

СОДЕРЖАНИЕ

Николас Д. Надежность и авторитет в научных коммуникациях	3
Монжон Ф., Ларивьер В. Сотрудничество, за которое приходится платить дорогой ценой: влияние научного обмана на карьеру соавторов	11
Гёдерт В. Фасеты и типизированные отношения как инструменты для процессов рассуждения в информационном поиске	20
Лакович Д., Юрич М. Отчет о библиотеках в цифровом веке (LIDA 2014)	29
Авторский указатель к МФИ, 2014, т. 39	31

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Академик РАН **Ю.М. Арский** (Российская Федерация) — *главный редактор*,
ВИНИТИ РАН, 125190, Москва, ул. Усневича, 20. Телекс 411249

Проф. д-р. **Р.С. Гиляревский** (Российская Федерация) — *заместитель главного редактора*,
ВИНИТИ РАН, 125190, Москва, ул. Усневича, 20. Телекс 411249

С. Дж. Паркер (Канада) — *заместитель главного редактора*, IDCR, P.O. Box 8500,
Ottawa, Ontario K1G 3H9, Canada

А. Джикарайст (Великобритания) — CURA Consortium and GAVEL g.e.i.e,
38 Ship Street, Brighton BN1 1AB, UK

М. Дрейк (США) — Технологический институт шт. Джорджия, Библиотечный
и информационный центр, 704 Cherry Street, Atlanta, Georgia 30332-0900, USA

А. де Кемп (Германия) — Издательство “Springer-Verlag”, Postfach 10 52 80,
D-69042 Heidelberg, Germany

Д-р **Т. Кеннон** (Великобритания) — Отдел исследований и разработок
Британской библиотеки, 2 Sheraton Street, London W1V 4BH, UK

М. Миддлтон (Австралия) — Школа информационных систем, QUT Gardens
Point Campus, 2 George Street, Brisbane, 4000 QLD., Australia

Т. Молвиг (Норвегия) — Национальное управление по научной информации,
вузовским и специальным библиотекам, P.O. Box 2439 Solli, N-0201, Oslo,
Norway

Х. Ринкон Феррейра (Бразилия) — Бразильский институт информации по
науке и технике (IBICT), SAS — Quadra 5, Lote 06, Bloco H, 700-70-000 Brasilia
D.F., Brazil

С. Феррейро (Чили) — Чилийский университет, Системы информационных
и библиотечных служб, Casilla de Correo 10D, Santiago, Chile

Проф. **Ю. Фуздивара** (Япония) — Университет Цукуба, Институт электроники
и информатики, Tsukuba-shu, Ibaraki, 305 Japan

Д-р **М. Хименес** (Испания) — Испанское общество по научной документации
и информации, Fuencarral, 123-6° dcha., 28010, Madrid, Spain

Надежность и авторитет в научных коммуникациях*

Дэвид НИКОЛАС
(David NICHOLAS)

Колледж коммуникации и информации,
Университет шт. Теннесси, США

Общая цель статьи заключается в изучении того, насколько появившиеся виды цифрового поведения перспективны и, вероятно, меняют понятия надежности и влияния в научной среде. Обсуждается, как ученые определяют и проверяют авторитет и достоверность научных источников и каналов, которые они выбирают для использования, цитирования и публикации. Такими источниками/каналами являются журналы, сетевые сайты, базы данных и социальные медиа. В частности, анализируется, как ученые ориентируются в современной насыщенной, динамичной, разнообразной и неопосредованной научной цифровой среде, где все труднее установить качество, достоверность, авторство и авторитетность (значимость) просматриваемой информации. Также исследуется, сказывается ли рост в использовании соцмедиа и публикаций открытого доступа в научных целях на традиционных практиках установления авторитета и надежности информационных источников и каналов. Это исследование в целом является международным, хотя данная статья фокусируется на ситуации в США и Великобритании и в основном на науке и социологии. Три методологии (группы интереса, крайне необходимые случайные интервью и опросы) были использованы в более широком исследовании, а здесь мы сконцентрируемся на результатах интересующих нас групп, т.е. применяется первая методология. Выводы показывают, что традиционные критерии доверия — рецензируемые журналы, импакт факторы и персональные знания, а также сети — до сих пор занимают значительное место, хотя в использовании молодыми учеными прогрессируют соцмедиа.

ВВЕДЕНИЕ

Доверие, авторитет и надежность являются основным девизом научной коммуникации; система научной коммуникации опирается на гарантию качества. Однако все может меняться. Таким образом, цифровая революция в научных коммуникациях, насчитывающая сегодня более десятка лет, но все еще находящаяся в развитии (впереди у нее долгий путь становления), начинает «мутить воду относительно доверия», возможно проверяя некоторые долго процветающие взгляды и практики. В сегодняшней насыщенной, динамичной, разнообразной, многоплатформной и неопосредованной научной цифровой среде становится все труднее определить, чья реально эта информация, а также можно ли ей доверять. Все возрастающее соперничество среди ученых, усиливающееся ростом числа ученых и глобального рынка, по-видимому, приводит к опоре на уполномоченных по гарантии качества и метрики, такие как импакт факторы,

и к использованию этих метрик. И в таком случае, конечно, имеет место появление некоторого очень перспективного информационного поиска и использование поведения, унаследованного от поколения Google (рожденных в цифровой век людей) и пользователей смартфонов.

Чтобы понять, что происходит в результате этих изменений и возможных проблем, исследовательская группа CIBER и Университет шт. Теннесси при финансовой поддержке фонда А. Р. Sloan Foundation заняты сегодня изучением того, насколько появившиеся цифровые поведения/платформы перспективны и как они меняют понятия доверия и влияния в научном мире. В частности, они хотят установить:

- Как университетские ученые определяют и проверяют авторитет и надежность научных источников и каналов, которые они выбирают для использования, цитирования и распространения.
- Сказываются ли соцмедиа и открытый доступ на традиционных практиках установления авторитета и надежности.

*Перевод Nicholas D. Trust and authority in scholarly communications//VII Encontro Ibérico EDICIC, Porto (Portugal), 4-6 November 2013. — 2013.— <http://eprints.rclis.org/22903>

Это исследование является глобальным, но с особым акцентом на ситуации в Великобритании и США, представляющих собой доминирующих игроков в области научной коммуникации. Тематический уклон в основном делается на университетских ученых в науке и социологии, хотя некоторые данные гуманитарных областей собирались в целях проверки. Исследование проводилось в октябре 2012 г. - ноябре 2013 г. (см. <http://cics.cci.utk.edu/cicsprojects/Sloan>).

МЕТОДОЛОГИЯ

Совокупность методов использовалась для этого проекта в целом:

а) Было организовано четырнадцать интересующих нас групп в Великобритании и США, чтобы очертить границы исследования и определить важные проблемы и вопросы. Как правило, в каждой интересующей нас группе принимало участие 6-9 ученых. Участники отбирались в целях обеспечения равновесия по половой принадлежности, сроку работы по тематике и должности.

б) Было проведено 80 крайне необходимых случайных интервью один на один в Великобритании и США относительно недавно опубликованной учеными статьи с целью получить подробные детали их поведения. Интервью велись лицом к лицу, а также по телефону или по Скайпу.

в) Что касается потребности в информации от интересующих групп и интервью, то опрос был разработан с помощью программы Survey Monkey. Используя списки электронной почты издателей (шесть основных сотрудничающих международных издателей), приглашение участвовать в опросе было отправлено авторам по всему миру. Ответили почти 3700 человек.

Ко времени написания статьи крайне необходимые случайные интервью и изучение данных опроса не были полностью завершены, поэтому представленные здесь данные являются в основном результатами 14 интересующих нас групп, включающих около 100 ученых из Великобритании и США.

ОБЗОР ЛИТЕРАТУРЫ

С попыткой проведения научного исследования, неизменно перемещаемого в цифровую среду с репутацией неясного характера множества информации, которую она может предложить, беспокойства относительно качества и надежности интернет-источников и каналов быстро принимают угрожающие размеры. То есть, когда предполагается синдром «публикуйся или погибнешь», связанный со значимости и зависимости некоторого знания, производимого и передаваемого в эти дни, признан эндемическим. В таком случае неудивительно, что возможность неестественных изменений в качественном обеспечении потребностей научного сообщества, свободно определяемых и понимаемых как потребности, имеющие отношение к практикам и процедурам, которые систематически отслеживают различные аспекты процессов научной коммуникации в целях обнаружения, исправления и гарантии удовлетворения стандартов качества, кажется в самом деле вполне реальной.

В современной цифровой среде, трансформируемой в том виде как она есть за счет новых технологий, вместе называемых «web 2.0», качество, достоверность и надежность информационных источников и каналов трудно определить. Это в основном происходит из-за того, что рога изобилия доступной в сети научной информации не только с жадностью всеми используется, но также постоянно пополняется каждым, начиная от правитель-

ственных, научных или коммерческих контент-провайдеров, включая общественность; наши традиционные понятия о том, что может считаться высококачественным, авторитетным, достоверным, достойным уважения и поэтому вызывающим доверие и надежным, обязательно подвергаются радикальным изменениям. Таким образом, авторы [1] полагают, что современные потребители информации, сталкиваясь, как это обычно и происходит, с большим сомнением относительно того, кому и чему можно доверять, а также с проблемой, кто или что отвечает за информацию, которую они встречают, должны выработать новые навыки и стратегии оценки информационного источника или канала.

В то же самое время значительно измененная научная культура в современных, движимых рынком системах высшего образования также, кажется, должна подвергнуть риску способность ученых доверять информационным источникам и сайтам, которые они используют в научной работе. Эта измененная научная культура, неизбежное следствие призыва к университетам стать прибыльными интеллектуальными предприятиями, обеспечивающими прямые и эффективные ответы на потребности производства и рынка труда в ответ на финансовую поддержку [2-4], предъявляет научной продуктивности наивысшее приоритетное требование в современной научной среде. Созданная атмосфера, в которой навязанные извне нормы пытаются поддерживать и давать оценку научной продуктивности, в дальнейшем расширяет строгий, касающийся ментальности, диктат, – «публикуйся или погибнешь» [5], который сегодня на какое-то время управляет научными усилиями. Что касается научного успеха в рамках продвижения в карьере, достижения и поддержания хорошей репутации и сохранения финансирования дальнейших проектов, все более зависящих от качества статей, публикуемых в высокоранжируемых журналах, и числа ими получаемых ссылок [6-7], то Артуро Касадевалл, главный редактор *mBio*, высказывание которого процитировано в недавней статье *New York Times* [8], вероятно, слишком прав, говоря, что наука превращается в «игру, в которой победитель получает все, с порочными стимулами, которые заставляют ученых идти напрямик и в некоторых случаях совершать недобросовестные поступки».

РЕЗУЛЬТАТЫ

Сначала я коснусь наблюдений, охватывающих все три основных вида научной деятельности и связанных с ними (публикация/распространение, цитирование и чтение/использование), которые изучались, а затем рассмотрю каждое из них в отдельности.

ОБЩИЕ ВОПРОСЫ

Характеристики надежного и достоверного источника/канала

В интересующих нас группах несколько ученых действительно в первую очередь объяснили то, что они считают надежностью, качеством и достоверностью. Обсуждения, как правило, перерастали в дебаты об относительных новых достоинствах индивидуального определения надежности и качества в отличие от испытанных методов (например, метрик цитирования). Вообще достоверность (*trustworthiness*), слово, которое они в основном предпочитали использовать, сводилось к репутации автора или журнала области их научной специализации. Качество наилучшим образом могло бы определяться с помощью личного наблюдения и оцен-

ки, но из-за нехватки времени, слишком большого объема материала для попадания к адресату и строгой критики руководителей и учреждений (например, относительно публикации в журналах с высоким импакт фактором) оно не всегда осуществимо.

Тогда как почти каждый полагает, что достоверность и качество являются важными вопросами, никто в действительности не думает, что они представляют большую или ключевую проблему. Эти проблемы, по-видимому, могут быть сняты. В конце концов, большинство ученых хорошо ориентировались в науке и имели большой опыт оценки доступной литературы. Выгодность и прагматизм, безусловно, играют свою роль в современной, испытывающей давление и быстро развивающейся (меняющейся) научной среде; иногда нужно просто ухватиться за идею того, что полезно.

Действующие лица

Научные сотрудники выступают в разных ролях – авторы, редакторы, рецензенты, те, кто осуществляет цитирование и пользователи/читатели данной литературы, и они могут выдвигать различные надежные оценки, чтобы выполнять каждую функцию. Однако, когда они говорят о доверии и авторитете в научных коммуникациях, они редко предвзвешивают то, о чем они говорят, констатируя, что например, как пользователь или цитирующий я поступаю так. Возможно это из-за того, что когда они думают над этими вопросами, они обычно думают о них в связи с каким-либо источником – например, журналами; для ученых журналы служат носителем качества.

Редакторы в частности, кажется, придерживаются иной позиции; их точка зрения определялась собственным взглядом на себя как на информационных провайдеров, людей, предлагающих другим заслуживающие доверия массивы документов. Они оказались более традиционными и ортодоксальными в своих взглядах. Они считали, что существующая система действует хорошо (по шаблону), и что все остается таким же, каким это было всегда. Для них это настоящий предмет кризиса, какого кризиса? Те, кто находится вне рамок журнальной системы, весьма вероятно были готовы выступить с критикой существующей научной системы.

Что касается различий между распространением, поведением в цитировании и в чтении, то ученые больше выделяют первый вид деятельности и меньше – использование. В сфере распространения основной целью является появление в высокоранжируемом журнале, и ученые весьма этим обеспокоены. Поведение в цитировании стремится быть строже, более сфокусированным и сложным, чем в использовании. Ученые явно имеют значительно больше свободы относительно того, что они используют – они могут использовать блоги, публикации открытого доступа и т.д. – в дополнение к основному контенту, но они, как правило, не будут цитировать такие источники.

Рецензирование

Казалось, что рецензирование обычно обеспечивает определенную степень уверенности относительно качества продукта. Оно показывает, что кто-то приложил усилия и что его усилия подтверждены сообществом ученых. Поэтому оно является важным научным атрибутом доверия, стимулирующим ученых на поиск, использование, цитирование и распространение.

С одной стороны, ученые хотят публиковаться в журналах, имеющих надежные механизмы рецензирования (несмотря на душевную боль и разочарования,

сопровождаящие отказ и критику), а с другой стороны, они не хотят беспокоиться по поводу своего цитирования и цитирования рецензируемого контента.

Хотя существует сильная привязанность к рецензированию, большинство людей начинают свое выражение доверия к рецензированию с признания того, что имеются проблемы в способе его осуществления. Однако, когда вы пытаетесь углубиться в действительные механизмы рецензирования, то вообще обнаруживается очень мало согласия относительно всего того, что могло бы его улучшить. Всего лишь частица демократии, как предположил один ученый.

До сих пор мы обсуждаем рецензирование в редакции, но еще одна форма рецензирования беспокоит многих британских ученых – развивающийся подход научного превосходства (Research Excellence Framework - REF), правительственный эксперимент по поощрению научного качества в университетах Великобритании (<http://www.ref.ac.uk>). Утверждалось, что самое большое влияние, которое имел REF, заключалось в создании системы *учрежденческого* рецензирования в университетах по всей Великобритании, системы, по которой ученые не только оцениваются при представлении публикации, но и отслеживаются во времени.

Метрики

Имелось некоторое свидетельство, показывающее, что научный мир начал развиваться и управляться с помощью алгоритмов и последствие этого выражается в том, что творчество и новые идеи вытесняются высокой основанной на метриках культурой, стандартизирующей поведение научной коммуникации и делающей научные статьи менее интересными.

Были некоторые различия в этом отношении в зависимости от дисциплины ученого. Большинство ученых в значительной степени не задавали вопросов о важности метрической системы; социальные ученые были немного обеспокоены, но чувствовали, что нет никакого реального выбора, кроме как примириться с этим; ряд ученых гуманитарных областей в интересующих нас группах явно чувствовали дискомфорт в плане культуры и отчуждение, но с дискомфортом или без него, они были очень большой её частью.

Самый неожиданный пример, касающийся ученых, руководствующихся в своей работе метриками, происходит не из науки, как можно было ожидать, а из сферы бизнеса/экономики, где ученым ясно говорят, где публиковаться и что цитировать согласно руководству по качеству научных журналов, выпускаемому Ассоциацией бизнес-школ (<http://www.associationofbusinessschools.org/content/abs-academic-journal-quality-guide>). Если журнал приводится в списке данного руководства, тогда все обстоит благополучно для представления статьи, а если нет, тогда инструкция не рекомендует представлять туда статью, так как это не будет способствовать продвижению в карьере.

Начинающие карьеру ученые в областях социальных и гуманитарных наук скорее стоят особняком в этой группе и утверждают, что считают себя «рабами» по отношению к сфокусированной, основанной на метриках, журнальной системе; они вынуждены придерживаться правил, чтобы подняться по карьерной лестнице, но полагают, что лестница сломана. На самом деле один из членов интересующей нас группы выразил это даже более решительно: *Журналы являются демонстрацией всего «плохого», что есть в системе научной коммуникации. Да все это безумие, но я здесь не для того, чтобы преобразовывать научное сообщество.*

Соцмедиа

Только несколько человек, большей частью молодые ученые, думают, что соцмедиа являются не просто социальными и преходящими. Для большинства других ученых не было вопроса относительно того, что им следует доверять, и, безусловно, маловероятно, что они угрожают рецензируемому журналу. Тем не менее, ученые, хотя и выражающие отсутствие заинтересованности в самих соцмедиа, говорят, что знают людей, которые используют соцмедиа, и чувствуют себя виновными в недостаточном их привлечении. Их неиспользование и общее отсутствие интереса могут в основном приписываться тому факту, что соцмедиа не содержат заслуживающих доверия источников. Однако существуют и другие причины:

- а) многие ученые, с которыми мы говорили, были новичками в вопросах, касающихся соцмедиа;
- б) они вообще были враждебно настроены в отношении соцмедиа, что-то в них их раздражало;
- в) они говорили, что не имеют свободного времени, чтобы экспериментировать с соцмедиа;
- г) они отклонили саму идею идти этим путем в пользу современного общественного настроения по вопросам высшего образования в Великобритании и США, отдающего предпочтение распространению материала в высокоранжируемых журналах;
- д) неформальный язык соцмедиа не подходит для научного дискурса.

Имелось общее признание того, что соцмедиа хорошо подходят для некоторых аспектов, особенно для: а) получения новых/различных идей, а следовательно и стимула; б) самопродвижения публикаций особенно в отношении распространения – сферы практиков и общестственности. Что в наибольшей степени заинтересовало участников (при разговоре об этом) так это тот факт, что соцмедиа могут способствовать продвижению их работы путем увеличения числа ссылок, получаемых их публикациями.

Однако начинающие карьеру ученые интенсивно использовали соцмедиа, но боялись охватить их в полной мере и скрывали их использование от своих старших коллег, которые не одобряли такое поведение. Для начинающих ученых имеется много научной выгоды в использовании соцмедиа:

- а) помощь в разработке персональной сети,
- б) облегчение сотрудничества,
- в) поиск коллег для совместной работы (в реальном времени),
- г) пребывание в курсе того, что происходит,
- д) отслеживание авторов, чьими работами они заинтересовались
- е) облегчение поиска приверженцев особых точек зрения.

Открытый доступ

Ученые склонны относиться с недоверием к публикациям открытого доступа, как с точки зрения автора, так и читателя. Недоверие значительно снижается (но не уходит совсем), когда журналы открытого доступа публикуются авторитетным издателем, как в случае с Sage. Удивительно, но некоторые ученые отдавали себе отчет в инициаторских попытках Plos One (<http://www.plosone.org>), показывающих, что журналы открытого доступа могут стать высокопрестижными, популярными и быстрыми в отношении публикации статей. Хотя, когда им рассказали историю успеха PLOS

One, большая часть ученых проявила к нему значительный интерес.

Ученые из университетов интенсивного обучения были более позитивно настроены относительно распространения открытого доступа главным образом на основании того, что он стал более открытым и доступным. Они также почувствовали, что публикация их статей в журналах открытого доступа может помочь их карьерному росту. Этот аргумент состоит в следующем: университеты могут ранжировать журналы импакт фактора более высоко, а поисковая система Google ранжирует публикации открытого доступа выше в списках поисковых ответов, поэтому если вы хотите оказаться в списке первым, чтобы впечатлить «охотников за головами», тогда публикация открытого доступа может дать вам преимущество.

Начинающие карьеру ученые также в принципе любят открытый доступ опять же за его открытость и доступность, но они боятся им пользоваться, поскольку чувствуют, что научный мир не принял решение о нем. Если они публикуются в журнале открытого доступа или цитируют его, то, как говорится, они «могут оседлать неправильную лошадь» и обнаружить, что опубликовались во второсортном журнале. Как сказал один ученый, *вероятно угроза репутации.*

Стало ли положение лучше или хуже спустя десять лет?

Ученые признали, что вокруг было много «плохого», так как система научной коммуникации только сейчас становится более доступной и существует гораздо больше возможностей опубликоваться; но в целом они почувствовали, что со временем качество действительно выросло. Рост относительно качества означает, что они могут смириться с увеличением *плохого и скучного материала.*

Рост качества в целом является результатом того, что больше людей приходит в данную область науки, а с этим появляется и больше уровней соперничества. Считалось, что рост качества в основном имеет место в ведущих журналах, и что именно ниша специальных журналов получает полную выгоду от обилия обычного материала. Поэтому журналы открытого доступа не представляют рост плохого или среднего качества контента, а большой рост наблюдается в нише подписных журналов.

Сегодня существует огромное море материалов посредственного качества, так как стало проще публиковаться, а чем выше уровень, тем лучше качество благодаря более совершенному обучению, большей конкуренции и поощрениям за публикации.

Использование и чтение

Ученые действительно являются «клавишами тихого звучания», когда речь идет о вопросах доверия и влияния относительно использования и чтения научного контента. Они занижают проблемы использования и доверия, так как, если «что-либо», в чем они заинтересованы, отсутствует, то их давно зарекомендовавшие себя персональные сети им об этом скажут; у них нет необходимости это искать. Как прокомментировал один ученый, *Именно интуиция подсказывает, куда вам обратиться за информацией. Но в основе этого лежит все, что касается сетевых связей людей, которые вы развивали годами.*

Другой ученый говорит от имени большинства, когда объясняет, как он определяет то, что заслуживает внимания. *Если я не знаю автора, тогда надо быть внимательным; более того, если вы не знаете принадлежность учреждения, будьте еще внимательнее, а если вы не знаете и журнал, тогда определенно это даже не стоит просматривать.*

Импакт фактор журнала не был так важен в определении того, что читать, так как:

а) он обеспечивает слишком узкий взгляд на научную литературу;

б) он привязан к журналу, а не к статье;

в) иногда это лотерея относительно того, какой журнал получит импакт фактор;

г) его наличие означает, что журнал будет атакаться авторами низкого ранга (большинство из развивающихся стран), пытающимися повысить свою карьеру – это общепринятая плата за успешную публикацию;

д) журналы с высоким импакт фактором стремятся к традиционному стилю и им не хватает инновационных и свежих статей;

е) многие ученые просто не знают, какой импакт фактор у данного журнала.

Чем больше научная область, тем более важным является импакт фактор, так как в больших областях, таких как бизнес-исследования, ученые не знают всех авторов и импакт факторы выступают в качестве гарантов качества. По схожим причинам импакт факторы важны в периферийных областях. Одной из привлекательных черт журналов с высоким импакт фактором, с точки зрения использования, является то, что они обычно имеют лучших рецензентов, поэтому статьи в них, как правило, высокого качества, благодаря вкладу рецензентов.

Рецензируемые публикации обычно получают больше доверия, а знание (и доверие) журнала, редактора и редколлегии становится более важным. Но само рецензирование необязательно гарантирует качество, поскольку *вне его существует некий сомнительный рецензируемый материал.*

Ученые не знают о возможностях предложения относительно фильтрации материала с помощью использования факторов или данных загрузки. Однако, когда их информируют о таких возможностях, ученые становятся полностью негативными в своих суждениях. Они чувствуют:

1) Использование метрик является слишком легким делом.

2) Высокоиспользуемые статьи не являются лучшими с точки зрения качества (по оценке мнений редакции).

3) Загрузки не представляют реального количества прочтений – люди загружают материал, который потом не читают.

4) Огромный рост медиа значительно увеличивает профиль статей и утверждается, что он служит искажающим фактором.

Альтметрия обеспечивает иной способ фильтрации использования (<http://altmetrics.org/manifesto>) на основе применения и рекомендаций. Он обеспечивает нетрадиционные критерии научного влияния, основанные на видах деятельности в сетевых средах. Эти метрики отслеживают связанные взаимодействия и деятельность, позволяя ученым получить более полную картину сферы и влияния научного исследования. Альтметрия обычно охватывает книжные закладки, лайки, предпочтения и рекомендации. PloS One широко использует альтметрию, чтобы помочь ученым в отборе контента для чтения.

Ученые отмечали, что они ищут не только качественный или авторитетный контент; иногда они ищут интересный и оригинальный контент. *Имеются случаи, когда авторитет/ранг являются второстепенной задачей в определении использования, когда ученые ищут новый, свежий и творческий контент.* Высказывающимися это учеными являются

ся, как правило, ученые социальных областей науки. Ряд ученых социальных областей упоминал, что «быстрый и грязный» поиск в Google/Google Scholar был особенно продуктивным в предоставлении интересного материала.

Хорошее изложение материала также является определителем того, что прочитывается. Признано, что, если автор имеет явный авторитет, вы потратите время на прочтение его материала независимо от того, насколько хорошо он написан, но для большей части материала у вас есть выбор и этот выбор всегда будет склонять вас в сторону хорошо написанного материала.

Рефераты (резюме) играют очень важную роль в определении того, что читается и чему можно доверять. Слишком много публикуется и слишком мало времени у большинства ученых для чтения того, что доступно, и это приводит к широко распространенному выбору формы «быстрого просмотра» искомой информации, которая включает быстрое движение вдоль цифрового пространства, с частыми легкими контактами или сменной направленности поиска. Такое поведение ведет к зависимости от прокси (рефератов) в целях навигации по большому объему материала. Экранирование и перекрестные сравнения производятся на уровне реферата. Рефераты настолько важны, что один ученый сказал, что хотел бы их видеть с контролируемым качеством, должным образом рецензированными. *Нам необходимо доверять реферату.* Тот факт, что рефераты свободно предлагаются для просмотра, вносит значительный вклад в их значимость.

Хотя наличие полнотекстового доступа может быть важным в получении оценки относительно доверия, ряд ученых на самом деле читают весь текст целиком. Фактически, чтение статьи означает просмотр 5-10% ее содержания.

Были упомянуты другие библиографические и показательные для доверия к контенту импульсные точки: а) методология (более свойственна науке; ученые социальных областей считают ее слишком механистической мерой), б) выводы, в) библиография (кто цитируется?), г) теоретическая позиция в случае социальных наук.

Типы используемых источников

Цифровые источники. Доверие улучшается, если цифровая публикация также доступна в печатной форме, так как это придает ей ощущение реальности. Некоторые ученые чувствуют, что у статьи имеются достоинства, и часть их связана с качеством/надежностью. Таким образом: а) считается, что статья вызывает больше доверия, так как подразумевается, что она менее эфемерна; б) так как имеются четкие ограничения по объему при публикации статьи, то это заставляет автора сконцентрироваться, и, следовательно, выдается лучший вариант. Однако начинающие карьеру ученые придерживаются весьма противоположного мнения; они не могут верить тем, кто так думает, поскольку полагают, что цифровой мир реален, а печатная статья предназначена для музеев.

Данные. Некоторые беспокойства по поводу доверия появились относительно данных из-за предположительного отсутствия рецензирования. Рецензирование считается необходимым, так как невозможно самому решить вопрос о том, надежны данные или нет, учитывая размер и сложность баз данных. Имелось ощущение, что это может помочь, если данные относятся к статье, то статья придаст им авторитет. Имя автора даже более важно в установлении авторитета данных.

Труды конференций. В Великобритании понимают, что труды конференций не приемлемы для использования

подхода научного превосходства (REF), поэтому обычно их широко не рассматривают. Утверждалось, что а) механизмы рецензирования слабы, отсутствуют или не касаются реферата; б) наблюдается *растущее число внешних «акул», позволяющих замять названия конференций* (проходящих под эгидой организаций, больше заинтересованных в получении денег, чем представлении качественного контента); в) несмотря на попытки лучшего индексирования в *Web of Knowledge и других базах*, труды не имеют должной обзримости и положения. Но есть мнение, что в некоторых областях они могут привлекать огромное число ссылок.

Поиск

Большая, широкая «вызывающая доверие» информационная труба, к которой подключаются ученые, не стремится быть публикационной платформой, университетским каталогом или федеральной поисковой системой. На самом деле используются универсальные поисковые системы, такие как Google или Google Scholar. Система Scholar считалась *удивительно* хорошей, по мнению многих ученых. Библиотекари утверждают, что «нахождение материалов» становится сегодня все более трудным, учитывая открытый доступ и учрежденческие архивы, вносящие неразбериху, но ученые с этим не согласны. Они действительно считают, что многое становится более легким.

Библиотеки лишь упоминались учеными, но, когда речь заходила о них, в основном высказывания носили негативный и ностальгический характер. Ученые чувствовали, что роль библиотекарей в основном ограничивается пределами зданий, в которых они работают. Считается, что библиотеки, первоначальные хранители качества, не играют роли в сегодняшней ситуации. Ученые просто не видят библиотеки в качестве отправной точки к информации, которую они ищут. Библиотеки рассматривались как *несовершенные* источники информации, и ученые не доверяют библиотекарям в принятии критических решений относительно того, что, по их мнению, есть и чего не хватает в обнесенном стеной саду.

Ученые упоминали ряд вызывающих доверие (и традиционных) стратегий поиска: 1) следовать ссылкам из достоверных источников и затем проверять рефераты с целью установления их ценности и качества; 2) вводить название классической работы по своей области в Google Scholar и затем смотреть ссылки на нее. Они предпочитают простой поиск – *любой барьер заставит вас превратить поиск необходимого*. Начинающие карьеру ученые придерживаются мнения людей и издателей, но большое различие заключается в том, что они делают это не через базу данных издателя, а через сайты соцмедиа.

Влияние публикаций открытого доступа и соцмедиа

Все ученые, с которыми мы говорили, были получателями выгоды от крупного соглашения, которое даст им легкий доступ к сотням подписных журналов, поэтому открытый доступ не являлся важным. Хотя никто фактически не признался, что скорее не будет использовать материал открытого доступа, даже если бы и хотел, неясно то, как они могли прежде всего видеть различия. В конце концов все издатели подписных журналов также публикуют материал открытого доступа.

Беспокойства, которые ученые выразили по поводу слабого (или отсутствующего) рецензирования статей открытого доступа, не привели к массовой проверке политики рецензирования журналов открытого доступа. Но интересно то, что ряд ученых, которые действительно

но так поступали, вероятно, не делали это для подписных журналов, безусловно об этом не упоминалось.

Научные твиты почти всегда используются для того, чтобы сослаться на журнальную статью. Ученые используют Википедию, они любят ее и с удовольствием об этом говорят. Academe.edu, онлайн сообщество, было упомянуто с позитивной точки зрения для оценки трудного по получению доступа материала; материалу, кроме важного, двери были закрыты.

ПОВЕДЕНИЕ ЦИТИРОВАНИЯ

Поведение цитирования не было настолько либеральным, как поведение использования или чтения. Авторы тщательно взвешивают источники перед тем, как цитировать, обычно отклоняя в процессе соцмедиа. Здесь имеются политические вопросы для рассмотрения; им (авторам) необходимо прикрыть тылы; есть люди, которых вы должны процитировать, чтобы вашу статью приняли; у вас имеется только реальный шанс процитировать себя несколько раз. В самом процессе цитирования имеется много моментов для того, чтобы, как говорится, показать товар лицом.

Был упомянут ряд практик цитирования и требований:

- 1) Цитируйте собственную работу, чтобы увеличить свой h-индекс.
- 2) Цитируйте статьи журнала, в который вы представляете свою статью.
- 3) Рецензенты иногда просят вас процитировать их собственные работы, и вы это делаете ради перспективы опубликования.
- 4) Цитируйте очень высококачественные статьи, так как они поднимают злободневные проблемы и представляют вершину науки (гало-эффект).
- 5) Цитируйте апостериорно – используйте ссылки для подкрепления своей позиции, придавайте вашим идеям больший вес; это особенно важно, если ваши идеи новы или противоречивы.
- 6) Цитируйте первый и самый последний опубликованный источник по теме.
- 7) Цитируйте обзорные статьи как библиографическую стенографию.

Мы знаем, что люди хитрят, когда дело касается ссылок, так как говорят, что не делают этого, но знают тех, кто так поступает.

Начинающие карьеру ученые находились под давлением своих наставников относительно того, чтобы цитировать рецензируемые статьи, а не соцмедиа. *Вы находите интересные материалы где-то еще (т.е. в соцмедиа), но не можете их использовать/цитировать*. Некоторые подходят к этому творчески, цитируя источники соцмедиа как «личные связи». Хотя для большинства ученых соцмедиа представляют собой табу. *Твиттер не используется как источник информации: это может расцениваться как цитирование беседы в баре. Многие блоги – это просто потоки сознания*.

Ученые из университетов интенсивного обучения также различаются по их поведению; многие из них понимают, что не могут цитировать то, что они не читали, и были бы счастливы цитировать все, включая торговые публикации.

ИЗДАНИЕ/РАСПРОСТРАНЕНИЕ

Самой большой и наиболее важной проблемой для ученых был не столько гигантский рост литературы, сколько то, кто отвечает за управление этим ростом, увеличивающееся давление публиковаться. По мнению большинства ученых, эта тенденция создает лавину *плохих/посредственных журнальных статей*. Практика «публи-

куйся или погибнешь» множится. *Имеется огромное количество макулатуры, появляющейся вокруг (научной) системы.* Это проявляется в массовом отклонении статей журналами; 50% составляет «мусор», происходящий отовсюду, но большей частью из развивающихся стран.

Второй самой большой проблемой, свойственной Великобритании, является огромное влияние REF. Как сказал один ученый, REF распространяется повсюду и выглядит убедительным. Политика REF говорит, что оценка ссылок – это еще не все, но никто в это не верит и продолжает стремиться публиковаться в высококачественных журналах. Один ученый спросил, *кто подвергнет сомнению статью из «Nature»?* В конце концов это приводит к негативному влиянию на творчество и извращению понимания того, где статьи действительно должны помещаться.

REF также обвиняется в принуждении людей публиковаться больше, чем бы они хотели, что приводит к очень высоким объемам плохого опубликованного контента. Все больше и больше университетов рассматривают метрики как гарантию качества, полагая, что университеты более низкого ранга присоединяются к ним в этом вопросе, тогда как реально это не происходит.

Издание: рецензирование

Ряд ученых думали, что процесс рецензирования разрушен, хотя начинающие карьеру ученые были, по крайней мере, оптимистично настроены в отношении системы. Для многих ученых издатели несут ответственность за организацию рецензирования, это их центральная роль. Процент отклонения статей считался знаком качества, и ученые буквально были готовы торговаться за столом интересовавших нас групп.

Что касается существующей практики рецензирования, то: а) ученым нравится анонимное рецензирование – рецензенты могут свободно делать замечания по статьям; б) ученые имеют неоднозначные мнения относительно пользы предлагаемых автором рецензентов (недостаток состоит в том, что авторы могут предложить своих друзей; достоинство заключается в том, что можно избежать рецензентов, которых автор подозревал в нечестности (предвзятости)); в) ученым не нравится поверхностное рецензирование, так как тщательное рецензирование приводит к лучшим статьям; г) ученые чувствуют, что открытое рецензирование сдерживает рецензентов; д) ученые не выражают уверенности относительно рецензирования после публикации; е) ученые осознают, что окончательными судьями должны быть редакторы; им следует быть справедливыми и не всегда обращать внимание на замечания рецензентов; ж) ученые считают, что редакторы должны действовать в качестве выпускного клапана процесса рецензирования в случае неудачи, чтобы сгладить различие во мнениях, внести свежесть и инновацию; з) ученые, тем не менее, полагают, что рецензенты улучшают статью, даже если они ее отклоняют. *Так что стоит представлять статьи в «Nature», даже если у вас нет шанса быть принятым, только для того, чтобы получить качественную обратную связь.*

Основной слабостью данной универсально используемой системы является то, что она слишком медлительна. Существовало общее мнение, что есть потребность в получении решения в течение двух месяцев. Слабостью издателей, эксплуатирующих открытый доступ, было получить полное преимущество (если не в реальности, то хотя бы в своей рекламе). Также проблемой было качество рецензирования, которое считалось слишком разнообразным ради обес-

печения комфорта, и дело доходило до давления на рецензентов, чтобы они проделали работу быстрее; как следствие *страдает качество.*

Плагиат и неэтические практики

Со шквалом статей приходит некорректный контент. Считается, что он должен быть более традиционным и широко распространенным, чем обычно предполагается, хотя данные заслуживающего доверия издателя, показанные автору статьи, демонстрируют уровни плагиата и дублирования, которые, как обнаружилось, в последние 3 года – снизились от 30% до 10% рукописей.

Хотя начинающие карьеру ученые соглашались, что плагиат – это табу, они, кажется, менее антагонистичны по отношению к ограниченному поведению – «вырезать и вставлять», обеспечивающему данные права. Несмотря на то, что ученые из университетов интенсивно обучаются также не поддерживали плагиат, они не столь критичны как ученые из университетов интенсивных исследований. Они полагали, что это нечеткая область и понимали причины/давления его осуществления. Самоплагиат они, по-видимому, считали, менее серьезным нарушением, а, может быть, и вовсе не нарушением.

Также обстоит дело и с фабрикацией материалов. Каждый знает такие случаи, включая старых ученых. Существует непреодолимое давление такого поведения в некоторых частях мира, где все зависит от публикаций. Утверждалось, что фабрикации можно обнаружить в 10% - 20% рукописей в биологических науках.

Издание открытого доступа

Некоторые ученые признались в публикации в открытом доступе; имело место явное глубокое неодобрение открытого доступа многими, но не всеми учеными. Некоторые понимали, что он скорее навязан им по политическим причинам, а не потому, что они видят в нем то, что им действительно необходимо для того, чтобы помочь в их усилиях по распространению. Тем не менее, он открыт, сказал один ученый, подразумевая под этим, что почти все можно найти в сети. Удивительно то, что лишь немногие осознали, что традиционные издатели выпускают статьи и журналы открытого доступа.

Имеются некоторые беспокойства относительно открытого доступа:

1) Материал открытого доступа является низкого качества. Одним из используемых для его описания слов было слово «чепуха», другие определения – *тифосное издание и самообманывающийся автор.* Почему *где-то на начальном этапе карьеры вы хотите публиковаться и платите за это там, куда проще войти и не иметь ни репутации, ни известности.* Для меня журнал должен иметь бренд, чтобы рассматривать его.

2) Бизнес-модель основы открытого доступа (автор платит) должна была подорвать процесс строгого рецензирования; имеются проблемы относительно двух возможных прослеживаемых процессов рецензирования в отношении статей открытого доступа, с которыми обращались более снисходительно из-за привлекаемых ими денежных средств; имелась определенная сложность с моделью «автор платит», подводящей фундамент для издательства золотого открытого доступа. Перевод денежных средств, по-видимому, коррумпирует транзакцию. Некоторые ученые считают, что так можно купить свое издание.

3) Это явно плохой и непрофессиональный подход: *он не признает, что за издательством стоит много профессионализма и стандартов.*

4) Были беспокойства, что из-за давления со стороны учреждений ученые будут вынуждены публиковаться в журналах открытого доступа и что подписные журналы прекратят представлять различный контент и авторов.

Соцмедиа

Едва ли кто-то рассматривал соцмедиа даже в качестве плохой альтернативы изданию журнала или книги. Однако для начинающих карьеру ученых в социальных областях науки считалось, что именно социальные медиа должны обеспечить им способ передачи идей и информации, которые они не могут опубликовать в мире журналов с наличием импакт фактора, находящихся в руках кабальной политики старших ученых. Короче говоря, молодые, нацеленные на карьеру работники говорят о соцмедиа как о способе их общения.

ВЫВОДЫ

Главными выводами интересующих нас групп являются:

- Традиционные метрики качества (например, импакт фактор) все еще заслуживают доверия и, вероятно, используются больше, чем когда-либо, хотя связанные с ними проблемы широко признаются.
- Научные политики, касающиеся пребывания в должности, как университетские, так и правительственные (например, REF) сильно влияют на то, где публиковаться ученым и где цитировать.
- Импакт фактор журнала является более важным для публикации, чем для чтения или цитирования. Ученые, как правило, вольны читать то, что им нравится.
- Читатели обращают внимание на контент (реферат, методология и ссылки) и репутацию автора для определения достоверности. Их собственная персональная сеть значительно помогает им в указании на качественный контент.

• Некоторые аспекты меняются в мире научной коммуникации, но, возможно, не так быстро, как следовало бы ожидать, учитывая стремительные технологические изменения. Соцмедиа и издательство открытого доступа все еще борются за место во главе стола научного общества. Начинающие карьеру ученые просто могут помочь им туда добраться, но судей там нет.

В дальнейшем мы увидим, подтвердит ли изучение международного опроса эти результаты.

ЛИТЕРАТУРА

1. *Rieh S.Y., Danielson D.R.* Credibility: A multidisciplinary framework// Annual Review of Information Science and Technology. 2007. — Vol. 41.— P. 307-364.
2. *Gibbons M., Limoges C., Nowotny H., Schwartzman S., Scott P., Trow M.* The new production of knowledge: The dynamics of science and research in contemporary societies. — London: Sage, 1994.
3. *Delanty G.* The idea of the university in the global era: From knowledge as an end to the end of knowledge?// Social Epistemology. — 1998. — Vol.12, No.1. — P. 3-25.
4. *Nedeva M., Boden R., Nugrobo Y.* Rank and file: Managing individual performance in university research// Higher Education Policy. — 2012. — Vol. 25, No.3. — P. 335-360.
5. *Wilson L.* The academic man: A study in the sociology of a profession. — New York: Oxford University Press, 1940.
6. *Mebo L.I.* The rise and rise of citation analysis. — 2006. — arXiv preprint physics/0701012.
7. *Nicolini C., Nozza F.* Objective assessment of scientific performances world-wide// Scientometrics. — 2008. — Vol. 76, No.3. — P. 527-541.
8. *Zimmer C.* A sharp rise in retractions prompts calls for reform// The New York Times, 16 April, 2012. — 2012.

Сотрудничество, за которое приходится платить дорогой ценой: влияние научного обмана на карьеру соавторов*

Филипп МОНЖОН
(Philippe MONGEON)

Школа библиотекосведения
и информатики, Монреальский университет,
г. Монреаль, Канада

Винсент ЛАРИВЬЕР
(Vincent LARIVIÈRE)

Межуниверситетский центр исследований
по науке и технологии,
Университет Квебека в Монреале,
г. Монреаль, Канада

За последние несколько лет научное сообщество было потрясено рядом случаев крупного научного обмана. Число ежегодных отзывов (возвратов) в публикации работ также чрезвычайно растет, особенно в области биомедицины, и научная этика отвечает больше, чем за половину таких отзывов. Предполагается, что соавторы отозванных работ подвергаются влиянию некорректного поведения своих коллег; цель данного исследования — предоставить эмпирические свидетельства влияния отзывов в области биомедицинских исследований на научную карьеру соавторов. Используя данные из Web of Science (WOS), мы измерили продуктивность, влияние и сотрудничество 1123 соавторов 293 отозванных работ за пятилетний период до и после отзыва. Обнаружено явное свидетельство того, что сотрудничающие стороны страдают от последствий некорректного поведения коллег и что отзыв в случае обмана (fraud) имеет более значимые последствия, чем в случае ошибки (error). Наши результаты также предполагают, что степень распространения этих последствий тесно связана с ранжированием соавторов по отношению к отозванной статье, она ярко выражена для первых авторов, за которыми следуют последние авторы, тогда как для авторов, расположенных в середине, влияние менее ощутимо.

ВВЕДЕНИЕ

Отзывы публикации научных статей резко росли в последнее десятилетие, особенно в медицине [1]. Более 500 статей отозваны в PubMed в 2012 г. и 2013 г., что составляет более одной статьи ежедневно и 20-кратный рост по сравнению со средним числом 25 возвратов в год, начиная с 1990-х гг. Этот огромный рост отзывов, а также тот факт, что половина из них была сделана по причине научного обмана [2], вызвали множество дискуссий относительно некорректной научной этики как в научном сообществе, так и в обществе в целом.

Преыдущее исследование большей частью фокусировалось на росте отзывов статей [3,4], их причинах [2,5] и продолжающемся цитировании этих статей [6-9]. Другие авторы изучали в своих работах доминирование научного обмана [4, 10, 11], их возможные последствия для науки и для общества в целом [12], а также вероят-

ные способы предотвращения, обнаружения и воздействия на научный обман [13]. В то время как некоторые исследования рассматривали последствия обмана для отдельных областей [14] и научных групп [15], наше исследование впервые сравнивает авторскую продуктивность до и после отзыва, влияние и практики сотрудничества всех отдельных соавторов отозванных статей.

Многие исследователи видят завершение своей карьеры в позоре после обвинения в научном обмане, но уличенные в нем ученые редко работают поодиночке. Только один прецедент с голландским психологом Дидериком Стапелем бросил тень на работу более 30 его соавторов. Учитывая рост сотрудничества начиная с 1950-х гг. [16], ожидается, что все большее число ученых может в конечном счете подвергнуться влиянию некорректной этики со стороны своих соавторов. Сообщения об отзыве или исследования, как правило, выявляют отдельных авторов в качестве ответственных за обман, но предполагается, что несмотря на заявление об освобождении от ответственности за неэтичное поведение, невинные соавторы все еще испытывают на себе некоторые последствия обмана [17]. Однако пока нет эмпирического свидетельства, которое обеспечивало бы поддержку такого предположения. Данная статья восполняет этот пробел с помощью исследования продук-

* Перевод Mongeon P., Larivière V. Costly collaborations: The impact of scientific fraud on co-authors' careers.— 2014.— <http://arxiv.org/ftp/arxiv/papers/1409/1409.0129.pdf>. Данная работа выполнена при поддержке Межуниверситетского центра исследований по науке и технологии (CIRST) и канадской программы научных кафедр.

тивности после отзыва (число статей, опубликованных за год), влияния (число ссылок на статью) и практик сотрудничества соавторов (число авторов, учреждений и стран на статью), вовлеченных в случаи нарушения научной этики в области биомедицины.

Роберт Мертон [18] описал научную систему как расслоенное и иерархическое пространство, в котором ученые вносят вклад в общую сокровищницу знаний в обмен на различные формы признания или на то, что Бурдьё [19] называет символическим капиталом. В свою очередь ученые используют символический капитал для получения большего доступа к ресурсам и увеличения объема или важности своего вклада, а значит получения большего символического капитала и улучшения своей позиции в структуре науки. Этот механизм аккумуляции символического капитала усиливается тем, что Мертон [18] назвал эффектом Матфея, который можно определить как награждение большим признанием тех, кто уже имеет, и отказ в признании тем, у кого его нет. В контексте совместного исследования символический капитал, полученный с помощью определенного вклада, делится между его соавторами, и можно ожидать, что отрицательный капитал, приобретенный в случае обнаружения обмана, также в некоторой степени делится. Если это так, то такой негативный эффект можно наблюдать, по крайней мере, в одном из трех моментов: 1) ученые могут публиковать меньше статей, 2) их публикации могут меньше цитироваться коллегами и 3) они (ученые) могут быть менее склонны к сотрудничеству либо потому, что стали больше сомневаться в отношении сотрудничества, стали более избирательными в плане людей, с которыми хотят работать, либо потому, что другие ученые менее склонны с ними сотрудничать из-за их связи с делом о нарушении этики. Таким образом, эта статья анализирует продуктивность, научное влияние и практики сотрудничества соавторов; значительное снижение этих показателей, следующее за отзывом статьи, покажет цену, которую надо заплатить за связь с человеком, замешанным в публично известном деле о нарушении этики.

Чтобы оценить степень этих последствий, мы изучили карьерный рост 1038 ученых в области биомедицины, которые принимали участие в авторстве отозванных статей в 1996-2006 гг. Для пятилетнего периода до и после отзыва мы подсчитали число публикаций авторов, а также характеристики их статей, такие как средние относительные ссылки (Average Relative Citations – ARC) и число соавторов, учреждений и стран. Для сравнения использовали контрольную группу (control) из 1862 соавторов, которые не были причастны ни к одному известному делу об обмане. Отдельно нами были изучены первые, средние и последние авторы отозванных статей, так как в биомедицинской области распределение доверия (и ответственности) между соавторами, как правило, имеет U-образную форму с первым автором (first author), обычно, проводившим работу, последним автором (last author), руководившим этой работой, и менее вовлеченным в нее средним автором (middle author) [20]. Таким образом, можно утверждать, что последствия научного обмана для авторов должны быть пропорциональны их индивидуальному уровню ответственности. Мы также отдельно изучили авторов статей, отозванных в случае обмана и в случае ошибки. Хотя существует общее мнение, что честно признанные ошибки являются нормой в научной сфере и что они «должны рассматриваться не как источники смущения или неудачи, а скорее как возможности для обучения и совершенствования» [21], обман является серьезным ук-

лонением от основных принципов и цели научного исследования. Поэтому мы полагаем, что отзыв в случае обмана будет иметь больше влияния на карьеру ученого, чем в случае ошибки.

Это исследование является по нашим сведениям первым, обеспечивающим эмпирическое свидетельство последствий, когда отзывы статей в биомедицинской области (наиболее важны те, которые случаются в делах, касающихся научного обмана) влияют на карьеру соавторов, формально не несущих ответственности за обман или ошибку. Это также первое исследование, которое предоставляет данные о последующем научном выходе авторов отозванных статей. Его концентрация на биомедицинской области является необходимостью и ограничением, первое — потому что большинство отозванных статей принадлежит этой области, включая лишь некоторые отзывы, встречающиеся в других областях, а второе — потому что оно ограничивает обобщение наших наблюдений рамками биомедицинской области.

ДААННЫЕ И МЕТОДЫ

Отозванные статьи

Чтобы получить выборку отозванных статей, мы проводили поиски в PubMed всех сообщений об отзыве (тип сообщения – «Отзыв публикации») и всех отозванных статей (тип сообщения – «Отозванная статья») и нашли всего 2451 отзыв и 2299 отозванных статей. Сгруппировали в пары отозванные статьи с соответствующим сообщением об их отзыве в целях нахождения года отзыва и задержки отзыва (т.е. количества лет, прошедших между публикацией и отзывом) всех отозванных статей. Затем мы нашли эти статьи в библиометрической версии Web of Science (WoS) издательства Thomson Reuters, предоставленной Обсерваторией наук и технологий. Она является реляционной базой данных, позволяющей связать любую переменную с другой, полученной из источника данных XML, предоставленного издательством Thomson Reuters и хранящегося на сервере Microsoft SQL. Мы ограничили поиск статьями, написанными на английском языке и опубликованными в биомедицинских журналах и журналах по клинической медицине. Оставили только те статьи, которые были отозваны в 1996-2006 гг., чтобы получить достаточный временной интервал для оценки продуктивности соавторов, влияния и сотрудничества до и после отзыва. Это обеспечило нам выборку, включающую 443 отозванные статьи.

В анализе отозванных статей, найденных в PubMed, авторы работы [14] классифицировали отозванные статьи по причине отзыва. Они сделали это, используя информацию, найденную по сообщению об отзыве, а также любую другую информацию, найденную в Интернете. Мы использовали их данные для категоризации статей в нашей выборке в соответствии с причиной отзыва: обман, включающий фабрикации данных, фальсификацию или плагиат (N=179) и ошибка (N=114). Статьи, отозванные по другим причинам (N=150), не вошли в наш анализ.

Авторы отозванных статей

Мы составили список всех авторов статей, отозванных из-за обмана или ошибки. После устранения неоднозначности имен авторов, что было сделано путем просмотра каждой отозванной статьи авторов, имеющих два или более отзыва публикаций, список насчитывал 1098 отдельных авторов. Затем для каждого из этих авторов мы провели поиск в Web of Science всех статей,

опубликованных в течение пяти лет до и после отзыва. Для авторов множества отозванных в разные годы статей мы искали статьи, опубликованные в течение пяти лет до первого отзыва и пяти лет после последнего отзыва. Пятилетний интервал использовался для того, чтобы собрать достаточное количество публикаций для каждого ученого и гарантировать, что наблюдаемые изменения происходили не только из-за долгосрочных тенденций или обычных краткосрочных колебаний.

После устранения неоднозначности имен авторов окончательная выборка содержала всего 15333 отдельные статьи, опубликованные в 1991-2011 гг. Устранение неоднозначности имен авторов производилось вручную с использованием фамилий и инициалов авторов, их мест работы, дисциплины журнала и ключевых слов к названиям статей. В случае, когда эта информация не позволяла нам различить омонимы, мы смотрели саму статью или искали в сети резюме ученого (curriculum vitae).

Время для отзыва

Чтобы перегруппировать статьи, опубликованные в определенный период, связанный с годом отзыва (например, пять лет, предшествующие отзыву), было необходимо перевести год публикации в другую переменную, которую мы называем «время для отзыва» (time to retraction, T). Значение T колеблется от -5 до 5, при этом 0 считается годом отзыва. В случаях множественных отзывов в течение многих лет, переменная T=0 для всех статей, опубликованных между годами первого и последнего отзыва включительно. В большинстве частей нашего анализа мы делили публикации на две группы: до отзыва и после. Первая группа включает статьи, опубликованные между T= -5 и T= -1, а вторая – статьи, опубликованные между T= 1 и T= 5.

Нечестные и честные ученые

Мы использовали данные из работы авторов [14], а также сообщений об отзыве и поиска в Интернете для установления авторов, ответственных за обман или другие причины отзыва. Нам удалось определить авторов, ответственных за большинство случаев обмана (82 автора, ответственные за 159 из 179 случаев обмана), в то время как ответственный автор в случаях ошибки определялся редко (3 автора, ответственные за 5 из 114 случаев ошибки). Не было найдено ни одного ответствен-

ного автора за 20 случаев обмана (4 случая фабрикация данных или фальсификации и 16 случаев плагиата). Так как в таких случаях невозможно отличить нечестного автора от честного, все авторы этих 20 статей (N=66) были удалены из нашей выборки в целях гарантии того, что она содержит только честных ученых.

Также мы разделили соавторов на три эксклюзивные группы в соответствии с их рангом в списке авторов отозванных статей. В случаях авторов с многочисленными отзывами статей, они приписывались к группе «первых авторов», если они были первыми авторами, по крайней мере, одной отозванной статьи. Группа «последние авторы» содержит авторов, которые были приведены в списке как последний автор, по крайней мере, одной отозванной статьи и кто не значился в списке как первый автор любой отозванной статьи. Группа «средние авторы» содержит авторов, не приписанных ни к группе «первых авторов», ни к группе «последних авторов». Что касается статей с одним автором, этот автор считался первым, а в случаях статей с двумя авторами второй автор приписывался к группе «последний автор».

Контрольная группа

Для каждой статьи, отозванной в 1996-2006 гг., в нашем первоначальном списке (т.е. включающем статьи, отозванные по другим причинам, отличным от обмана или ошибки), мы произвольно выбирали не отозванную статью с тем же самым числом авторов, опубликованную в том же номере того же журнала. Это обеспечило нам список в 1862 автора, для которых мы искали все публикации за пять лет до и после отзыва их соответствующей отклоненной статьи, всего 55036 статей. Авторы контрольной группы были также категоризованы в соответствии с их рангом в статье, опубликованной в том же журнале, что и отозванная статья.

Окончательная выборка

Окончательная выборка авторов, включая контрольную группу, представлена в табл. 1.

Авторы, не имевшие публикаций в период ни до, ни после отзыва, были удалены из выборки, используемой для оценки научного влияния соавторов (см. рис. 2 в разделе «Результаты») и практик сотрудничества (см. рис. 3 там же). Эта подвыборка, используемая для таких анализов, представлена в табл. 2.

Таблица 1

Выборка авторов

Ранг	Обман	Ошибка	Контрольная группа	Всего
Первые авторы	45	108	411	564
Средние авторы	346	366	1046	1758
Последние авторы	77	102	405	584
Всего	468	576	1862	2906

Таблица 2

Подвыборка авторов с, по крайней мере, одной публикацией в период до и после отзыва

Ранг	Обман	Ошибка	Контрольная группа	Всего
Первые авторы	28	83	354	465
Средние авторы	253	276	860	1389
Последние авторы	64	89	382	535
Всего	345	448	1596	2389

Библиометрические показатели

Чтобы измерить продуктивность соавторов, мы использовали число статей, публикуемых за год. Число статей было нормализовано на индивидуальном уровне с помощью деления значения за данный год на его среднее значение за период пяти лет, предшествующих пяти годам после отзыва. Мы называем это конечное число индивидуальной относительной продуктивностью (Individual Relative Productivity - IRP) ученого. Эта нормализация на индивидуальном уровне позволяет сравнение данного показателя между учеными. Чтобы измерить научное влияние соавторов, нами использовалось среднее число ссылок, полученных их статьями. Число ссылок для каждой статьи было нормализовано на уровне дисциплины путем деления числа ссылок статьи на среднее число ссылок, полученных всеми статьями, опубликованными в одной и той же области и в одном и том же году, область определялась по журналу, в котором данная статья была опубликована, а дисциплина журнала – журнальной классификацией Национальной организации содействия развитию науки (National Science Foundation – NSF). Итоговый показатель назван средними относительными ссылками (ACR). Наконец, мы использовали число авторов, учреждений и стран, приведенных в списке публикаций соавторов, также нормализованное на уровне дисциплины, как показатель их практик сотрудничества.

Статистические тесты

Чтобы оценить значимость отличий, наблюдаемых между группами, мы использовали тест Mann-Whitney U-test. Этот тест был предпочтен тесту t-test, так как он более надежен и из-за непараметрического характера сравниваемых распределений. Нулевая гипотеза (H_0) теста Mann-Whitney U-test состоит в том, что сравниваемые группы имеют одинаковую медиану, а альтернативная гипотеза (H_1) заключается в том, что медианы неравны. Статистически значимое отличие ($P < 0,1$) означает, что существует вероятность менее 10% в отношении того, что наблюдаемые отличия случайны и что медианы фактически одинаковы. Поэтому, когда мы получили значение P меньше, чем 0,1, мы отклонили нулевую гипотезу и приняли альтернативную.

РЕЗУЛЬТАТЫ

Нечестные авторы

Прежде чем посмотреть на честных соавторов, мы посмотрели на 79 соавторов, официально установленных в качестве ответственных за обман ($N=79$). Из этих соавторов 45 не имели публикаций в течение пяти лет, следующих за отзывом статьи, предположительно из-за того, что они ушли из научной области, хотя медианное IRP 34 оставшихся авторов снизилось на 64,6%. Эти результаты подтверждают, что обнаружение обмана будет сильно влиять на карьеру нечестных ученых, приводя ее к концу в большинстве случаев. Эти 79 соавторов (а также 3 автора, установленные в качестве ответственных за ошибки) были исключены из выборки для другого анализа, который фокусируется на честных участниках сотрудничества.

Честные участники сотрудничества

Наиболее поразительным является число соавторов, для которых мы не нашли публикаций в течение пяти лет, следующих за отзывом. Как показано в табл. 3, это

представляет случай для 27,6% и 20,3% соавторов статей, отозванных из-за обмана или ошибок, соответственно. Было гораздо меньше соавторов (12,1%) в контрольной группе без публикаций в течение пятилетнего периода после отзыва.

Таким образом, неудивительно, что рис. 1 показывает – медианная индивидуальная относительная продуктивность (IRP) соавторов отозванных статей быстро снижается в период, следующий за отзывом. Рис. 1 также демонстрирует, что степень этого снижения продуктивности зависит от причины отзыва (потери продуктивности более значительны после отзыва из-за обмана, чем после отзыва из-за ошибки) и позиции соавтора в авторском коллективе отозванной статьи (IRP первых авторов снижается более резко, чем средних и последних авторов). Рис. 1 также отображает то, что IRP авторов в контрольной группе эволюционирует с различием для первых, средних и последних авторов. Хотя медианное IRP вполне стабильно в течение всего 11-летнего периода для последних авторов, мы наблюдаем более высокие изменения для первых и средних авторов. Что касается первых авторов, то имеется отчетливый рост в период до отзыва, за которым следует небольшое снижение. Относительно средних авторов медианное IRP также растет в период до отзыва, а схожее снижение наблюдается в период после отзыва. Такая вариация наиболее вероятна из-за того, что последние авторы, как правило, представляют ведущих ученых со стабильной карьерой, тогда как первые и средние авторы могут быть преходящими авторами, которые, возможно, не ставят своей целью научную карьеру.

Медианная доля статей, опубликованных до и после отзыва (табл.4), показывает, что соавторы испытывают значительное снижение в IRP независимо от причины отзыва, исключение составляют последние авторы статей, отозванных из-за ошибки, для которых спад в продуктивности не является статистически значимым. Случай обмана, а не ошибки, имеют значительное влияние на IRP средних авторов. Тем не менее, средние авторы контрольной группы также демонстрируют важное снижение в IRP. Поэтому наши результаты предполагают, что отзыв статьи имеет меньшее влияние на последующее число публикаций средних авторов, по сравнению с первыми и последними. Таким образом, степень последствий, ощущаемых соавторами, кажется, распределяется так, что она подобна распределению полученного авторами доверия, первые и последние авторы больше подвержены влиянию научного обмана со стороны соавторов, чем средние авторы.

В оценке эффекта отзывов на научное влияние соавторов мы должны исключить из нашей выборки ученых, для которых мы не нашли ни одной публикации в период либо до, либо после отзыва. Это значительно сократило нашу выборку (см. табл. 2 в разделе «Методы»), поскольку, как мы видели ранее, многие ученые не имели публикаций после отзыва. Рис. 2 показывает для каждой группы соавторов медианный спад в ARC (ось x) и индивидуальные относительные публикации (ось y) авторов из этой подвыборки. Группы с ростом медианной IRP и медианных ARC должны располагаться в первом квадранте графа (верхний правый), тогда как группы со спадом обоих показателей – в третьем квадранте (нижний левый).

Таблица 3

Доля (%) авторов без публикаций в течение пяти лет, следующих за отзывом

Ранг	Обман		Ошибка		Контрольная группа	
	N	%	N	%	N	%
Первые авторы	17	39,1	24	22,2	63	11,3
Средние авторы	96	27,7	81	22,1	266	15,2
Последние авторы	16	20,8	12	11,8	29	5,0
Всего	129	27,6	117	20,3	358	12,1

Таблица 4

Отличие между медианной индивидуальной относительной продуктивностью (IRP) до и после отзыва

Ранг	Группа	Вариация IRP (%)	P-значение*
Первые авторы	Обман	-65,2	0,000**
	Ошибка	-35,5	0,006**
	Контрольная	0	-
Средние авторы	Обман	-50,0	0,000**
	Ошибка	-44,2	0,000**
	Контрольная	-25,0	-
Последние авторы	Обман	-46,7	0,000**
	Ошибка	-23,5	0,102
	Контрольная	-17,4	-
Все авторы	Обман	-50,0	0,000**
	Ошибка	-39,2	0,000**
	Контрольная	-18,3	-

Примечание: соавторы, опубликовавшие статьи только при T=0, сюда не включены. Это случай 8, 10 и 14 соавторов, замеченных в обмане, ошибке и тех, кто вошел в контрольные группы, соответственно.

*P-значение — результат сравнения группы авторов с контрольной группой с использованием теста Mann-Whitney U-test.

** Спад в публикациях для этих групп соавторов статистически важен (P<0,01).

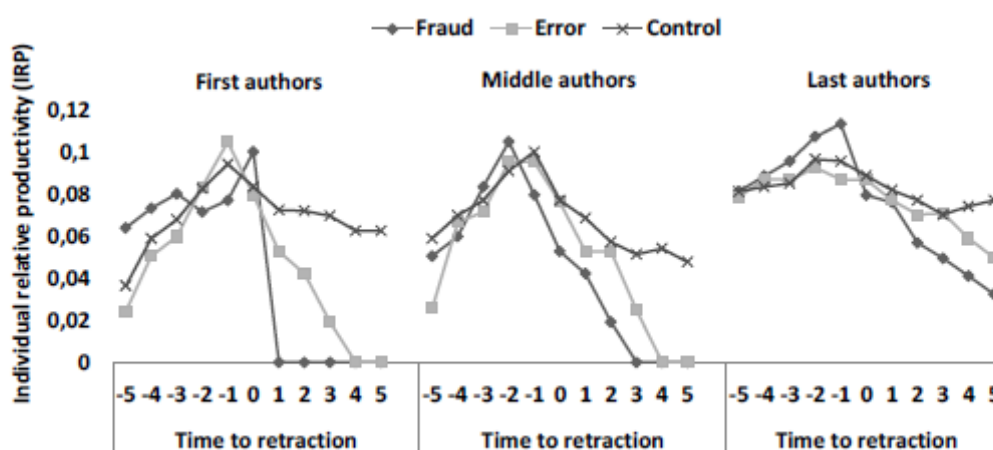


Рис. 1. Медианная индивидуальная относительная продуктивность (IRP) — от пятилетнего периода, предшествующего отзыву, до пятилетнего периода после отзыва.

Рис. 2 четко иллюстрирует различия между соавторами, которые отзывали статью из-за обмана, теми, кто отзывал статью из-за ошибки, и теми, кто принадлежит к контрольной группе. Как можно ожидать, медианная продуктивность и влияние авторов, принадлежащих к группе обмана, располагаются в нижнем левом квадранте графа, показывая, что оба эти показателя подвержены негативному влиянию обмана. Также, как ожидается, медианы для контрольной группы расположены ближе к середине графа, что означает относительную стабильность в периоды до и после отзыва. Интересно, что авторы, отзывавшие статью из-за ошибки, по-видимому, должны чувствовать снижение их научного влияния чуть меньше, чем авторы контрольной группы, и даже рост в случае средних и

последних авторов. В условиях самого научного влияния большинство наблюдаемых различий не является статистически важным ($P > 0,05$), за исключением авторов статей, отзыванных из-за ошибки, для которых медианное влияние значительно растет ($P < 0,05$) (см. табл. 5). Это ограничение в нашем исследовании может быть отчасти объяснено тем фактом, что многие ученые, которые не публиковались после отзыва, исключены из этой части анализа. Однако показанные в табл.5 результаты подтверждают, что при делении авторов только по причине отзыва (а не рангу статьи) обман имеет значительное негативное влияние на ссылки в интервале доверия в 90% ($P < 0,1$), хотя с другой стороны, ошибки имеют положительное влияние на ссылки в интервале доверия в 99% ($P < 0,01$).

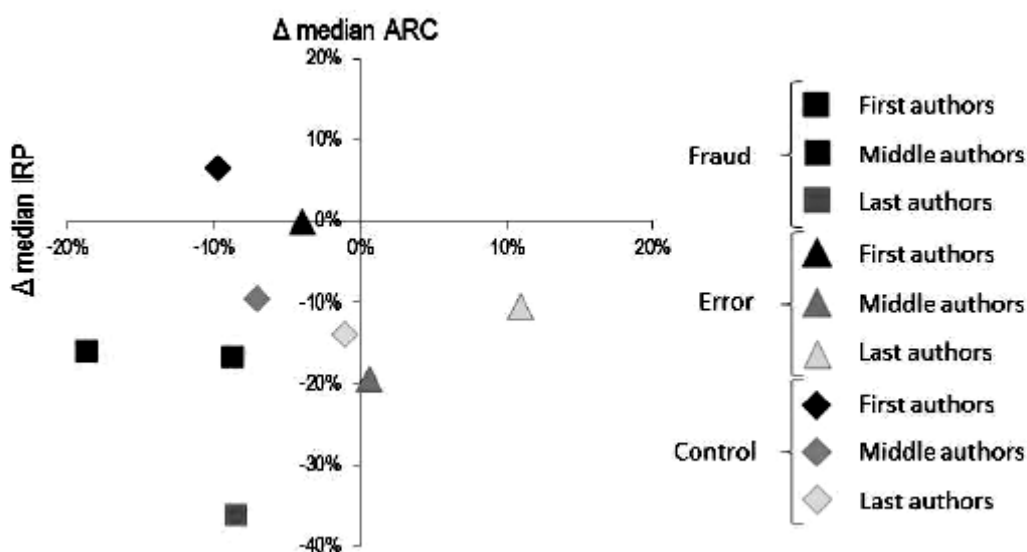


Рис. 2. Отличие между медианной индивидуальной относительной продуктивностью (IRP) и медианными средними относительными ссылками (ARC) до и после отзыва.

Таблица 5

Отличие в медианах средних относительных ссылок (ARC) до и после отзыва

Ранг	Группа	Варнация ARC (%)	P-значение*
Первые авторы	Обман	-8,7	0,986
	Ошибка	-4,0	0,274
	Контрольная	-9,7	-
Средние авторы	Обман	-18,7	0,092
	Ошибка	0,6	0,013***
	Контрольная	-7,1	-
Последние авторы	Обман	-8,5	0,524
	Ошибка	11,0	0,108
	Контрольная	-1,0	-
Все авторы	Обман	-17,6	0,056****
	Ошибка	2,0	0,003**
	Контрольная	-7,5	-

* P-значение — результат сравнения группы авторов с контрольной группой на основе использования теста Mann-Whitney U-test.

** Отличие в средних относительных ссылках для этих групп авторов является статистически значимым ($P < 0,01$).

*** Отличие в средних относительных ссылках для этих групп авторов является статистически значимым ($P < 0,05$).

**** Отличие в средних относительных ссылках для этих групп авторов является статистически значимым ($P < 0,1$).

Это может быть связано с предыдущим наблюдением авторов работы [22], которые показали, что отзывы, о которых сообщается самостоятельно (большой частью, ошибки), приводят к росту в ссылках для авторов более ранних работ. Наши результаты должны в таком случае предполагать, что это может происходить и с авторами последующих работ. Поскольку отзывы имеют относительно небольшое влияние на нормализованное по областям среднее число ссылок, полученных будущими статьями, то, тем не менее, имеется более значимый спад в общем влиянии, измеренном как общее число полученных ссылок, в силу наблюдаемого в публикациях снижения (рис. 1).

Сотрудничество

Наконец, мы оценили влияние отзыва на уровень сотрудничества соавторов. Мы не нашли никакого видимого отличия в числе авторов, учреждений или стран на статью, нормализованном на уровне дисциплины в периоды до и после отзыва. Например, рис. 3 показывает уровень межучрежденческого сотрудничества за целый период, который, по-видимому, должен быть одинаковым для всех групп. Таким образом, отзывы статей, кажется, не должны иметь влияния на практики сотрудничества соавторов.

ОБСУЖДЕНИЕ

Мы обнаружили явное свидетельство, что сотрудничающие люди испытывают последствия незтичного поведения коллег, и что обман имеет больше последствий, чем ошибки. Наши результаты также предполагают, что степень этих последствий тесно связана с ранжированием соавторов по отношению к отозванным статьям, эти последствия острее ощущаются первыми авторами, за которыми следуют последние авторы, тогда как влияние отзыва менее важно для средних авторов. Примечательно, что важное различие во влиянии отзыва на первого и последнего авторов наиболее вероятно, благодаря тому факту, что хотя пер-

вый (автор) является часто молодым ученым с еще не устоявшимся профессиональным статусом, последний, как правило, является авторитетным ученым, чья позиция в социальной структуре науки может помочь в преодолении шока, вызванного отзывом. Тем не менее, наши наблюдения согласуются с идеей, что большее вовлечение участников в научный процесс ведет не только к большему доверию, но и к большей ответственности [23], а следовательно — к большим последствиям в случаях обмана. В контексте совместной работы символический капитал, полученный через публикацию открытия, делится, хотя и неодинаково, между соавторами. Таким образом, можно сказать, что отзыв статьи, считающейся обманной, несет ее соавторам отрицательную форму символического капитала (например, позор вместо признания), которая также делится неравнозначно (автор, подозреваемый в обмане, вероятно всего, получит самую большую долю). Следовательно, связь наших результатов с теорией кумулятивного преимущества и недостатка [18] предполагает, что отзыв обманной статьи может инициировать негативную направленность, которая будет дорого стоить некоторым ученым, приведя многих из них (большой частью, менее опытных) к полному уходу из данной области науки.

Как упоминалось в «Введении», мы ожидали, что последствия отзыва ошибочной статьи будут минимальными. Тем не менее, наши результаты показывают, что ошибки тоже имеют последствия (хотя и не такие важные, как в случаях обмана) для сотрудничающих лиц с точки зрения продуктивности. Эти результаты можно отчасти объяснить тем фактом, что отзывы происходят, как правило, в случаях основных ошибок, делающих результаты ненадежными в целом, тогда как минимальные ошибки приводят, вероятно всего, к исправлениям. Однако, ошибки, кажется, должны иметь положительное воздействие с точки зрения ссылок. Необходимо больше исследований, чтобы получить лучшее понимание этого феномена.

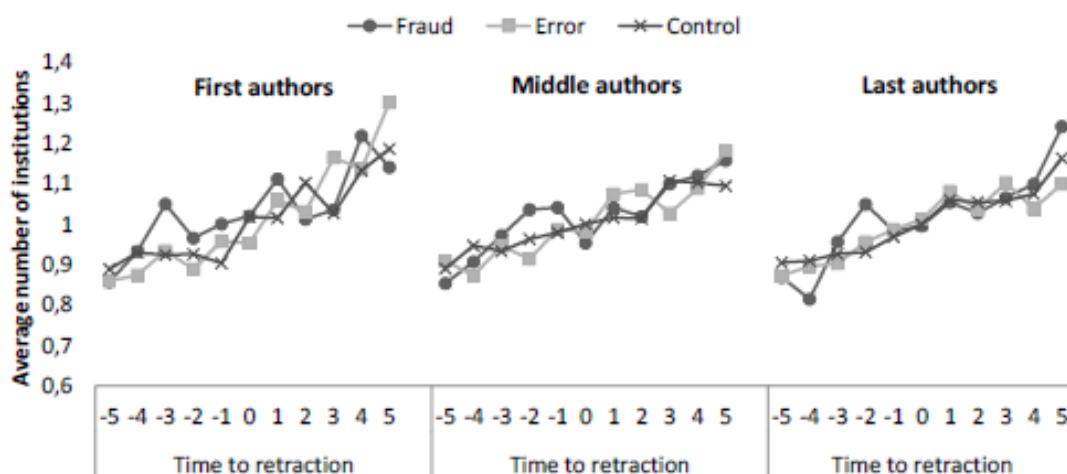


Рис. 3. Среднее число учреждений на статью, нормализованное на уровне дисциплины, от пяти лет до отзыва и пяти лет после отзыва.

Одно из ограничений нашего исследования состоит в том, что из-за большого числа ученых, переставших публиковаться после отзыва (в основном из-за случаев обмана), мы остались с весьма маленькой выборкой, необходимой для нашего сравнения влияния и сотрудничества до и после отзыва. В связи с этим дальнейшее исследование, повторяющее нашу методологию, может оказаться выгодным из-за гораздо большего числа отзывов, встречающихся в биомедицинской области в последние несколько лет. Другим ограничением является то, что используемая здесь методология не позволяет нам контролировать другие, нежели отзывы, факторы, что может влиять на продуктивность ученых, научное влияние и сотрудничество. Тем не менее, использование нами большой контрольной группы и критериев значимости, полученных с помощью статистических тестов, усиливают доверие к надежности наших результатов.

Будучи ограниченными биомедициной, это исследование не позволяет нам распространить наши результаты на другие области. Таким образом, дальнейшее исследование сможет оценить последствия отзывов для соавторов в других областях. Однако отзывы в других областях являются не такими доминирующими, как в биомедицине, что может усложнить сбор достаточного числа отзывов для проведения такого анализа, как представлен здесь. Дальнейшая работа может также проанализировать, как варьируется воздействие отзывов в зависимости от социокультурного или учрежденческого факторов. Это становится возможным благодаря большому числу отзывов в последние годы.

ЗАКЛЮЧЕНИЕ

Наши результаты показывают, что научный обман в биомедицинской области не только вреден для науки в целом, но также и на индивидуальном уровне уличенного в обмане ученого, также он вреден и для честных ученых, чья ошибка заключается только в неверном выборе коллег для работы. Действительно, многие из соавторов обманных статей, включенных в наше исследование, опубликовали меньше статей в следующие после отзыва обманной статьи годы. В меньшей степени, но среднее число ссылок, полученных их статьями, также снизилось после отзыва. Однако отзыв, кажется, не влияет на практики сотрудничества. Эти изменения в продуктивности и влиянии могут быть связаны с ассоцированием соавторов со случаем обмана, так как ни соавторы нашей контрольной группы, ни те, кто отзывал статьи из-за ошибки, не показывают подобных тенденций. Последние даже отмечают рост среднего числа ссылок на их статьи после отзыва.

Имеющие место в случае участия в научном обмане последствия выходят за рамки снижения роста статей или потерь научного влияния для уличенных в обмане авторов и их коллег. Некоторые последствия могут быть психологическими (например, ученые теряют доверие к науке, коллегам и учреждениям) или пустой тратой научных усилий и финансов. Судебное дело в отношении Хендрика Шёна в области физики служит хорошим примером такой затраты усилий: он сфальсифицировал «основополагающие» результаты, которые мировые ученые пожелали воспроизвести и основываться на них. Было потрачено много времени и денежных средств на эти безуспешные попытки, а обнаружение обмана вынудило некоторых обескураженных ученых (многие из которых доктора наук и аспиранты), работающих над этими проектами, оставить идею сделать карьеру в науке

[24]. Более того, случаи обмана, обнаруживаемые почти ежедневно, представляют, вероятнее всего, только вершину айсберга: метаанализ исследований показал, что 2% ученых признались в фальсификации или фабрикации данных, а 14% сказали, что знают тех, кто так делает [10]. В США обвинения в обмане, полученные Управлением по проблемам ответственности в научных исследованиях (U.S. Office of Research Integrity – ORI), достигли такого размаха, что только небольшая их часть может действительно быть расследована [25]. Поэтому вероятно, что число судебных дел будет продолжать расти и что все больше участвующих в сотрудничестве лиц поймет, что их карьере не суждено реализоваться. Наконец, имея возможность получить материал из первых рук, посмотрите на работу своих коллег перед представлением для публикации, соавторы являются первым звеном цепочки рецензирования. Учитывая возможные последствия соавторства в обманной статье, подумайте, как это может отразиться на карьере, в ваших интересах отнестись к этому серьезно.

ЛИТЕРАТУРА

1. Zhang M., Grieneisen M. The impact of misconduct on the published medical and non-medical literature, and the news media // *Scientometrics*.—2012.—Vol. 96, No.2. — P.1–15.
2. Steen R. G., Casadevall A., Fang F. C. Why has the number of scientific retractions increased? // *PLoS ONE*.— 2013.— Vol. 8, No.7, e68397.
3. Cokol M., Ozbay F., Rodriguez-Esteban R. Retraction rates are on the rise // *EMBO Reports*.— 2008.— Vol. 9, No.1.— P. 2.
4. Steen R. G. Retractions in the scientific literature: Is the incidence of research fraud increasing? // *Journal of Medical Ethics*.— 2011.— Vol. 37, No.4.— P. 249–253.
5. Fang F. C., Steen R. G., Casadevall A. Misconduct accounts for the majority of retracted scientific publications // *Proceedings of the National Academy of Sciences*.— 2012.— Vol.109, No.42.— P.17028–17033.
6. Furman J. L., Jensen K., Murray F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine // *Research Policy*.— 2012.— Vol. 41, No.2.— P. 276–290.
7. Neale A.V., Northrup J., Dailey R. K., Marks E., Abrams J. Correction and use of biomedical literature affected by scientific misconduct // *Science and Engineering Ethics*.— 2007.— Vol. 13, No.1.— P. 5–24.
8. Neale A. V., Dailey R. K., Abrams J. Analysis of citations to biomedical articles affected by scientific misconduct // *Science and Engineering Ethics*. —2010.— Vol. 16, No.2.— P. 251–261.
9. Pfeifer M. P., Snodgrass G. The continued use of retracted, invalid scientific literature // *JAMA*. — 1990.— Vol. 263, No.10.— P. 1420.
10. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data // *PLoS ONE*. — 2009. — Vol. 4, No.5, e5738.
11. Sovacool B. K. Exploring scientific misconduct: Isolated individuals, impure institutions, or an inevitable idiom of modern science? // *Journal of Bioethical Inquiry*. — 2008.— Vol. 5, No.4.—P. 271–282.
12. Steen R. G. Retractions in the medical literature: How can patients be protected from risk? // *Journal of Medical Ethics*. — 2012. — Vol. 38, No.4. — P. 228–232.

13. *Steneck N.* Fostering integrity in research: Definitions, current knowledge, and future directions// *Science and Engineering Ethics*. — 2006. — Vol. 12, No.1. — P. 53–74.
14. *Azoulay P., Furman J. L., Krieger J. L., Murray F. E.* Retractions// *NBER Working Paper Series*. — 2012. — P. 18449.
15. *Jin G. Z., Jones B., Lu S. F., Uzzi B.* The reverse Matthew effect: Catastrophe and consequence in scientific teams// *NBER Working Paper Series*. — 2013. — P.19489.
16. *Wuchty S., Jones B., Uzzi B.* The increasing dominance of teams in production of knowledge// *Science*. — 2007. — Vol. 316, No. 5827. — P.1036–1039.
17. *Bonetta L.* The aftermath of scientific fraud// *Cell*. — 2006. — Vol. 124, No.5. — P. 873–875.
18. *Merton R. K.* The Matthew effect in science// *Science*.— 1968.— Vol. 159, No.3810.— P. 56–63.
19. *Bourdieu P.* The forms of capital// *J. G. Richardson* (Ed.), *Handbook of theory and research for the sociology of education*.— New York: Greenwood, 1986.
20. *Pontille D.* La signature scientifique: Une sociologie pragmatique de l'attribution. — Paris: CNRS, 2004.
21. *Nath S. B., Marcus S. C., Druss B. G.* Retractions in the research literature: Misconduct or mistakes? //The *Medical Journal of Australia*. — 2006. — Vol. 185, No.3. — P. 152–154.
22. *Lu S. F., Jin G. Z., Uzzi B., Jones B.* The retraction penalty: Evidence from the Web of Science// *Scientific Reports*. — 2013. — Vol. 3.
23. *Birnholz J. P.* What does it mean to be an author? The intersection of credit, contribution, and collaboration in science// *Journal of the American Society for Information Science and Technology*.— 2006.— Vol. 57, No.13.— P. 1758–1770.
24. *Reich E. S.* *Plastic fantastic: How the biggest fraud in physics shook the scientific world*. —New York: Palgrave Macmillan, 2009.
25. Seven days: 26 April–2 May 2013// *Nature News*. — 2013.—<http://www.nature.com/news/seven-days-26-april-2-may-2013-1.12899#/trend>

Фасеты и типизированные отношения как инструменты для процессов рассуждения в информационном поиске*

Винфрид ГЁДЕРТ
(Winfried GÖDERT)

Кельнский университет прикладных наук,
Институт информатики,
г. Кельн, Германия

Фасетное расположение объектов и типизированные отношения для отображения различных ассоциаций между объектами служат утвердившимися средствами в представлении знания. В статье обсуждается предложение, сочетающее оба средства, чтобы прийти к выводам относительно цепочек отношений. Такой подход может принести новую выгоду для процессов информационного поиска, особенно при моделировании неоднородных сред в семантической сети. Фасетное расположение можно использовать как инструмент отбора для семантического знания, моделируемого в рамках представления знания. Типизированные отношения между объектами различных фасетов могут применяться в качестве ограничений для их отбора внутри фасетов.

ВВЕДЕНИЕ

Представление знания, рассматриваемое как инструмент информационного поиска, должно находить решения для изображения элементов знания, выгодного для использования людьми. Традиционные языки индексирования сильно ориентированы на когнитивную интерпретацию представлений их содержания, объектов. Такая интерпретация изначально основана на знании пользователя и формально поддерживается только обеспечением некоторых различных типов отношений между объектами. Как правило, не дается никакого формального определения объекта, например с помощью указания множества атрибутов, характеризующих содержание. Приписанные отношения между терминами должны проверяться главным образом правдоподобием. Также нет никаких критериев для атрибутов, которые выбраны с целью решения установить только эти и никакие другие отношения.

В силу этих обстоятельств представленное знание обычно может быть найдено только на основе объектов. Стоящие за выводами относительно цепочек отношений процессы не могут быть выполнены. Только формы подробно изученных поисков, поддержанных иерархической структурой представления знания, иногда могут обрабатываться. Качество окончательных массивов сильно зависит от транзитивности установленных иерархий, формального свойства, которое не всегда может восприниматься на веру [3,5,18]. В этом отношении языки индексирования отличаются от структур знания, по-

строенных на принципах формальных представлений знания и считающихся основой семантических утверждений в средах семантической сети.

Мы выдвигаем предложение по сочетанию двух подходов, чтобы построить поисковые системы большей мощности. Цель – распознать *близость* (aboutness)* определенных документов, поскольку рассматривается их *апостериорное* содержание. Часть этого предложения состоит в сравнении между пониманием людьми представленного знания и должна гарантировать формальную правильность объяснения цепочек отношений путем машинных выводов. Основными методологическими средствами являются фасетные представления объектов, типизированные отношения между объектами и выводы относительно цепочек отношений.

Начнем обсуждение представлением краткого изложения основ, которые необходимо соблюдать для задачи индексирования, понимаемой как утверждение о содержании документов, их близости.

БЛИЗОСТЬ И ИНДЕКСИРОВАНИЕ

Обеспечение термов может быть оправдано в том случае, если соответствующие понятия охватываются вопросами внутри контекста документа. Как правило, это нельзя рассматривать только по отношению к изолированным терминам. Каждое понятие является частью семантического контекста и может рассматриваться в синтаксических связях с другими понятиями. Сумма всех

* Перевод Gödert W. Facets and typed relations as tools for reasoning processes in information retrieval. — 2014.— <http://arxiv.org/ftp/arxiv/papers/1409/1409.0491.pdf>

* Мы отдаем предпочтение термину «близость» вместо слов содержание, предмет или тема; ср. [http://www.iva.dk/bh/coreconcepts in lis/articles a-z/aboutness.htm](http://www.iva.dk/bh/coreconcepts%20in%20lis/articles%20a-z/aboutness.htm)

охваченных понятий и их включение в семантические или синтаксические отношения могут рассматриваться как близость документов. Индексирование должно отражать эту близость посредством представления знания, содержащего концептуальные понятия и семантические отношения между ними. Мы будем обсуждать требования, необходимые при решении некоторых нетривиальных задач индексирования и поиска.

Семантический контекст провоцирует рост установления *априорных* связей между элементами языка индексирования. Обычно существуют три типа отношений – *синонимия*, *иерархия* и *ассоциация*. Решение относительно того, какой тип отношения должен быть установлен между двумя понятиями, основывается на интеллектуальном содержательном анализе понятий. Формальные характеристики, как правило, не задаются ни для определения содержания элементов, ни для отношений. Следовательно, решение, надо ли использовать элемент для представления части близости документов как терм, также основывается на содержании элементов с помощью интеллектуальной интерпретации.

Апостериорные отношения определяют контекст элементов содержания на втором уровне. Та же самая сумма понятий может представлять различную близость, если они связаны разными синтаксическими отношениями. Например, можно представлять роль «быть агентом действия» или «быть объектом действия». На сегодняшний день существует спорная дискуссия относительно того, до какой степени можно иметь дело с этим типом отношений. Их общее число (отношений) представляется слишком большим для включения их всех в среду индексирования и поиска. Более того, как инструменты точности они требуют дополнительного объема работы.

Наша цель заключается в обсуждении некоторых характерных признаков близости и ее представления путем семантических особенностей языка индексирования. Приведем несколько простых примеров. Предположим, нас интересуют *все* документы по *певчим птицам*. Можно думать, что обработка такого поиска требует только одного термина, например, *поющие птицы*, чтобы найти полный результат запроса. Но имеются документы о видах певчих птиц, которые проиндексированы более определенным термом, например, *синицы*. В идеале интеллектуальная структура термов представляет такие связи с помощью иерархических отношений. Автоматически сбор множества всех объектов иерархической цепи может в результате помочь сформировать полный результат запроса для *певчих птиц*. Успех такого подхода зависит от наследования свойств, другими словами, иерархическое отношение должно отражать свойство транзитивности.

Наш второй пример описывается с помощью тем, требующих больше одного термина в качестве формы координатного индексирования и поисковых терминов для посткоординатного поиска, например, сочетания термов путем булевых операторов или других синтаксических элементов. Для наглядности, наш вышеупомянутый пример может быть расширен до темы *миграционное поведение певчих птиц*. Этот случай требует, чтобы *миграционное поведение*, а также *поющие птицы* были проиндексированы и их можно будет искать в сочетании.

Что более важно, так это тип тем, представленных в третьем примере, *певчие птицы с инстинктом миграции*. В этом случае *инстинкт миграции* не должен быть затрагиваемой темой, поэтому он может не индексироваться для документов поиска. Миграционный инстинкт явля-

ется только ограничивающим признаком для отбора соответствующих видов певчих птиц. Существует много видов певчих птиц без такого поведения. Поэтому *миграционный инстинкт* должен быть смоделирован как ограничение для отбора соответствующих объектов представления знания для последующего генерирования результата запроса.

Мы можем обобщить наш пример в основном вопросе: как должны быть даны такие ограничения в представлении знания, чтобы извлечь пользу от индексирования и поддержать механизмы выводов в поисковых средах? Мы вернемся к этому примеру и его обобщению в разделе «Типизированные отношения и выводы».

ИНДЕКСИРОВАНИЕ И ФАСЕТЫ

Основным инструментом интеллектуального индексирования является структурированная совокупность контролируемых терминов. Понятия такого языка индексирования, как правило, расположены в соответствии с их характеристиками, как определяется интеллектуальным анализом. Для построения подходящих иерархий на нескольких этапах цепочек отношений должно наблюдаться наследование этих свойств. Это означает, что каждое иерархически подчиненное понятие имеет все характеристики родового понятия и, по крайней мере, еще одно. Это требование наилучшим образом может быть удовлетворено аннулированием полниерархий или имеющих много измерений связей. Это главным образом означает, что дополнительные характеристики должны происходить из области общего аспекта или определенного категорийного фасета. Если родовый контекст для определения характеристик изменяется, полниерархии становятся неизбежными и результат чаще является докомбинационной содержательной систематизацией с недостаточно выраженными критериями.

В качестве примера можно рассмотреть следующие слова: стол, деревянный стол, стеклянный стол, кухонный стол, стол в гостиной, стол в стиле ар-нуво, рабочий стол, пеленальный стол, приставной стол. Любая попытка организовать эти понятия, или даже расширения, в одну иерархию должна потерпеть неудачу из-за указанных причин. Поэтому подчинение путем изменяющихся аспектов, как это часто случается в докомбинационных классификациях или тезаурусах с расширенным применением концептуальной композиции (например, использованием составных существительных), не может быть подходящим условием для получения умозаключений с помощью относительных выводов.

Чтобы подтвердить выводы, нам необходима фасетная структура, сохраняющая родовые контексты внутри каждой иерархической цепочки. Общеизвестно, что фасетное расположение элементов знания наилучшим образом поддерживает требования индексирования и поиска [2]. Использование терминов фасетного индексирования для представления близости документов позволяет определить каждый компонент содержания на желаемом уровне детализации (гранулярности), не игнорируя структурные требования построения транзитивных цепочек отношений. Рис. 1 может служить в качестве визуализации фасетной структуры, подходящей для темы *Иллюстрированная книга европейских певчих птиц*, составленной по предметам трех аспектов.

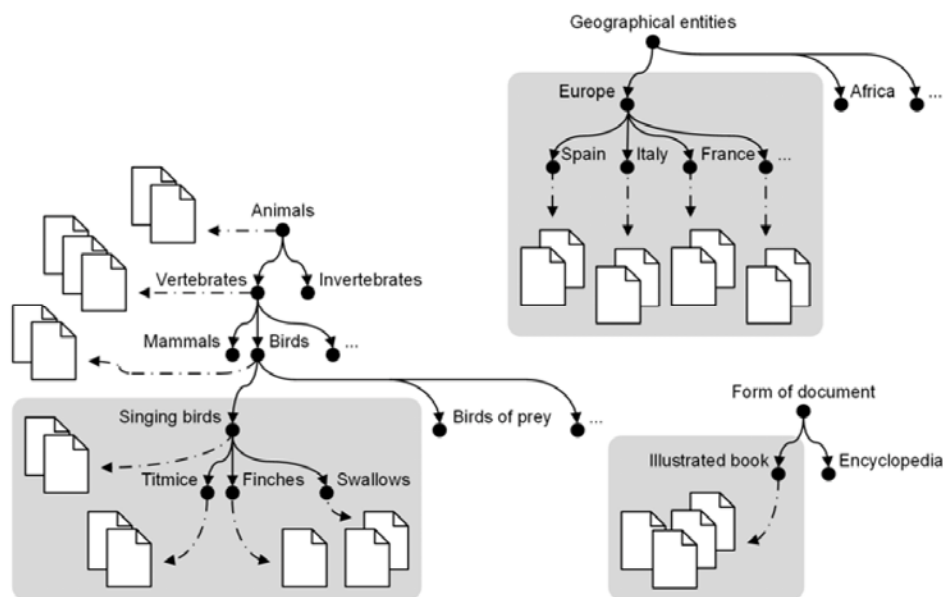


Рис. 1. Структура знания с фасетами, иерархическими выводами и соответствующими документами

Такое расположение имеет большие преимущества для установления правильных иерархий, избегающих неопределенности. Связи, установленные использованием общей культуры, научной традицией или статистическими совместными встречаемостями, могут рассматриваться как полезные, но не могут выражаться отношениями внутри фасетов. Объединение такой темы как *Европейские певчие птицы с инстинктом миграции* в один объект создаст докомбинационную структуру знания с множеством связей между объектами. Она не может быть должным образом связана с объектами только одного фасета без нарушения требований транзитивности и прозрачности окончательной структуры. Мы должны найти другое решение по представлению таких тем с помощью объектов и структурных особенностей представления знания.

ФАСЕТЫ И ТИПИЗИРОВАННЫЕ ОТНОШЕНИЯ

Концептуальные части нашего примера могут различаться по фасетным компонентам. Один фасет может характеризоваться как таксономический фасет (*певчие птицы*), а второй – как поведенческий фасет (*миграционный инстинкт*). Связь между обоими фасетами может быть представлена отношением, определенным как тип ассоциации из перечня типизированных отношений. Общая модель таких связей приведена на рис. 2, поясняющим пример трех типов ассоциативных связей между иерархически организованными объектами четырех фасетов.

Анализ для разработки таких перечней с полученными предложениями приводятся в литературе [12, 19]. Основой этих предложений послужили некоторые теоретические исследования [8,17]. Мы будем использовать краткий перечень [6, 7] для нашего обсуждения в разделе «Типизированные отношения и выводы». При условии, что он характеризуется свойствами транзитивности, можно будет прийти к выводам о цепочках типизированных ассоциативных отношений между объектами различных фасетов, а не только между объектами, связанными иерархическими отношениями. Таким образом мы хотим продемонстрировать потенциал сочетания фасетных структур с выводами о цепочках комбинаций различных типизированных отношений.

ТИПИЗИРОВАННЫЕ ОТНОШЕНИЯ И ВЫВОДЫ

Основанное на правилах (формальное) обоснование поддерживается выводами о соответствующем знании. Для индексирования и поиска это знание задается объектами и структурой между ними. Оба элемента, индексированные объекты, а также отношения между ними, должны поэтому использоваться для формирования окончательных множеств.

В рамках нашего подхода мы различаем два типа процессов фильтрации. Один тип позволяет прийти к выводам об *априорных* утверждениях, моделируемых в представлении знания. Второй тип применяется к результатам процесса индексирования, представляющих *апостериорные* утверждения документа. Этот второй тип является более общим, обычно реализуемым путем использования булевых операторов или других синтаксических элементов.

Рис. 3 представляет визуализацию требований для первого процесса путем использования абстрактных обозначений. Эта иллюстрация подчеркивает важность структурных свойств таких представлений знания по сравнению с простой интерпретацией содержания семантических объектов на интеллектуальном уровне. Обозначения должны читаться следующим образом:

- $E_j F^i$ – Объект j в Фасете i
- $E_x F^i$ – Объект в Фасете i , т.е. в a , являющимся отношением к $E_j F^i$, $x \in \{a-z\}$
- $E_x y F^i$ – Объект в Фасете i , т.е. в a , являющимся отношением к $E_x F^i$, $x \in \{a-z\}$, $y \in \{a-z\}$
- Rel_i - Отношение типа i

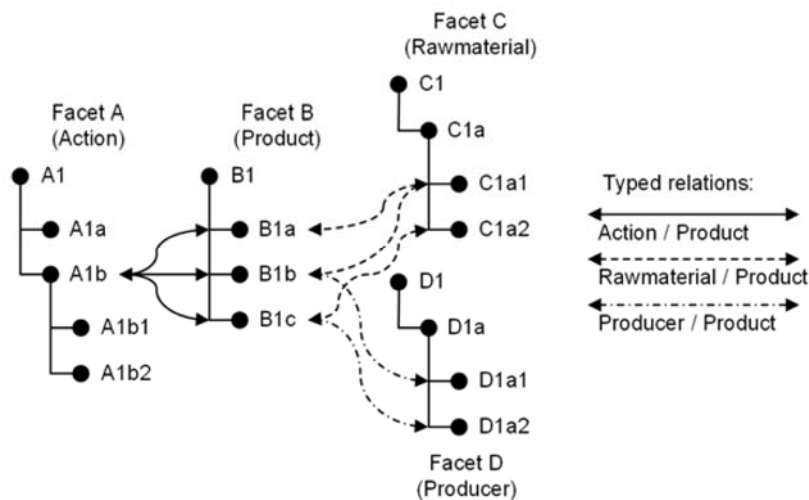


Рис. 2. Фасетная систематическая структура с типизированными отношениями между фасетами

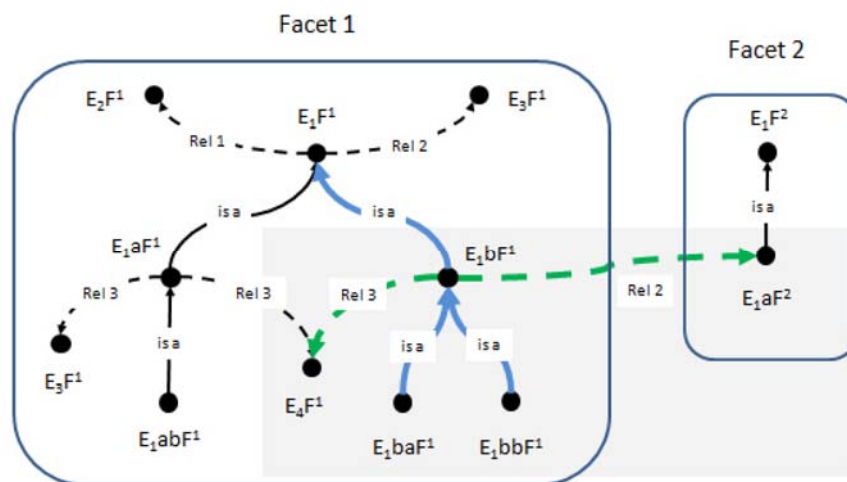


Рис. 3. Представление знания с передачей типизированных свойств и выводами

Эта визуализация поддерживает поисковый интерес, требующий собрания *всех* объектов, характеризуемых как «все объекты фасета 1, которые могут описываться как объекты, относящиеся к E_1F^1 путем a , являющимися отношением в соответствии с ограничением, задаваемым объектом $E_{1a}F^2$ фасета 2», короче говоря $E_{1b}F^1 \xrightarrow{Rel_2} E_{1a}F^2$. Ни объект E_1F^1 , ни объект $E_{1a}F^2$ не должны использоваться в качестве поискового термина, поиск должен обрабатываться только объектами $E_{1b}F^1$.

Рассматривая *певчих птиц с инстинктом миграции* в качестве примера этого абстрактного расположения, ни *певчие птицы*, ни *инстинкт миграции* не должны приписываться в качестве термов к релевантным документам. Мы ищем документы, содержащие утверждения о некоторых понятиях (*певчие птицы*) с определенными свойствами (*имеющие инстинкт миграции*) без самих этих свойств, затрагиваемых в документах.

Фасет 1 представления нашего знания содержит знание об объекте E_1F^1 (*полющие птицы*), например о том, что $E_{1b}F^1$ является видом певчих птиц. Типизированное отношение (Rel_2) по отношению к объекту $E_{1a}F^2$ фасета 2

накладывает ограничивающее условие (*имеет инстинкт миграции*) для всех объектов $E_{1b}F^1$ фасета 1, что не справедливо для E_1F^1 . Вывод об иерархических цепочках отношений позволяет образовать множество всех подходящих певчих птиц и таким образом провести соответствующий поиск. Пунктирная линия, связывающая $E_{1b}F^1$ с $E_{1a}F^2$, дает иллюстрацию к данному выводу.

Этот пример обеспечивает надлежащую иллюстрацию вышеупомянутого абстрактного правила выбора соответствующего иерархического узла для его связи с другим узлом второго фасета путем типизированного отношения, чтобы избежать необходимости в упоминании исключений. Для лучшего понимания мы проиллюстрируем это утверждение с помощью обсуждения трех рисунков, представляющих различные малые структуры знания. Не представляет труда обобщить примеры структурного содержания до общей ситуации.

Рис. 4 прежде всего показывает приписывание типизированных отношений к объекту *Полющие птицы*, что должно квалифицироваться как некорректное. Одно из

этих отношений задано – «*Поющие птицы*, имеющие *Миграционную модель*».

Безусловно, есть певчие птицы, имеющие миграционный инстинкт, но это не является справедливым для *всех* видов. Решение смоделировать этот факт может быть вызвано тем, чтобы связать только подчиненные виды с *Миграционной моделью*, которые действительно имеют миграционный инстинкт. Несмотря на множество усилий, это решение ведет к отсутствию прозрачности в рамках структуры знания. Этот факт показан на рис. 5 для связи только очень небольшого числа примеров. Он должен пониматься как представление общей ситуации, возможно, с сотнями связей.

Другим решением этой проблемы может служить применение исключений на основании принципов экспертных систем.

Наше предложение, данное на рис. 3, сочетает главным образом два элемента:

- выбор соответствующего иерархического узла для его связи с другим узлом путем типизированного отношения,
- выводы относительно цепочек отношений.

Рис. 6 разъясняет общую установку для нашего примера. Не все виды певчих птиц связаны с *Миграционной моделью*, а только те, у кого есть миграционный инстинкт. Во избежание опасности построения структур с меньшей прозрачностью и для обеспечения гарантии ориентированной на содержание последовательности формальных свойств только виды соответствующего высшего иерархического уровня были выбраны для связей.

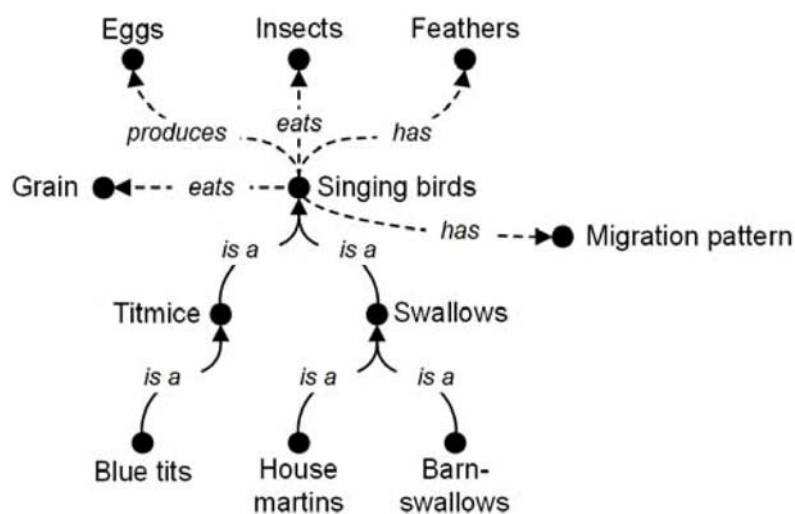


Рис. 4. Структура знания с некорректной связью типизированных отношений

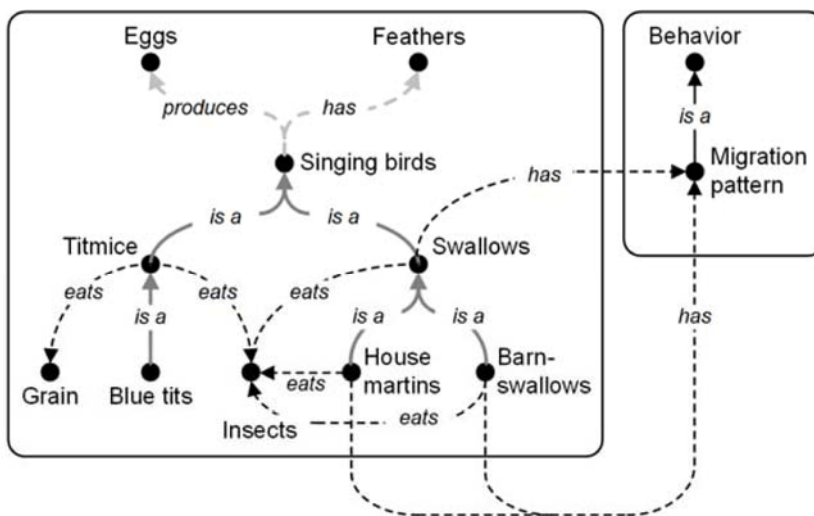


Рис. 5. Структура знания с передачей типизированных свойств, но без выводов

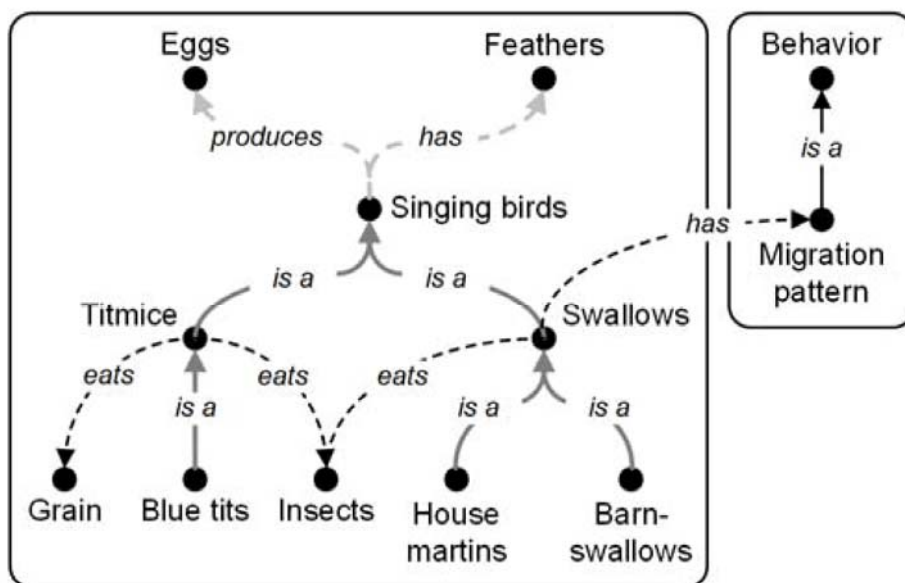


Рис. 6. Представление знания с передачей типизированных свойств и выводами

Дальнейшее улучшение для использования выводов в области информационного поиска задается построением множеств наравне с цепочками отношений сочетаний различных связей. Как упоминалось, такой механизм недостаточно хорошо поддерживается традиционными языками индексирования. Поэтому может быть выгодным проанализировать различные комбинации семантических *априорных* отношений, чтобы выдвинуть некоторые предложения о транзитивных выводах относительно цепочек отношений. Из-за основанных на содержании решений о типе отношений в языках индексирования невозможно предоставить точные доказательства для таких утверждений транзитивности. Особенно для ассоциативных отношений наблюдается ограничение аргументами правдоподобия. Представляем краткое изложение наших наблюдений в трех таблицах, где используем следующий улучшенный перечень типичных отношений [6,7].

- Эквивалентность
 - Синоним
- Иерархия
 - Абстракция, родовой контекст
 - Целое/Часть
 - Абстракция, родовой контекст
 - Целое/Часть
- Хронологический контекст
 - Раньше/Позже
 - Позже/Раньше
 - Раньше/Позже
 - Позже/Раньше
- Ассоциация
 - Неопределенная ассоциация

- Сырье/ продукт
- Причинность (причина - следствие)
- Лицо как деятель/ действие
- Учреждение как деятель/ действие
- Лицо как деятель/продукт
- Учреждение как деятель/ продукт
- Действие/ продукт

Руководящим принципом для наших утверждений транзитивности всегда является наследование свойств. Как уже обсуждалось, для иерархических отношений это поддерживается фасетным подходом, с тем чтобы избежать неопределенностей. Более подробный анализ и представление аргументов даны в [7].

Утверждения транзитивности всегда указаны в последней колонке символами:

- +: Транзитивность дана
- : Транзитивность нельзя ожидать
- O: Не допустимо для языков индексирования

Табл. 1 показывает утверждения транзитивности для сочетания отношений одного вида.

Табл. 2 показывает утверждения транзитивности для сочетания различных типов иерархических отношений.

Наконец, табл. 3 показывает утверждения транзитивности для сочетания различных типов иерархических отношений с типизированными ассоциативными отношениями. Включены сочетания только с позитивным утверждением.

Такой список утверждений транзитивности может использоваться для формального отображения в представлениях знания при характеристике свойств сочетаний типизированных отношений для разработки семантически улучшенных поисковых сред.

Транзитивность в случае одного вида отношений

Тип отношения	Отношение 1	Отношение 2	Транзитивность
Эквивалентность	Синоним	Синоним	О
Иерархия	Абстракция, родовой контекст	Абстракция, родовой контекст	+
	Целое/Часть	Целое/Часть	+
	Абстракция, родовой контекст	Целое/Часть	-
	Целое/Часть	Абстракция, родовой контекст	-
Хронологический контекст	Раньше/Позже	Раньше/Позже	+
	Позже/Раньше	Позже/Раньше	+
	Раньше/Позже	Позже/Раньше	-
	Позже/Раньше	Раньше/Позже	-
Ассоциация	Неопределенная ассоциация	Неопределенная ассоциация	-
	Сырье/продукт	Сырье/продукт	+
	Причинность (причина- следствие)	Причинность (причина- следствие)	+
	Личность как деятель/действие	Личность как деятель/действие	-
	Учреждение как деятель/действие	Учреждение как деятель/действие	-
	Личность как деятель/продукт	Личность как деятель/продукт	-
	Учреждение как деятель/продукт	Учреждение как деятель/продукт	-
	Действие/продукт	Действие/продукт	-

Транзитивность в случае различных типов иерархических отношений

Тип отношения	Отношение 1	Отношение 2	Транзитивность
Иерархия	Синоним	Абстракция, родовой контекст	+
	Синоним	Целое/Часть	+
	Абстракция, родовой контекст	Синоним	О
	Абстракция, родовой контекст	Синоним	О
	Целое/Часть	Синоним	О
	Целое/Часть	Синоним	О
Хронологический контекст	Синоним	Раньше/Позже	+
	Синоним	Позже/Раньше	+
	Раньше/Позже	Синоним	О
	Позже/Раньше	Синоним	О
	Абстракция, родовой контекст	Раньше/Позже	+
	Абстракция, родовой контекст	Позже/Раньше	+
	Раньше/Позже	Абстракция, родовой контекст	+
	Позже/Раньше	Абстракция, родовой контекст	+
	Целое/Часть	Раньше/Позже	+
	Целое/Часть	Позже/Раньше	+
	Раньше/Позже	Целое/Часть	+
	Позже/Раньше	Целое/Часть	+

Транзитивность для сочетаний типизированных ассоциаций с иерархическими отношениями

Тип отношения	Отношение 1	Отношение 2	Транзитивность
Ассоциация	Неопределенная ассоциация	Абстракция, родовой контекст	+
	Неопределенная ассоциация	Целое/Часть	+
	Неопределенная ассоциация	Раньше/Позже*	+
	Сырье/продукт	Абстракция, родовой контекст	+
	Сырье/продукт	Целое/Часть	+
	Сырье/продукт	Раньше/Позже*	+
	Действие/продукт	Абстракция, родовой контекст	+
	Действие/продукт	Целое/Часть	+
	Действие/продукт	Раньше/Позже*	+
	Личность как деятель	Абстракция, родовой контекст	+
	Личность как деятель	Целое/Часть	+
	Личность как деятель	Раньше/Позже*	+
	Учреждение как деятель	Абстракция, родовой контекст	+
	Учреждение как деятель	Целое/Часть	+
	Учреждение как деятель	Раньше/Позже*	+
	Причинность (причина-следствие)	Абстракция, общий контекст	+
	Причинность (причина-следствие)	Целое/Часть	+
	Причинность (причина-следствие)	Раньше/Позже*	+
	Личность как деятель/продукт	Абстракция, родовой контекст	+
	Личность как деятель/продукт	Целое/Часть	+
	Личность как деятель/продукт	Раньше/Позже*	+
	Учреждение как деятель/продукт	Абстракция, родовой контекст	+
	Учреждение как деятель/продукт	Целое/Часть	+
	Учреждение как деятель/продукт	Раньше/Позже*	+

* Также: Позже/Раньше

СЕТЕВОЙ ПОИСК И СЕМАНТИЧЕСКОЕ ВЗАИМОДЕЙСТВИЕ: ПЕРСПЕКТИВА

Представленные в этой статье идеи могут быть использованы в любых средах индексирования и поиска. Но они особенно полезны при применении в неоднородных информационных средах, например в контексте применений семантической сети и ее стандартов для семантического представления. Чтобы продемонстрировать эту выгоду в принципе, прототип с поисковой формой был построен в *Институте информатики в Кельнском университете прикладных наук*, он может быть доступен через сетевой интерфейс (<http://ixtrieve.fh-koeln.de/ghn/>). Этот прототип был основан на существующих данных, взятых из базы данных *Literatur zur Informationserschließung (Literaturpa no информатике)* (<http://ixtrieve.fh-koeln.de/LitIE>). Охват документов соответствует выдержке из тезауруса *ASIS&T Thesaurus* [1], которая была преобразована в *тематическую карту* ([4,10,15,16]) путем использования программного обеспечения *Ontopia* (<http://www.ontopia.net>). Эта среда выбрана за ее особый потенциал представлять объекты и отношения между ними с помощью типизированных отношений как тематические карты. Характеристики *Ontopia* включают визуализацию тематической карты и возможность поддержки сетевой поисковой системы. Мы воздерживаемся от предоставления подробностей или примеров в этой статье, больше материала приводится в [6,7].

Сетевой поиск и неоднородные информационные среды содержат вопросы семантического взаимодействия. По сравнению с уже представленными предложениями в рекомендациях *SKOS* [9,13,14] или в *ISO 25964* [11] использование типизированных отношений может обеспечивать современную поддержку внедрения процессов выводов и методов фасетного поиска в неоднородных языках индексирования. Стандарты семантического пред-

ставления *RDF* или *OWL* поддерживают совершенствование существующих отношений без необходимости полной реорганизации языка индексирования.

Проекты дальнейшего исследования и практического применения будут необходимы для получения выгоды в реальных информационных средах, представляющих запись данных, индексированные одной или несколькими хорошо известными системами организации знания. Эта статья могла представить только схему этого подхода. Любая реализация установки в обстановке существующего мира требует сочетания уже выработанных методик связанных данных с интенсивным реструктурированием систем организации знания, используемых в неоднородных средах индексирования и поиска.

ЛИТЕРАТУРА

1. *ASIS&T Thesaurus of Information Science, Technology and Librarianship*/A. Redmond-Neal, M.M.K. Hlava (Eds.). 3rd ed. — Information Today, Medford, NJ, 2005.
2. *Broughton V.* Faceted classification as a basis for knowledge organization in a digital environment: The Bliss bibliographic classification as a model for vocabulary management and the creation of multi-dimensional knowledge structures// *New review of hypermedia and multimedia*. — 2001.— Vol. 7, No.1.— P. 67-102.
3. *Buxton A.B.* Computer searching of UDC numbers// *Journal of Documentation*. — 1990.— Vol. 46, No.3.— P. 193-217.
4. *Garsbol L.M.* Metadata? Thesauri? Taxonomies? Topic maps!: Making sense of it all// *Journal of Information Science*. — 2005. — Vol. 30, No.4. — P. 378-391.
5. *Gödert W.* Die Dezimalklassifikation im Online-Retrieval// *Nachrichten für Dokumentation*. — 1990. — Vol. 41, No.3.— P. 155-158.

6. *Gödert W.* Ein Ontologie basiertes Modell für Indexierung und Retrieval// Information - Wissenschaft und Praxis. — 2014.— Vol. 65, No.2. — P.83-98; an English version has been accepted for publication in the Journal of the Association for Information Science and Technology. A preliminary version can be found under: <http://arxiv.org/abs/1312.4425>
7. *Gödert W., Hubrich J., Nagelschmidt M.* Semantic knowledge representation for information retrieval. — DeGruyterSaur, Berlin, 2014.
8. *Green R., Bean C. A.* Aligning systems of relationships. Knowledge organization, information systems and other essays: Professor A. Neelameghan Festschrift/ K.S. Raghavan, K.N. Prasad (Eds.)// Ess Ess Publ., New Delhi. — 2006. — P.111-128.
9. *Isaac A., Summers E.* SKOS simple knowledge organization system primer// W3C working group note. — 2009. — <http://www.w3.org/TR/skos-primer>
10. *ISO 13250.* Topic maps. Information technology.— ISO, Geneva,1999.
11. *ISO 25964.* Thesauri and interoperability with other vocabularies: Part 1: Thesauri for information retrieval. Part 2: Interoperability with other vocabularies. — ISO, Geneva, 2011-2013.
12. *Michel D.* Taxonomy of subject relationships. — 1997. —<http://web2.ala.org/ala/alctscontent/CCS/committees/subjectanalysis/subjectrelations/mrsrscu2.pdf>// Part of: Final Report to the ALCTS/CCS Subject Analysis Committee. — June 1997. — <http://web2.ala.org/ala/alctscontent/CCS/committees/subjectanalysis/subjectrelations/finalreport.cfm>
13. *Miles A., Bechhofer S.* SKOS Simple Knowledge Organization System// W3C recommendation. — 2009. — <http://www.w3.org/TR/skos-reference>
14. *Miles A., Brickley D.* SKOS mapping vocabulary specification. — 2004. — <http://www.w3.org/2004/02/skos/mapping/spec/2004-11-11.html>
15. XML topic maps: creating and using topic maps for the Web/J.Park , S. Hunting (Eds). — Addison-Wesley, Boston, 2003.
16. *Pepper S.* The TAO of topic maps: Finding the way in the age of infoglut. — <http://www.ontopia.net/topicmaps/materials/tao.html>
17. Relationships in the organization of knowledge/ C. A. Bean, R. Green (Eds.). — Kluwer Academic, Boston, 2001.
18. *Shiri A.* Powering search: The role of thesauri in new information environments. — ASIST, Medford, NJ, 2012.
19. *Tudhope D., Alani H., Jones C.* Augmenting thesaurus relationships: Possibilities for retrieval// Journal of Digital Information. — 2001. — Vol. 1, No.8. — <https://journals.tdl.org/jodi/index.php/jodi/article/view/181>

Отчет о библиотеках в цифровом веке (LIDA 2014)*

Дарко ЛАКОВИЧ

(Darko LACOVIC)

Осиекский университет,
г. Осиек, Хорватия

Матэ ЮРИЧ

(Mate JURIC)

Задарский университет, г. Задар, Хорватия

«Библиотеки в цифровом веке» (Libraries in the Digital Age – LIDA) – это международная, проходящая раз в два года, конференция, центром внимания которой является переход библиотек и информационных служб в цифровую среду. Она состоялась 16 – 20 июня 2014 г. в г. Задар (Хорватия). Основной темой конференции LIDA была «Оценка и использование библиотек и библиотечные пользователи», затрагиваемая многими презентациями, рабочими семинарами и постерами. Данный отчет посвящен сессиям, на которых присутствовали авторы и которые отражают их собственные научные интересы.

ВВЕДЕНИЕ

В признании развивающихся социальных и технологических влияний, представляющих как проблемы, так и возможности, «Оценка» была приоритетной темой LIDA 2014. Тематика конференции была поделена на две части: в первой затрагивались преимущества в качественной оценке методов и практик, а вторая охватывала методы оценки, включающие альтернативные метрики на основе социальных медиа и широкий ряд видов коммуникационной деятельности, обычно относящихся к альтметрии.

Конференция LIDA проводилась 16-20 июня 2014 г. в г. Задар, Хорватия. Общее число участников – 109 человек. 60 представляло 21 страну, 27 участников были из Хорватии и 22 являлись студентами области библиотековедения и информатики. Было представлено 34 доклада, написанных 75 авторами, включая почетного гостя и 7 приглашенных докладчиков. Трое аспирантов участвовали в форуме аспирантов, спонсируемом ASSIST. Также было проведено 4 семинара в рамках этой конференции, 19 постеров представлено 32 авторами. Конференция LIDA была организована Татьяной Апарак-Елусич (Отделение библиотековедения и информатики, Задарский университет, г. Задар, Хорватия) и Тефко Сарацевичем (Школа коммуникации и информации, Ратгерский университет, шт. Нью-Джерси, США).

Данный отчет фокусируется на сессиях, на которых присутствовали авторы и которые отражают их научные интересы. Дополнительную информацию можно найти на сетевом сайте конференции (<http://ozk.unizd.hr/lida/>).

ЧАСТЬ 1: КАЧЕСТВЕННЫЕ МЕТОДЫ

Первая часть конференции LIDA была посвящена качественным методам в оценке библиотек, пользователей и использования: применения, результаты. Она возглавлялась Дэвидом Боуденом (Центр информатики, Университет Сити, Лондон). Целью этой части темы «Оценка» было изучить попытки, понятия и результаты в использовании качественных методов в оценке библиотечного влияния, значимости, эффективности и использования новых и старых служб. Результаты качественной оценки могут улучшить библиотечные службы и стимулировать более эффективное использование благодаря множеству пользователей в развивающихся технологических и социальных средах.

Конференция открылась основным докладом проф. Боудена под названием «Наивысшее удовольствие?: о получении понимания от качественного исследования в библиотековедении/информатике». В своей презентации он особо выделял отношение между качественными и количественными методами в библиотековедении и информатике, утверждая, что они могут сочетаться. Более того, Боуден привел интересные примеры такого подхода сочетания методов из естественных наук.

Приглашенный докладчик Элке Грейфенедер (Королевская школа библиотековедения и информатики, Дания) провел презентацию под названием «Большие данные не равны большой картине», показывая, как некоторые направления в исследованиях информационного поведения влияют на сбор данных, и приводя примеры того, как интерпретация результатов онлайн интервью и случайных исследований не всегда приводит к правильным выводам. Поэтому, «большая картина» не всегда является «правильной картиной».

Название презентации, подготовленной Россом Дж. Тоддом и Пунитом Дадлани (Ратгерский университет) – «Совместное использование информации студентами высшей школы в цифровой образовательной среде».

* Перевод Lacovic D., Juric M. Report on Libraries in the digital age (LIDA 2014). — 2014. — <http://www.dlib.org/dlib/july14/lacovic/07lacovic.html>

Демонстрируемые результаты их исследования показали, что студенты ценят допустимость создания групповой работы в отношении создания знания, образовательных результатов, распределения рабочей нагрузки, справедливости образования, коллегиальности и сотрудничества. Росс изучил понятие «социальная справедливость», которое родилось из собранных данных, и выяснил, что оно является важным фактором, который необходимо учитывать при планировании совместных задач.

Саньика Фалетар Танакович, Дарко Лакович и Гордана Гасо (университет Й.Ю. Штрессмайера, г. Осиек, Хорватия) представили результаты качественного исследования, ненавязчивого наблюдения исследуемых пространств в академической библиотеке, показавшего, что преобладающими видами деятельности студентов были чтение, составление записей и использование компьютеров в научных целях (например, MS Office) и для развлечения (Фейсбук, Твиттер и т.д.). В результате библиотека улучшила некоторые свои пространства и услуги (столы стали больше, выросла численность компьютеров) в целях лучшей поддержки поведения студентов во время работы с документами.

Приглашенный докладчик Дональд Кейс (Университет шт. Кентукки, США) уделил внимание недавнему росту измерений информационных результатов, представляя анализ контента выборок за 62 года изучения информационного поведения и делая предположения о развитии новых видов измерений. Эти представления могут быть весьма полезны для осуществления дальнейших исследований информационного поведения людей.

Почетный гость проф. Гари Марчонини (Университет шт. Северная Каролина на Чепел Хилл, США) провел презентацию под названием «Библиотеки и грамотность: от «я» к «мы»». Он объяснил, как библиотеки могут «собрать людей», помогая им правильно управлять своими разнообразными профилями времяпрепровождения и киберидентичности, такими как электронные записи о здоровье, персональном геноме, цифровых установлениях личности (например, паролях), электронные записи об обучении и финансовые профили и отчеты о кредитах, собранные из таких источников, как Google Scholar, MS Academic, ResearchGate и Klout, используя доверия идентичности для создания «соединительных сетей» и облегчения процессов чтения и письма. Докладчик показал, как библиотеки могут обслуживать различные группы людей.

ЧАСТЬ 2: АЛЬТМЕТРИЯ

Вторая тема конференции, проводившейся под председательством проф. Блейза Кронина (Школа информатики и вычислительной техники, Университет шт. Индиана, США), охватывала методы оценки, включающие альтернативные метрики на основе социальных медиа, и широкий круг коммуникационных видов деятельности, обычно относящихся к «альтметрии». Хотя качественные методы имеют давно существующую традицию применения, альтметрия является совокупностью новых подходов к фильтрации информации и оценке научной коммуникации и исследований с возможным применением также и к библиотекам.

Кронин представил превосходное введение к сессиям по альтметрии в презентации, названной «Бетховен в противоположность Биберу: о содержательности

(альт)метрии». Одним из его заключений было то, что временный интерес часто неправильно интерпретируется как показатель высокого качества или даже как показатель более глубокого социального влияния. Кронин предполагает, что альтметрия и даже социальные метрики могут использоваться как показатели профессиональной ориентации и мотивирования работников, но не как факторы определения уровня заработной платы. Причиной этого служит то, что такие метрики не являются вполне надежными и заслуживающими доверие мерами рабочей эффективности. Тем не менее, метрики — очень полезные показатели при соответствующем использовании. Другим интересным замечанием было то, что интернет-профили ученых и графы цитирований в реальном времени питают их нарциссизм. Это не упоминалось в докладе, но, с другой стороны, умеренное использование таких интернет-профилей полезно для здоровой самоактуализации.

Приглашенный докладчик Пол Уотерс (Центр научных и технологических исследований, Нидерланды) представил научный проект «Научные карьеры, понимаемые посредством измерений и норм», сокращенно «ACUMEN» (Academic careers understood through measurements and norms). Метрический ACUMEN включает источники данных по библиометрии и вебометрии, а также использует описание, представляющее самоперспективу и придающее большое значение достижениям, амбициям и интересам ученых. Метод ACUMEN, как ожидается, будет полезен в планировании карьеры и помощи ученым в процессах оценки. Остается открытым вопрос, будет ли этот метод действительно использоваться широким научным сообществом, особенно потому, что у многих ученых нет времени или желания помимо их постоянной работы заниматься самотражением и самооценкой.

Примоз Южнич, Полон Вилар и Томаж Бартол (Люблянский университет, Словения) представили результаты исследования, показывающего, что ученые из гуманитарных областей имеют преимущество в использовании данных о числе загрузок, и как ожидается, они, как правило, выступают оппонентами использования данных цитирования. Большая часть ученых из Словении по крайней мере частично согласилась, что оценка научных результатов должна применять средства социальной сети.

Приглашенный докладчик Кассиди Сугимото (Университет шт. Индиана, США) в своей презентации «Метрическая коллекция: Твиттер, лайки и другие метрики социальных медиа в библиотеке» отметила, что альтметрия не должна быть альтернативой метрик по цитированию, а должна быть чем-то за рамками этого. Она также настоятельно советовала научному сообществу серьезно подумать об альтметрии и всех метриках и задаться вопросом: являются ли они измерительными объектами? Служат ли улучшению общественного достояния? Реально ли измеряют социальное влияние?

В конце сессии приглашенный докладчик Мария Брайди-Вукович (Загребский университет, Хорватия) рассказала о проблемах хорватской науки в презентации «Проблемы практик глобальной оценки в контексте полупериферийного и локального производства знания». Она обсуждала, как научное сообщество предупреждается о последствиях, связанных с «международным поворотом», для хорватского общества и культуры, хорватского языка как языка науки и хорватских научных журналов и издательств.

ЗАКЛЮЧЕНИЕ

Представленные доклады и дискуссии часто освещали важность учета различных особенностей многих областей исследования. Идея, затронутая во множестве дискуссий, состоит в том, что сведение данных в одну единственную меру является непониманием сути проблемы. Наличие множества показателей, включая усовершенствование альтметрии, важно не как замещение метрик по цитированию, а как нечто, выходящее за рамки этого. Этим «нечто» является более широкое влияние и более полезные результаты. Обе темы конференции – качественные методы и альтметрия – требуют надежного измерения результатов и показателей влияния. Поэтому ряд докладчиков осветил важность точной организации различных показателей и строгих правил научной методологии.

Тема использования социальных медиа в библиотеках слегка охватывалась в презентациях докладов, а сессия постеров заполняла этот пробел. Конференция

удовлетворила свои цели тем, что она предложила ответы на многие вопросы относительно того, как оценивать библиотеки, их использование и пользователей, и как оценивать и собственно интерпретировать научные вклады. Может быть даже более важно то, что конференция также подняла новые вопросы об измерении результатов использования библиотек и научной работы. Более того, значимой целью LIDA являлось объединение ученых сферы информатики, и эта цель безусловно была достигнута.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Более подробная информация о 13-й конференции LIDA, включая ссылки на все презентации и доклады, находится на сайте конференции (<http://ozk.unizd.hr/lida/>). Конференция LIDA 2016 снова объединит ученых, преподавателей и участников со всего мира на форуме для обмена персональными мнениями, дискуссиями и обучением в очаровательном и особенно красивом г. Задар на берегу Адриатического моря.

Авторский указатель к МФИ, 2014, т. 39

	№	Стр.		№	Стр.		№	Стр.
Берш Б.	2	21	Кена Р.	2	21	Николас Д.	4	3
Бжезински М.	3	8	Кормод Г.	2	3	Рамайя Т.	2	29
Борнман Л.	2	13	Костас Р.	3	20	Тан К. Н.- Л.	2	29
Ван К.	3	3	Лакович Д.	4	29	Уотерс П.	3	20
Вуд М.	1	30	Ларивьер В.	4	11	Франнсен Т.Ф.	1	9
Гёдерт В.	4	20	Матакришнан С.	2	3	Юрич М.	4	29
Головач Ю.	2	21	Монжон Ф.	4	11	Ян Ц.	2	3
Далтон М.	1	14	Мриголд О.	2	21	Яффе К.	1	3
Захеди З.	3	20	Николаисен Й.	1	9			

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 10.12.2014 г.

Бумага “Хегох”. Формат 60x841/8. Гарн. литер. Печать цифровая

Усл. печ. л 4,00 Уч.-изд. л. 4,87 Тираж 38 экз.

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

Тел. (499) 155-44-95