее
Идентификация взаимосвязей	<u>Техническая работа</u> – подходы к извлечению данных для выборки ссылок на основе PID из опубликованной литературы <u>Институциональная работа</u> – взаимодействие с издателями и финансирующими организациями по открытой научной литературе для получения всестороннего анализа данных	<u>Техническая работа</u> – разработка агрегаторов взаимосвязей и связанные открытые сетевые службы <u>Институциональная работа</u> – 1. продвижение соответствующего использования PID. 2. разработка и адаптация подходов сообщества к распределению взаимосвязей (например, Scholix)
Подтверждение взаимосвязей	<u>Техническая работа</u> – подходы к извлечению данных для выборки неформальных ссылок на основе PID из опубликованной литературы с помощью оценки доверия <u>Институциональная работа</u> – взаимодействие с издателями и финансирующими организациями	<u>Техническая работа</u> – разработка/обновление моделей данных для поддержки взаимосвязи источника и достоверных утверждений <u>Институциональная работа</u> – разработка организационных сетей доверия для декларации отношений
Характеристика взаимосвязей	<u>Техническая работа</u> – соответствие нестандартным декларациям отношений в сообществе онтологий и схемах метаданных <u>Институциональная работа</u> – сочетание описаний объекта и взаимосвязей в потребности целевых сообществ и/или применений	<u>Техническая работа</u> – разработка/обновление моделей данных и схем метаданных для постоянно представленных взаимосвязей и объектов связи <u>Институциональная работа</u> – координация соглашений по семантике взаимосвязей внутри определенных сообществ или отдельных применений
Сохранение взаимосвязей	<u>Техническая работа</u> – сочетание сканирования ссылок и средств сетевого архивирования <u>Институциональная работа</u> – курирование информации о взаимосвязях во времени, итеративно обновленные взаимосвязи как необходимость поддерживать потребности пользователя	<u>Техническая работа</u> – разработка/адаптация пакетных средств, гарантирующих сохранность связей между ресурсами во времени <u>Институциональная работа</u> – 1. разработка подходов по поддержке для агрегаторов взаимосвязей. 2. курирование информации о взаимосвязях во времени

Прямой контакт и работа с создателями ресурса по установлению информации об отношениях в существующих ресурсах также имеет очевидные ограничения, и, вероятно, будет скорее всего происходить только в контексте особых проектов или применений при необходимости в данной информации. Например, в рамках нашего проекта EarthCollab изучались протоколы запросов ученых непосредственно по сбору информации о связях между данными и научными статьями, поскольку сами статьи, как правило, не содержат достаточной для нас информации, чтобы уверенно подтвердить сами отношения.

Таблица отражает различного рода работу, необходимую для поддержки широкого связывания научных ресурсов, фокусируясь на прошлом и будущем. Каждая ее ячейка далее разбивает потребности работы на технические и институциональные. Как обсуждалось ранее [37], курирование научных ресурсов содержит нечто большее, чем просто техническая работа. Институциональные факторы, такие как нормы и ожидания сообщества, доступность посредников в поддержке кураторской

работы, стандарты разработок и адаптации и установленный порядок, – всё играет важную роль в определении успеха кураторских усилий. Таким образом, в табл. подчеркивается, что и технические, и институциональные разработки важны для получения надежных инфраструктур связывания научных ресурсов. Таблица, кажется иллюстративной, но не исчерпывающей. Отличные от авторской точки зрения работы, цитируемые в статье, представляют дополнительные возможности акцентировать внимание на имеющейся и будущей работе в указанной сфере.

ЗАКЛЮЧЕНИЕ

Попытки связать научные ресурсы надежными и устойчивыми способами сталкиваются с многими сложными проблемами, распространяющимися на технические и организационные факторы. В статье обсуждаются три проекта, каждый из которых сфокусирован на одном или нескольких аспектах этих проблем. В ней синтезируются основные результаты и уроки, усвоенные в ходе этих проектов, и представлены принципы работы

четырёх основных сфер потребностей, главным образом, вопросов, касающихся потребности в идентификации, подтверждении, характеристике и сохранении информации о взаимосвязях между научными ресурсами. Также подчеркивается, как попытки построить связывающие инфраструктуры будущего сталкиваются с разными проблемами в отличие от инициатив по сбору и характеристике связей между уже существующими данными, программным обеспечением, публикациями и иными научными ресурсами.

Одновременный запуск разных проектов позволил нам использовать подход «параллельных усилий», чтобы понять, какого рода инициативы больше подходят для решения конкретных задач, относящихся к связыванию научных ресурсов. Удовлетворение локальных потребностей, например, желание связать две конкретные локальные системы, может быть найдено через ограниченную техническую трудность. Разработка перекрестных организационных и перекрестных институциональных решений по поддержке связей ресурсов, тем не менее, требует стандартов и координационной работы и на это потребуются длительное время. Ни одна техническая система или инфраструктура скорее всего не сможет обеспечить общее решение описанных в статье проблем из-за вовлеченности широкого спектра заинтересованных сообществ. Итерация и институциональная работа будут ключом к решению.

Информационные науки имеют опыт и научно-исследовательскую возможность стать крупными вкладчиками в предстоящие попытки связать научные ресурсы. Проблемы, затронутые в этой статье, касаются научной коммуникации, подходов к работе с метаданными, библиометрии, цифрового хранения, сетевой архитектуры, — которые все являются и исторически, и в настоящее время областями исследования в рамках дисциплин библиотекведения и информатики. Инфраструктуры связывания научных ресурсов должны будут найти способы усилить результаты и средства, относящиеся ко всем этим областям, чтобы успешно продвигаться вперед.

Благодарность. Хочу выразить признательность многочисленным сотрудникам Национального центра атмосферных исследований, Корнельского университета и UNAVCO, внесшим вклад в рассмотренные три проекта. Описанная в статье работа проводилась при поддержке грантов Национального фонда научных исследований США (гранты №№ 1449668, 1448480, 1440213 и 1440181). Национальный центр атмосферных исследований финансируется Национальным фондом научных исследований.

ЛИТЕРАТУРА

1. *Mitroff I.I.* Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists// *American Sociological Review*. — 1974. — Vol. 39, No. 4. — P. 579-595. — <https://doi.org/10.2307/2094423>.
2. *Fienberg S.E., Martin M.E., Straf M.L.* (Eds). Part I: Report of the Committee on National Statistics. In *Sharing Research Data*. — Washington, D.C.: National Academy Press, 1985. — <http://www.nap.edu/catalog/2033/sharing-research-data>
3. *Borgman C.L.* The conundrum of sharing research data// *Journal of the American Society for Information*

Science and Technology. — 2012. — Vol. 63, No. 6. — P. 1059-1078. — <https://doi.org/10.1002/asi.22634>

4. *Mayernik M.S.* Open data: Accountability and transparency// *Big Data & Society*. — 2017. — Vol. 4, No. 2. — <https://doi.org/10.1177/2053951717718853>

5. *Willinsky J.* The unacknowledged convergence of open source, open access, and open science// *First Monday*. — 2005. — Vol. 10, No. 8. — <http://ojphi.org/ojs/index.php/fm/article/view/1265/1185>

6. *Woelfle M., Olliaro P., Todd M.H.* Open science is a research accelerator// *Nature Chemistry*. — 2011. — Vol. 3, No. 10. — P. 745-748. — <http://doi.org/10.1038/nchem.114>

7. *Kriesberg A., Huller K., Punzalan R., Parr C.* An analysis of federal policy on public access to scientific research data// *Data Science Journal*. — 2017. — Vol. 16: paper 27. — <https://doi.org/10.5334/dsj-2017-027>

8. *Hitchcock S.* The effect of open access and downloads ('hits') on citation impact: A bibliography of studies. — University of Southampton, 2013. — <https://eprints.soton.ac.uk/354006/>

9. *Pinowar H.A., Vision T.J.* Data reuse and the open data citation advantage// *PeerJ*. — 2013. — Vol. 1: e175. — <https://doi.org/10.7717/peerj.175>

10. *Van de Sompel H., Payette S., Erickson J., Lagoze C., Warner S.* Rethinking scholarly communication// *D-Lib Magazine*. — 2004. — Vol. 10, No. 9. — <https://doi.org/10.1045/september2004-vandesompel>

11. *Pepe A., Mayernik M., Borgman C.L., Van de Sompel H.* From artifacts to aggregations: Modeling scientific life cycles on the semantic web// *Journal of the American Society for Information Science and Technology*. — 2010. — Vol. 63, No. 3. — P. 567-582. — <https://doi.org/10.1002/asi.21263>

12. *Kent W.* Data and Reality: Basic Assumptions in Data Processing Reconsidered. — New York: North-Holland, 2018.

13. *Bean C.A., Green R.* (Eds). *Relationships in the Organization of Knowledge*. — Boston, MA: Kluwer, 2001.

14. *Borgman C.L.* *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press, 2007.

15. *Lagoze C.J.* *Lost Identity: The Assimilation of Digital Libraries into the Web*. Ph.D. diss. — Cornell University, 2010.

16. *Van de Sompel H., Nelson M. L.* Reminiscing about 15 years of interoperability efforts// *D-Lib Magazine*. — 2015. — Vol. 21, No. (11/12). — <http://doi.org/10.1045/november2015-vandesompel>

17. *Mayernik M.S., Phillips J., Nienhouse E.* Linking publications and data: Challenges, trends, and opportunities// *D-Lib Magazine*. — 2016. — Vol. 22, No. (5/6). — <https://doi.org/10.1045/may2016-mayernik>

18. *Mayernik M.S., Hart D.L., Maull K.E., Weber N.M.* Assessing and tracing the outcomes and impact of research infrastructures// *Journal of the Association for Information Science and Technology*. — 2017. — Vol. 68, No. 6. — P. 1341-1359. — <https://doi.org/10.1002/asi.23721>

19. *Silvello G.* Theory and practice of data citation// *Journal of the Association for Information Science and Technology*. — 2017. — Vol. 69, No. 1. — P. 6-20. — <https://doi.org/10.1002/asi.23917>

20. *Mayernik M.S., Maull K.E.* Assessing the uptake of persistent identifiers by research infrastructure users// *PLoS ONE*. — 2017. — Vol. 12, No. 4, e0175418. — <https://doi.org/10.1371/journal.pone.0175418>

21. Borner K., Conlon M., Corson-Rikert J., Ding Y. (Eds.). VIVO: A Semantic Approach to Scholarly Networking and Discovery. — San Rafael, CA: Morgan & Claypool, 2012.
22. Mayernik M.S., Gross M.B., Corson-Rikert J., Daniels M.D., Johns E.M., Khan, H., Stott D. Building geoscience Semantic Web applications using established ontologies// Data Science Journal.— 2016.—Vol. 15, article 11.— P. 1-10. — <https://doi.org/10.5334/dsj-2016-011>
23. Lenfle S., Loch C. Lost roots: How project management came to emphasize control over flexibility and novelty// California Management Review. — 2013.— Vol. 53, No. 1.— P. 32-55. — <https://doi.org/10.1525/cmr.2010.53.1.32>
24. Carlson J. The use of life cycle models in developing and supporting data services// J.M. Ray (Ed.), Research Data Management: Practical Strategies for Information Professionals (pp. 63-86). — West Lafayette, IN: Purdue Univ. Press, 2014.
25. Lenhardt W.C., Abalt S., Blanton B., Christopherson L., Idaszak R. Data management lifecycle and software lifecycle management in the context of conducting science// Journal of Open Research Software.— 2014.— Vol. 2, No.1, e15. — <https://doi.org/10.5334/jors.ax>
26. Sheth A., Ramakrishnan C., Thomas C. Semantics for the Semantic Web: The implicit, the formal and the powerful// International Journal on Semantic Web and Information Systems. — 2005. — Vol. 1, No. 1. — P. 1-18. — <https://doi.org/10.4018/jswis.2005010101>
27. Michener W.K., Brunt J.W., Helly J.J., Kirchner T.B., Stafford S.G. Nongeospatial metadata for the ecological sciences// Ecological Applications — 1997. — Vol. 7, No. 1.— P. 330-342.
28. Edwards P.N., Mayernik M.S., Batcheller A., Borgman C.L., Bowker G.C. Science friction: Data, metadata, and collaboration in the interdisciplinary sciences// Social Studies of Science. — 2011. — Vol. 41, No. 5.— P. 667-690. — <https://doi.org/10.1177/0306312711413314>
29. Burton A., et al. The data-literature interlinking service: Towards a common infrastructure for sharing data-article links// Program. — 2017.— Vol. 51, No. 1.— P. 75-100. — <https://doi.org/10.1108/PROG-06-2016-0048>
30. Burton A., et al. The Scholix Framework for interoperability in data-literature information exchange// D-Lib Magazine. — 2017. — Vol. 23, No. (1/2). — <https://doi.org/10.1045/january2017-burton>
31. Thompson H.S. What is a URI and why does it matter? // Ariadne. — 2010. — Vol. 65. — <http://www.ariadne.ac.uk/issue65/thompson-hs/>
32. Duerr R., Downs R., Tilmes C., Barkstrom B., Lenhardt W.C., Glassy J., Bermudez L., et al. On the utility of identification schemes for digital earth science data: An assessment and recommendations// Earth Science Informatics. — 2011, Vol. 4, No. 3. — P. 1-22. — <https://doi.org/10.1007/s12145-011-0083-6>
33. Mayernik M. S. Bridging data lifecycles: Tracking data use via data citation workshop report. NCAR Technical Note, NCAR/TN-494+PROC. — Boulder, CO: National Center for Atmospheric Research (NCAR), 2013.— <https://doi.org/10.5065/D6PZ56TX>
34. Tilmes C., Fox P., Ma X., McGuinness D.L., Privette A.P., Smith A., ... Zheng J.G. Provenance representation for the National Climate Assessment in the Global Change Information System// IEEE Transactions on Geoscience and Remote Sensing. — 2011. — Vol. 51, No. 11. — P. 5160-5168. — <https://doi.org/10.1109/tgrs.2013.2262179>
35. Baker K.S., Duerr R.E., Parsons M.A. Scientific knowledge mobilization: Co-evolution of data products and designated communities// International Journal of Digital Curation. — 2015. — Vol. 10, No. 2.— P. 110-135. — <http://doi.org/10.2218/ijdc.v10i2.346>
36. Marchionini G., Lee C.A., Bowden H., Lesk M. (Eds.). Curating for Quality: Ensuring Data Quality to Enable New Science. Final Report: Invitational Workshop Sponsored by the National Science Foundation, Sept. 10-11, 2012, Arlington, VA. — 2012. — http://openscholar.mit.edu/sites/default/files/dept/files/altman2012-mitigating_threats_to_data_quality_throughout_the_curation_lifecycle.pdf
37. Mayernik M.S. Research data and metadata curation as institutional issues// Journal of the Association for Information Science and Technology. — 2016. — Vol. 67, No. 4. — P. 973-993. — <https://doi.org/10.1002/asi.23425>

Приглашаем российских и зарубежных авторов к сотрудничеству
в журнале «Международный форум по информации».
Оригинальные статьи и другие материалы (рецензии, письма)
можно присылать на русском или английском языке
по почтовому адресу, указанному в «Памятке для авторов»
или по электронной почте: mfi@viniti.ru.

Ответственный за выпуск *Л. В. Кобзева*

Компьютерная верстка *М. А. Филимонова*

ИД № 04689 от 28.04.2001 г.

Подписано в печать 16.03.2020 г.

Бумага офсетная. Формат 60х84 1/8. Гарн. литер. Печать цифровая

Усл. печ. л. 4,00 Уч.-изд. л. 4,22 Тираж 33 экз.

Адрес редакции: 125190, Россия, г. Москва, ул. Усиевича, д. 20

Тел. (499) 155-44-95