

СОДЕРЖАНИЕ

Кобзева Л. В., Овченкова Е. А. История и современность: журналу «Международный форум по информации» - 45 лет!	3
Бу И., Мюррей Д. С., Дин И., Сюй Ц., Ай П., Шэнь Ц., Ян Ф. Анализ научного сотрудничества с «гигантами» на основе важных этапов карьеры	5
Хенкок Д. Т., Нааман М., Ливи К. Опосредованная искусственным интеллектом коммуникация: определение, повестка исследований и этические размышления	19
Паскуа С., Ли К. Построение матрицы активности метаданных: концептуализация и обоснование связей между жизненным циклом исследований и метаданными	27
Хендерсон М. Почему необходимы мягкие и нетехнические навыки для успеха в профессии библиотекаря данных	33
АВТОРСКИЙ УКАЗАТЕЛЬ к МФИ, 2020, т. 45	37

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ГЛАВНЫЙ РЕДАКТОР

д.филол.н. ГИЛЯРЕВСКИЙ Р.С.

ЧЛЕНЫ РЕДКОЛЛЕГИИ:

к.т.н. БЫКОВ В.А. (РОССИЯ), к.физ.-мат.н. ВАРДАНЯН Г. Г. (АРМЕНИЯ),
д.т.н., проф. ВОЙТОВ И. В. (БЕЛАРУСЬ), МИРАЛИЕВ К. Х. (ТАДЖИКИСТАН),
МОЛДОШЕВА Д. А. (КЫРГЫЗИЯ)

РЕДАКТОРЫ:

КОБЗЕВА Л.В., ОВЧЕНКОВА Е.А.

История и современность: журналу «Международный форум по информации» – 45 лет!

Форум - неофициальное название журнала «Международный форум по информации». Сегодня слово *форум* можно встретить в СМИ и на телевидении. Оно звучит во многих контекстах: Дальневосточный экономический форум, Санкт-Петербургский экономический форум, форум как хранилище открытого доступа и т.п., что подтверждает его актуальность и многозначность. Наш журнал - тоже форум, печатный орган стран - членов Межгосударственного координационного совета по научной и технической информации (МКСНТИ).

В декабре текущего года нашему Форуму исполняется 45 лет. Поговорим о его достижениях.

Оглядываясь назад, отметим, что отдельные статьи, определенные выпуски и даже тома Форума связаны с работой и интересами основанной его организации – Международной федерации по документации (МФД), печатным органом которой Форум был до конца XX в. Тогда журнал был международной площадкой формирования новой научной области - информатики. Это золотое время Форума. Его авторы и читатели – ученые, разрабатывающие новую область науки – информатику. Это и Ч. Оппенхайм, и Д. Фоскетт (1918 – 2004), а также Б. Веккери (1918 - 2009), несомненно А. И. Черный (1929 - 2013), Р. С. Гиляревский, В. А. Маркусова и др.

Чем примечателен Форум - на его страницах ощущается современность. Когда шло активное становление новой сферы деятельности - научно-информационной (конец XIX в. – 1950-е гг.), нашего Форума не существовало, в это время происходил научный бум – шел сбор, накопление и распространение, а также обмен научными материалами, а с появлением в мире организаций, аккумулирующих эти процессы, они продолжили ускоряться. И в такой атмосфере в 1975 г. вышел в свет первый номер журнала «Международный форум по информации и документации» (МФИД) параллельно на двух языках - английском и русском. По его публикациям как раньше, так и сейчас можно отследить актуальное состояние одной из областей науки - информатики.

Наш Форум пользовался вниманием многих ученых из разных стран мира. Председатель редколлегии *IFLA Journal* Рассел Боуден так описывает значимость Форума: «Его значение как журнала, посвященного вопросам, связанным с коммуникациями в науке, традиционными и передовыми методами представления и распространения информации, прогнозированием и планированием тенденций развития информационных и библиотечных

служб и систем, теоретическими основами информатики, классификационно-информационными языками и подготовкой информационного персонала и пользователей информации, отводит ему центральную и значимую роль на этапе коммуникации для библиотечно-информационной науки и документоведческой профессии. *МФИД* и *IFLA Journal*, председателем редакционного комитета которого я имею честь быть, играют важную роль в коммуникации между специалистами во всем мире». Эта оценка роли нашего журнала во второй половине XX в. – живое подтверждение его значимости и востребованности со стороны ученых всего мира.

Начало XXI века Форум встретил изменениями, как в оформлении, так и в названии. Вернувшись в оформлении обложки к белому цвету, он сохранил свою связь с прошлым, но получил более ёмкое и актуальное название – «*Международный форум по информации*». Качество, стиль написания текста, авторский стиль материалов – все стало более информативным, лаконичным, идущим в ногу со временем. Объектом изучения журнала остается любой носитель фиксированной информации и его роль и функции в информатике, а также потребности пользователей информации.

Информация стала ключевым фактором развития нашего журнала. Он публикует материалы по следующим темам: информационная политика; научная коммуникация; направления в разработке форм отображения и проникновения информации (мультимедиа и гипермедиа); предсказание появления и внедрения систем, служб и сетей (международных, региональных, национальных); разработка и эксплуатация баз данных; основные теории информатики; классификации и информационные языки; информационные потребности; подготовка специалистов в области информации; обучение реципиентов; библиометрия; ценность информации. Разнообразие тем нашего Форума подтверждает его преемственность и свидетельствует о многосторонности интересов журнала.

Впервые на русском языке в Форуме были опубликованы фрагменты трудов классиков информатики – В. Буша «Возможный механизм нашего мышления» (2001), Ф. Махлупа «Культурное разнообразие в изучении информации» (2004), Д. Фоскетта «Пути коммуникации: книги и библиотеки в информационном веке» (2005), Б. Веккери «Мемопланета» (2006); «Информационная наука» (2006).

Наш Форум отслеживает новые тенденции в рамках приведенных выше рубрик. Он был в когорте журналов, подхвативших новое веяние в информатике, науку о данных. Так в Форуме за 2012 г. (Т. 37, № 3) была напечатана статья М. Дж. Грехема «Искусство науки о данных», осветившая новое направление в информатике. Форум не претендует на пальму первенства в раскрытии этой темы в России, но он внес свой вклад в интерес к ней. Тогда данная тема только начала свое развитие и не было сформулировано определение науки о данных. Спустя три года Форум снова обращается к этой теме – в 2015 г. он публикует статью авторов из Китая Я. Чжу и Ю. Сяона «Определение науки о данных». В ней авторы рассматривают науку о данных с трех оформленных на тот момент точек зрения. Статья ценна тем, что в ней предпринята попытка дать определение новой науке и описывается развитие темы, т.е. на ее страницах можно прочувствовать интерес авторов к этой формирующейся области знания.

От науки о данных вернемся в прошлое и рассмотрим вклад Форума в знакомство с новыми тенденциями в информатике. Так в 2010 г. Форум находит новый аспект – на первой полосе (Т. 35, № 3, с. 3-11) публикуется статья Э. Б. Кохена «Философия науки информирования». В ней снова поднимается вопрос о важности информирования специалистов актуальной для них информацией. Информирование было и раньше, но статья Э. Б. Кохена возродила интерес ученых к этой области науки. И данный интерес был столь серьезен, что в США создается журнал – *Informing Science: The international journal of an emerging transdiscipline*, и образуется отдельный институт – the Informing Science Institute. Появление в США данного журнала свидетельствует о новом витке в этой области, о малой освещенности такой темы в науке. И безусловно, учреждение института – the Informing Science Institute (США), занимающегося вопросами изучения проблем информирования, подтверждает осознанность интереса к этой проблематике. И Наш Форум знакомит своих читателей с новым взглядом на науку информирования.

Форум не боится экспериментировать, открывать для своих читателей новые имена - В. Ларивьера, И. Джэнга, С. Миццаро, Б. Кронина и Л. И. Мехо, а также других. Отдельные работы этих ученых можно найти в его выпусках.

В Форуме, как правило, не принято акцентировать внимание на работе и жизни отдельных ученых. Но в 2018 г. Форум отдает дань уважения одному из выдающихся ученых в информатике - Юджину Гарфилду

(1925-2017). Он печатает триаду публикаций, посвященных памяти выдающегося ученого. В докладах, подготовленных к церемонии его памяти (США, Филадельфия, 15 сентября 2017 г.), Юджин Гарфилд описывается и как ученый, и как человек теми авторами, которые на протяжении не одного десятка лет работали с ним рука об руку и знали его довольно хорошо. Эти авторы заслуживают того, чтобы их назвать: Александр М. Гримвэйд, Генри Смолл, Хенк Ф. Мосс, Гарриет Цуккерман. Их работы Форум публикует в 2018 г. и 2019 г. (Т. 43, № 4, с. 10-29; Т. 44, №1, с. 17-37).

Форум признателен своим читателям и делает все возможное, чтобы оставаться для них интересным. Так завершающий этот год выпуск журнала (Т. 45, № 4) содержит четыре новые статьи, а также авторский указатель. Представляемые статьи актуальны и разнообразны по своему содержанию. Первая статья касается вопроса научного сотрудничества ученых в области вычислительной техники с «гигантами» с точки зрения карьерного роста. Иными словами, в ней приводится анализ поэтапного развития карьеры молодых ученых области вычислительной техники в зависимости от сотрудничества с признанными и авторитетными учеными из той же области. И авторы предлагают провести аналогичный анализ на других областях науки. Вторая статья поднимает вопрос об использовании искусственного интеллекта в коммуникации между людьми и описывает острые проблемы этой сферы взаимодействия науки и технологии. Третья статья строит взаимосвязь между жизненными циклами исследования и данных и метаданными. А завершается выпуск статьей, отвечающей на вопрос, почему библиотекаря данных столь необходимы мягкие, а не только технические навыки, чтобы иметь успех в данной профессии.

Сегодня в условиях, когда вести отбор публикаций становится все сложнее, иногда что-то ускользает из нашего внимания, но Форум продолжает выходить, сохраняя традиции и отслеживая развитие своей области знания. Журнал делает все, чтобы, как и раньше, быть востребованным форумом!

Желаем нашему Форуму долгих лет, интересных номеров и откликов со стороны читателей!

Кобзева А. В., Овечкова Е. А.

Редакция МФИ

Анализ научного сотрудничества с «гигантами» на основе важных этапов карьеры*

И БУ
(Yi BU),

Дакота С. МЮРРЕЙ,
(Dakota S. MURRAY),

Ин ДИН
(Ying DING)

Университет шт. Индиана, США

Цзянь СЮЙ
(Jian XU)

Университет им. Сунь Ятсена, Китай

Пэн АЙ
(Peng AI)

Висконсинский университет в Мадисоне,
шт. Висконсин, США

Цзиньхуа ШЭНЬ
(Jinhua SHEN)

Университет Тунцзи, Китай

Фань ЯН
(Fan YANG)

Пекинский университет, Китай

Статья отражает исследование, касающееся изменения моделей сотрудничества на протяжении всей карьеры ученого, особенно сотрудничества с так называемыми «гигантами» — хорошо известными (авторитетными) представителями научной элиты с очень высоким влиянием. Карьера ученого делится на пять этапов, основываясь на числе лет, прошедших с момента получения докторской степени, для каждого этапа вычисляется ряд показателей, таких как подсчет публикаций, нормализованный подсчет ссылок, подсчет совместных работ, подсчет сотрудничества с «гигантами» и процентного соотношения непрерывного и продолжающегося присутствия. Анализ показывает, что совместная работа с «гигантами» может быть полезной для молодых ученых, они, весьма вероятно, смогут достичь большего научного успеха, если сначала будут работать независимо и только позже станут сотрудничать с «гигантами». Тем временем высоковлиятельные авторы стремятся работать с «гигантами» в начале своей карьеры, а не на поздних этапах, что заставляет нас пересмотреть укоренившиеся понятия относительно того, когда лучше всего сотрудничать с элитными учеными. Более того, обнаруживается, что высоковлиятельные авторы склонны иметь меньше соавторов на статью независимо от этапа своей карьеры, демонстрируя, что сотрудничество имеет важное значение, но и его размер также обладает значимостью.

ВВЕДЕНИЕ

Профессиональная карьера не имеет склонности быть единообразной — разные этапы карьеры могут внести изменения в задачи ученого, поведение и рабочие стратегии. Например, в начале карьеры подающий надежды молодой предприниматель готов учиться тому,

как представить недавно разработанные продукты на рынок или получить финансирование, тогда как старшие (опытные) бизнесмены, вероятнее всего, захотят идентифицировать торговые возможности вместо того, чтобы позволить себе погрузиться в разработку какого-либо продукта или маркетинговых тактик. Чтобы добиться успеха на каждом этапе профессиональной карьеры требуются *de facto* разные стратегии. Так и в науке успешные ученые применяют определенные стратегии на каждом этапе своей карьеры, такие, которые изменяют их модели сотрудничества [1].

* Перевод Bu Y., Murray D.S., Xu J., Ding Y., Ai P., Shen J., Yang F. Analyzing scientific collaboration with “giants” based on the milestones of career// Proceedings of 81st Annual Meeting ASIS&T. — 2018. — P. 29-38. — <https://www.asist.org/wp-content/uploads/Final-81st-Annual-Meeting-Proceedings-1.pdf>

Научное сотрудничество постепенно начинает преобладать во многих областях науки [2]. Сотрудничество даже может обеспечить некоторым ученым успех, что часто проявляется в более востребованных продуктах, а также порождать инновацию [4]. Однако, несмотря на то что успешные ученые могут быть мотивированы на широкое сотрудничество, характер сотрудничества может меняться в ходе их карьеры. Например, молодые ученые могут нести ответственность за написание статей и выполнять эмпирические исследования [5, 6], но постепенно, после нескольких лет, начать вести научные проекты, возможно, принимая в них роль указанного для переписки автора [7]. Альтернативно, в начале своей карьеры они могут пытаться искать сотрудничества с так называемыми «гигантами», хорошо известными и элитными авторами в их области, но в дальнейшем могут больше работать и со студентами.

«Стоять на плечах гигантов» – это часто повторяемая и глубоко укоренившаяся в культурном лексиконе науки фраза, поэтому ученые написали целые книги, проследившая ее значение и использование [9]. Часто, когда ученые употребляют это изречение, оно относится к кумулятивному характеру науки, принимая этих «гигантов» как образы знаменитых, давно ушедших ученых, таких как Исаак Ньютон и Альберт Эйнштейн; но другие подразумевают, что современные ученые тоже стоят на плечах пока живущих «гигантов». Автор работы [10] описал феномен, когда молодые ученые добиваются успешной карьеры с помощью объединения и получения пользы от экспертного знания своих наставников и профессиональных партнеров. Некоторые ученые с помощью усилия, удачи или того и другого получают доступ к знанию и наставничеству со стороны так называемых «гигантов» в рамках своей области исследования (это относится и к авторитетным, признанным авторам (authoritative authors), в данной статье – АА). Работа с АА обеспечивает молодым ученым уникальную возможность учиться у элиты, расширять свои профессиональные связи и достигать большего [8], позволяя лучше конкурировать с другими за ресурсы, признание и работу [11, 12, 13]. Как метафорически, так и практически стояние на плечах отдельных «гигантов» считается ценным, что может мотивировать молодых ученых на поиск и сотрудничество с такими гигантами.

В то время как предыдущее исследование изучало пользу работы с АА для молодого ученого, небольшое, более раннее исследование анализировало, могут ли ученые также сотрудничать с АА позже в своей карьере; ответ на этот вопрос может стать решающим в том, чтобы помочь нам определить модели успешного сотрудничества с точки зрения длительной перспективы карьеры. Статья рассматривает этот вопрос с помощью анализа (полного) цикла карьеры ученых-исследователей в области вычислительной науки, сотрудничающих с АА, и изучает взаимоотношения между этим сотрудничеством и его влиянием на каждом этапе карьеры.

ОТНОСЯЩИЕСЯ К ТЕМЕ ИССЛЕДОВАНИЯ

Различные этапы научной карьеры

Большинство исследователей начинают свою научную карьеру с присоединения к программе наставничества по подготовке к получению докторской степени [14, 15], где они опубликуют больше статей, чем те, кто в нее не вошел [16]. По сравнению с другими этапами

карьеры молодые ученые, с большей вероятностью, возьмут на себя ответственность за написание рукописей, проведение эмпирических исследований и воплощение идей [17] и с гораздо меньшей – руководство целой научной группой [18]. Между тем, во время своих исследований на докторскую степень молодые ученые стремятся получить больше публикаций в качестве первого автора [19]. Ряд исследователей заинтересованы в обеспечении обратной связи с молодыми учеными с целью помочь им достичь профессионального успеха и оценить карьерное развитие; например, Буддеберг-Фишер, Стамм, Буддеберг и Колагхофер (2008 г.) разработали одноразмерную шкалу карьерного успеха, нацеленную на молодых ученых в области физики, и предложили ряд подходов руководства или с целью получения высоких достижений. На этом раннем этапе их карьеры подразумевается, что ученые приобретут опыт и подготовятся к большим достижениям в дальнейшей карьере [20, 21].

Завершив курсовую работу, сдав кандидатские экзамены (например, по специальности) и защитив свои диссертации, младшие научные сотрудники готовы перейти из докторантов в независимые ученые [5]. После того, как они начнут новую карьеру, они должны иметь научных сотрудников или докторов наук на стажировке в их ВУЗе, работающих с ними [1]. Однако центр их внимания переходит от простой публикации статей к написанию применяемых для финансирования научных предложений и руководству молодыми учеными [22, 23]. На этой стадии карьеры ученые также могут начать испытывать стресс, связанный с продуктивностью, продвижением и пребыванием в должности, направленными на деятельность, которая напрямую выгодна их учреждениям [24]. Когда исследователи выдвигаются на следующую, более высокую позицию или продвигаются по должности, они часто служат в качестве интерпретаторов и посредников [25] и могут начать возглавлять научные группы [26]. Авторы работы [27] обнаружили, что старшие научные сотрудники имеют меньшую долю статей с правом первого автора, показывающую что они скорее могут быть руководителями исследовательских групп, нежели исполнителями работы. Однако старшие научные сотрудники стремятся потратить меньше времени на исследование [28], находя свое исследовательское время ограниченном растущим объемом административных обязанностей [29, 30] и преподавательской деятельности [31]. В этот период исследователи не стремятся искать руководителей (наставников) институциональным или официальным способом, как это обычно происходит в случае завершения их образования [11], но те, кто желает поддерживать взаимоотношения наставничества в целях карьерного роста все еще могут «искать руководителей из/в своих организаций» [32, с. 552], учитывая строгие обязательства, данные ими своей институциональной и общей культуре и ценностям [24]. Также обнаружено, что этот более поздний этап карьеры и стиль жизни имеет большую стабильность и меньший интерес к переопределению места [32]. Каждый этап карьеры ученого несет с собой различные центры внимания, поведения и активности.

Научное сотрудничество и карьера

Научное сотрудничество часто считается решающим для исследования в большинстве дисциплин и независимо от этапа карьеры [3]. Однако характер и частота

сотрудничества варьируется между различными этапами карьеры. Например, международный обзор 581 ученого в области экономики показал, что те, кто примкнул к их действующей организации меньше года назад, имеет наибольшую долю совместных статей, что не справедливо в отношении ученых, которые оставались со своим учреждением в течение десяти или более лет [1]. Это исследование также показало, что ученые в возрасте 56 лет и старше имеют значительно меньше совместных публикаций, чем ученые в возрасте 45 лет и моложе. Более того старшие ученые показывают большее разнообразие в числе соавторов на публикацию, чем более молодые ученые, с тенденцией иметь схожее число соавторов в статьях [33]. Авторы работы [34] разделили карьеры высококвалифицированных ученых на две части: до и после окончания учебного заведения и обнаружили различные модели сотрудничества на каждом этапе, соответствующие результатам предыдущего исследования (Крэнгт, 2015). Кроме того, на различных этапах карьеры подразделение задач в области сотрудничества также различно; характерно, как указано в работе [1], что нагрузка «в преобладающей степени лежит на плечах учеников» [р. 14] (или младших научных сотрудников), а не на руководителях-наставниках (или старших научных сотрудниках), которые вместо этого обеспечивают руководство своим научным сотрудникам, докторам, находящимся на стажировке в их ВУЗе, и докторантам.

Научное сотрудничество с «гигантами» (влиятельными авторами)

Мы часто замечаем, что ученые стараются встречаться и работать с «гигантами» определенной области, так как они могут обеспечить знание, опыт и ресурсы [10]. Авторы [1] обнаружили, что для младших научных сотрудников, повышающих свою продуктивность, наиболее важной причиной является сотрудничество с «гигантами»; другие причины этого: помощь молодым ученым в том, чтобы повысить вероятность привлекательности принятия статьи [35, 36]. Некоторые ученые демонстрируют выгоды сотрудничества с «гигантами». Например, авторы [37] освещают важную роль социальной сети для сотрудничества, демонстрируя, что ученые, вероятнее всего, обеспечат большее финансирование, если присоединятся к крупной исследовательской команде и объединятся с продуктивными исследователями, которые «имеют должный контроль над сетью сотрудничества и потоком информации» [37, р. 1]. Авторы [17] обнаружили, что в биомедицинских науках молодые ученые явно смогут найти новые идеи с помощью поощрения и руководства наставников из числа более опытных ученых.

Ньютон придерживался идеи о том, чтобы стоять на плечах гигантов; автор работы [10] использовал эту идею в качестве отправной точки аргументации, что ученые в области здравоохранения могут достичь успеха, работая с «гигантами» в своих областях. В работе [8] авторы расширили данную концепцию и изучили, как работа с «гигантами» влияет на успех молодых ученых. Они разделили авторов на две группы, тех, кто в какой-то момент работал с «гигантами» в своих областях, и тех, кто этого не делал. Затем подразделили первую группу на тех, кто работал с «гигантами» на начале своей карьеры, и тех, кто сотрудничал с «гигантами» на более позднем этапе карьеры. Результаты показали, что хотя сотрудничество с топовыми исследователями

имело важное позитивное влияние на возрастающую известность, работа в начале самостоятельно или с учеными среднего уровня, а затем позже сотрудничество с элитными авторами также может привести к успешному и высокому достижению. Основываясь на работе [8], данная статья ставит своей целью проанализировать влияние научного сотрудничества с «гигантами» в различных карьерных областях, используя разнообразные библиометрические показатели.

МЕТОДОЛОГИЯ

Данные и обработка

Массив данных является частью набора данных ArnetMiner [38], включающего публикации Ассоциации по вычислительному оборудованию (США) в области вычислительной техники. Массив ArnetMiner широко использовался и рассматривался в качестве одной из наилучшим образом курируемых баз данных статей по вычислительной технике [39]. Мы получили 2 092 356 статей, опубликованных в период с 1936 г. по 2014 г. и содержащих 1 207 061 уникального автора и 8 024 869 ссылок. Неоднозначность имен авторов была устранена в соответствии с алгоритмами, предложенными в работе [40]. Рис. 1 отражает процедуру обработки данных. Наши анализы проведены на ученых, чья первая публикация была издана в период 2000-2004 гг. Для каждой статьи в этом массиве данных подсчитывалось число ссылок, полученных каждой статьей из других статей в рамках БД ArnetMiner. H-индекс автора [41], широко использованный для оценки научных достижений ученого [42,43], получен в соответствии с числом публикаций автора, отраженных в массиве данных, и числом полученных ссылок.

Определение слова «гигант» является решающим в этом исследовании. Поскольку распределение h-индекса для всех авторов носит очень асимметричный характер, мы ссылаемся на работу [8] и определяем «гигантов» как тех, кто имеет h-индекс, равный 40 или больше, и называем их «авторитетными авторами» («authoritative authors» – AAs); в нашем массиве имеется 35 AAs (0, 002% из всех авторов). Чтобы изучить различие производительности между учеными, которые сотрудничали с AAs и которые не делали этого, мы подразделили авторов на две группы: Группа А включает тех авторов, которые, по крайней мере, однажды сотрудничали с АА за время своей карьеры (1 291 автор), и Группа В – содержит остальных. Чтобы изучить более дробные модели карьерных путей, мы потом разделили авторов Группы А на две подгруппы: группа А1, которая включает 613 авторов, публиковавших свои первые статьи с АА, и группа А2, включающая тех, кто в начале не публиковался с АА, но позже, по крайней мере, однажды сотрудничал. С помощью определения точного года получения докторской степени этими авторами мы можем приблизительно определить время начала их научной карьеры [14, 15]. Вручную получаем год получения образования на докторскую степень для всех авторов в группе А; мы смогли найти такую информацию только для 240 и 452 авторов в группах А1 и А2; называем эти две новые группы авторов с годами окончания учебы группами a1 и a2, которые составляют наш окончательный массив для анализа.

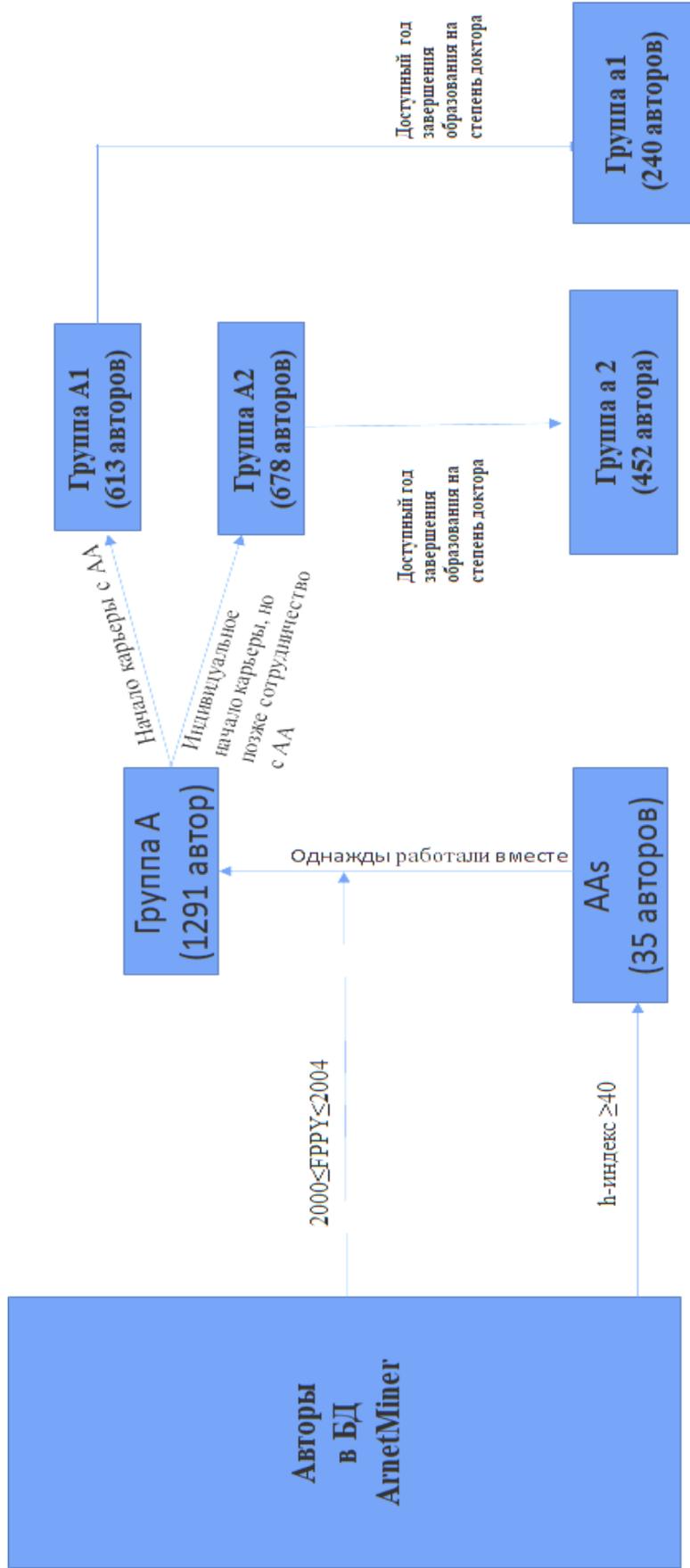


Рис. 1. Процедура обработки данных.
(Примечание: $\text{FRRU} = \text{first paper published year} - \text{год публикации первой статьи}$)

Категории влияния в Группах а1 и а2

Типы авторов		Авторы с высоким влиянием (h-индекс ≥ 10)	Авторы со средним влиянием ($10 > h\text{-индекс} \geq 5$)	Авторы с низким влиянием (h-индекс < 5)	Всего
Группа а1	Число (%)	50 (2,08%)	76 (31,7%)	114 (47,5%)	240 (100%)
Группа а2	Число (%)	158 (35,0%)	189 (41,8%)	105 (23,2%)	452 (100%)

Чтобы определить научную производительность авторов и лучше проанализировать развитие их карьеры, мы ссылаемся на работу [8] и делим всех авторов в Группах а1 и а2 на три типа на основе их h-индексов: авторы с высоким влиянием (h-индекс ≥ 10), авторы со средним влиянием ($5 \leq h\text{-индекс} < 10$) и авторы с низким влиянием (h-индекс < 5). Табл. 1 показывает, что доля исследователей с h-индексами 5 или выше в Группе а2 больше, чем в Группе а1, представляя, что исследователи, работающие самостоятельно в начале своей карьеры, а затем сотрудничающие с АА, обладают тенденцией иметь лучшие научные достижения, тем самым подерживая выводы работы [8].

Методы

В США ожидается, что многие студенты, специализирующиеся в вычислительной науке, получают свои степени в течение четырех - шести лет. Претенденты на докторскую степень стараются потратить первые годы своего обучения, непосредственно концентрируясь на курсовой работе, оставляя последние несколько лет до выпуска на то, чтобы посвятить их исследованию. После завершения учебы ученые, вероятно, должны работать в качестве постдокторанта, лица, получившего образование на докторскую степень, доцента или младшего научного сотрудника. Если они работают в университетах как члены профессорско-преподавательского состава, то продвижение по службе обычно происходит в период от четвертого до шестого года работы. Если они работают в индустрии или исследовательских учреждениях, то, вероятно, будут продвигаться до старшего научного сотрудника. Следовательно, мы определяем пять основных этапов карьеры ученых:

- Четыре года до окончания учебы на докторскую степень (обозначенные как годы -4-0): это, вероятно, докторанты;

- От одного до пяти лет после завершения образования на степень доктора (годы 1-5): это доктора наук, находящиеся на стажировке в ВУЗе (так называемые постдоки), доценты или младшие научные сотрудники [30, 44];

- От шести до десяти лет после завершения докторского образования (годы 6-10): это доценты/исследователи или продвинутые по должности адъюнкт-профессора/исследователи;

- От одиннадцати до пятнадцати лет после завершения образования на докторскую степень (годы 11-15): это адъюнкт-профессора/исследователи, профессора или старшие научные сотрудники [45];

- От шестнадцати до двадцати лет после завершения образования на докторскую степень (годы 16-20): это

старшие по возрасту и более опытные ученые, некоторые из которых, вероятно, находятся на стадии завершения карьеры [30].

Для каждого из этих периодов мы вычислим несколько библиометрических показателей. Число публикаций и ссылок являются традиционными показателями для измерения научной производительности ученых [46], поэтому они включены в данное исследование. Число публикаций – это просто число цитированных документов, опубликованных авторами за каждый период. Для каждого периода мы затем вычислим количество всех ссылок, сделанных другими публикациями в массиве данных для цитируемых документов, опубликованных данным автором за определенный временной период. Например, рассмотрим автора, который публикует две журнальные статьи, статья 1 в годы 1 - 5 и статья 2 в годы 6-10; автор не публикуется в годы 11-15. В годы один - пять статья 1 получила 10 ссылок. В годы шесть - десять статья 1 получила пять дополнительных ссылок, а статья 2 получила двадцать ссылок. В годы 11 - 15 обе статьи получают шесть ссылок. При вычислении общего числа (авторских) ссылок они (статьи) будут иметь 21 ($=10+5+6$) ссылок в годы один - пять, 26 ($=20+6$) ссылок в годы шесть - десять и ноль ссылок в годы одиннадцать - пятнадцать. После этого мы нормализуем подсчет ссылок с помощью использования результата числа публикаций и временного различия между временем опубликования и 2014 г., т.е. возрастом статьи в 2014 г., чтобы смягчить возможные предубеждения, вызванные авторской продуктивностью и возрастом статьи.

Другая цель нашего исследования – изучить научное сотрудничество с влиятельными авторами (АА - authoritative authors) и как это отражается на карьере ученого. Следовательно, мы также будем изучать взаимоотношения между этапом карьеры и числом определенных соавторов (показатель или соавторство), отдельных сотрудничающих авторитетных авторов (указывая влиятельных сотрудничающих лиц) и средний h-индекс сотрудничающих (указывая производительность сотрудничающих). Кроме того, нам интересно общее число сотрудничающих с ученым лиц. Мы используем эти показатели для анализа карьерной траектории каждого автора в Группах а1 и а2 на различных ступенях карьеры:

- число опубликованных статей;
- число полученных ссылок, нормализованных по возрасту каждой статьи;
- число отдельных сотрудничающих лиц;
- число отдельных сотрудничающих АА;
- средний h-индекс отдельных сотрудничающих лиц;
- среднее число сотрудничающих на опубликованную статью.

Для каждого из этих показателей мы проводим основанные на медиане тесты статистического значения по использованию необработанных значений из-за асимметричного распределения. Также мы применяем показатель непрерывного и продолжающегося присутствия (UCP- uninterrupted and continuous presence) для измерения постоянства (последовательности) достижений ученого. Здесь степень UCP определена как публикация по меньшей мере одной/трех/пяти статей в любые последовательные пять лет. Авторы [47] утверждают, что ученые, поддерживающие UCP в публикациях, могут достичь более высоких воздействий цитирования, чем те, кто этого не делает, и что UCP – это латентный показатель для измерения научной деятельности ученого и профессиональной общительности. Следовательно, чем выше показатель UCP ученого, тем больших достижений и стабильности он может достичь в своей научной карьере.

РЕЗУЛЬТАТЫ

Рис. 2-4 показывают результаты анализов карьеры по различным показателям Групп а1 и а2. Табл. 2 приводит статистическое значение результатов теста для всех показателей. Некоторые публикации исключены из этих подсчетов, так как они в некотором роде представляют собой «крайность» (экстрим), это, такие, которые имеют 15 или более соавторов и более 10 тыс. ссылок. Также мы исключаем определенных авторов, таких, как те, кто опубликовал более 200 публикаций в рамках любого 5-летнего временного интервала. Для некоторых групп нет значений за период с 11 до 15 лет, поскольку данные для более поздних периодов отсутствуют. Например, авторы, получившие свою докторскую степень после 2003 г. войдут в период (года 16-20) после 2014 г., который не включен в наш анализ.

С точки зрения числа статей, которое один автор имел на различных этапах карьеры (рис. 2 - левая часть и табл. 2) мы можем наблюдать возникновение различия между Группами а1 и а2 только после обозначенного 0 года, года завершения образования на докторскую степень. Это означает, что хотя авторы с одинаковым влиянием в Группе а1 и а2 публикуют одинаковое число статей до завершения образования на докторскую степень, а2 стремится быть более продуктивной после окончания обучения и таким образом более активной в научной деятельности [47]. Характерно, что пробел в числе публикаций между авторами высокого влияния в этих двух группах увеличивается по мере того, как авторы становятся старше. Между тем, в течение 15 лет с момента получения своих докторских степеней авторы с одинаковым влиянием имеют тенденцию публиковать меньше

статей в Группе а1 и а2. Однако в сравнении с а1 авторы в Группе а2 кажутся более усредненными, когда становятся старше. Следовательно, хотя Группа а1 была достаточно удачливой, чтобы работать с АА в начале своей карьеры, ее ученые не стремились быть такими постоянно продуктивными, как в Группе а2; наоборот, трудная работа авторов Группы а2 возможно привлекала внимание АА, которые тогда начали с ними сотрудничать.

Также стоит отметить, что авторы с высоким влиянием в Группе а2 публикуют больше статей в период (года 16-20) своей карьеры, чем в период (года 11-15). За период от 16 до 20 лет карьеры многие авторы смогли получить свои продвижения по занимаемой должности и, возможно, даже продвинулись до звания профессора. Хотя предыдущие исследования показывают, что продуктивность ученого склонна к снижению на поздних этапах карьеры [28], наш результат показывает, что авторы с высоким влиянием в Группе а2, те, кто начал свою карьеру независимо, и позже имеют шанс работать с «гигантами», постоянно публикуются и делают последовательные научные вклады. Независимо от группы, авторы с низким влиянием имеют самое малое количество публикаций, тогда как авторы с высоким влиянием – самое большое, вероятно исходя из факта, что h-индекс, наша мера импакта (влияния), зависит от числа публикаций [41].

Что касается нормализованного числа ссылок (рис. 2, правая часть и табл. 2), то публикации авторов Группы а2 имеют тенденцию к получению большего числа ссылок в результате удаления предвзятостей относительно продуктивности авторов и возраста статей. После получения своих докторских степеней авторы в Группе а2 стремятся получить несколько больше ссылок, чем авторы в Группе а1, показывая больший научный успех во время обучения и сразу после получения докторской степени, возможно в качестве предвестника успешной карьеры. В период более поздних этапов карьеры существует большее неравенство между авторами с высоким влиянием обеих групп, авторы с высоким влиянием в Группе а2 значительно превосходят авторов в Группе а1. Эти результаты согласуются с результатами авторов работы [8], которые обнаружили, что авторы Группы а2 имеют тенденцию к большему получению ссылок, чем авторы Группы а1, и такая разница возрастает по мере развития карьеры. Между тем, мы видим высокое число ссылок на этапе карьеры (годы 16-20) для авторов с высоким влиянием в Группе а2, согласующееся с показателем подсчета ссылок на рис. 2 слева. Это демонстрирует не только усердие этих авторов на протяжении всей карьеры, но также предполагает постоянное качество опубликованных статей на протяжении всей научной карьеры.

Таблица 2

Тест статистического значения для показателей в Группе а1 и а2 среди групп высокого, среднего и низкого импакта

Показатель	Высокий	Средний	Низкий	Показатель	Высокий	Средний	Низкий
Число публикаций	***	**	**	Число сотрудничающих с АА	**	*	***
Нормализованное число ссылок	**	*	**	Среднее число h-индекса сотрудничающих лиц	*	**	***
Число отдельных сотрудничающих лиц	***	***	**	Среднее число сотрудничающих на статью	***	***	*

(Примечание: * < 0,05; ** < 0,01; *** < 0,001)

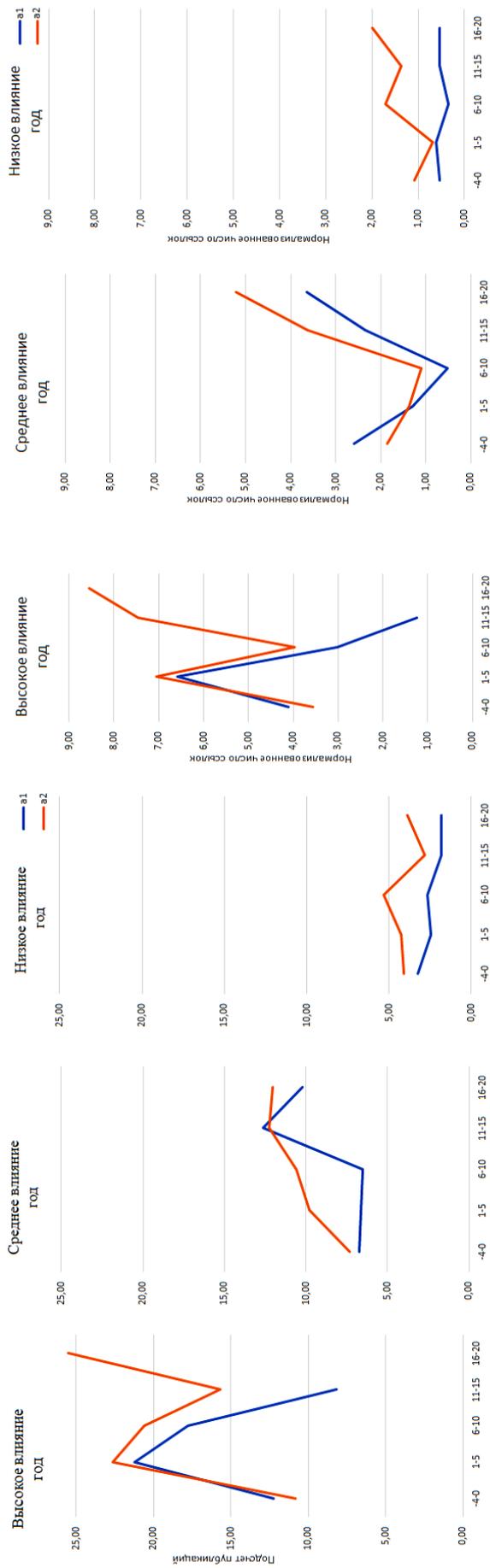


Рис. 2. Среднее число публикаций (слева) и среднее нормализованное число ссылок (справа) для Групп a1 и a2

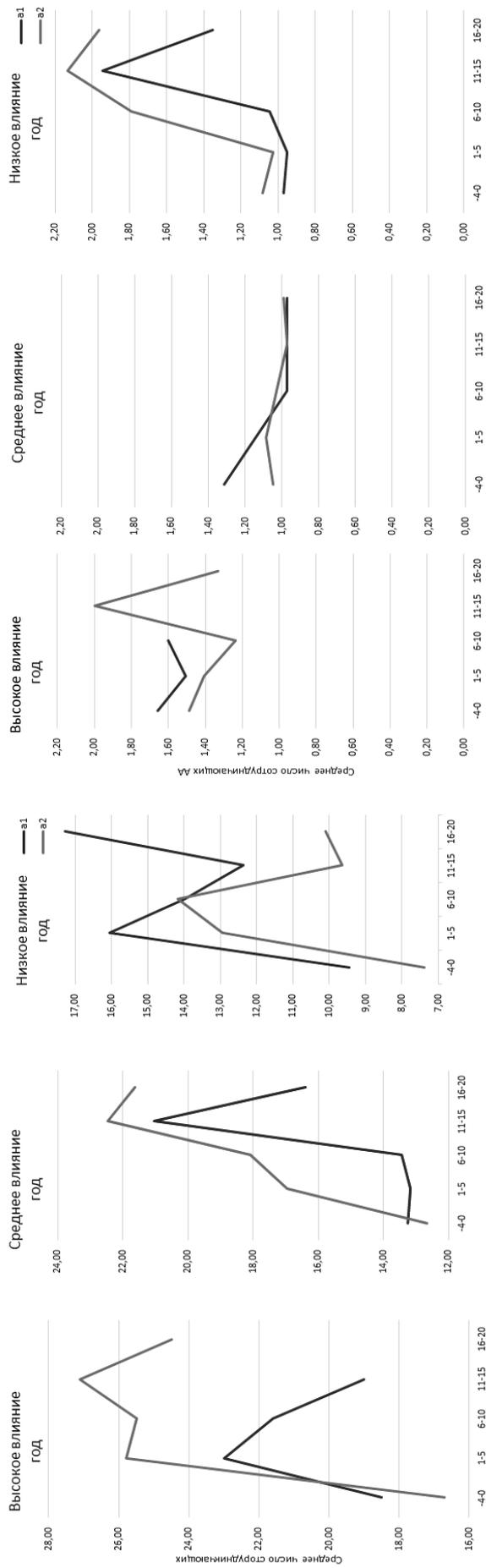


Рис. 3. Среднее число сотрудничающих лиц (слева) и среднее число сотрудничающих ЛА (справа) для Групп а1 и а2

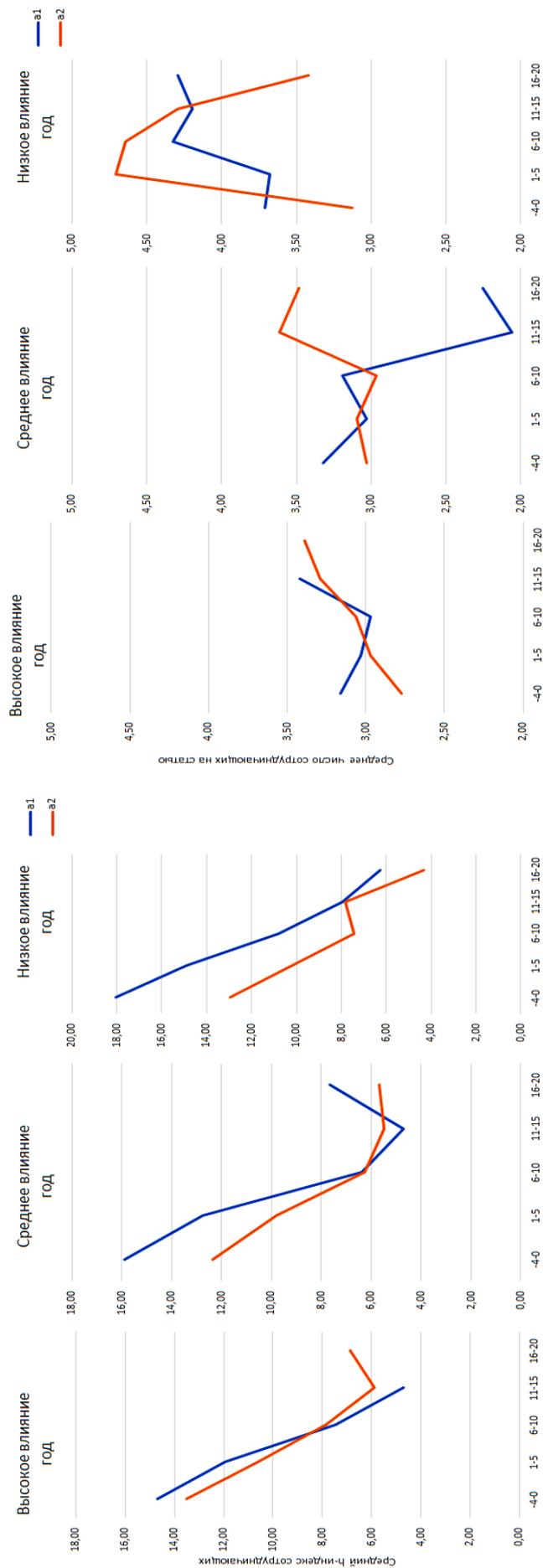


Рис. 4. Средний h-индекс (слева) и среднее число соавторов на статью (справа) для Группы a1 и a2

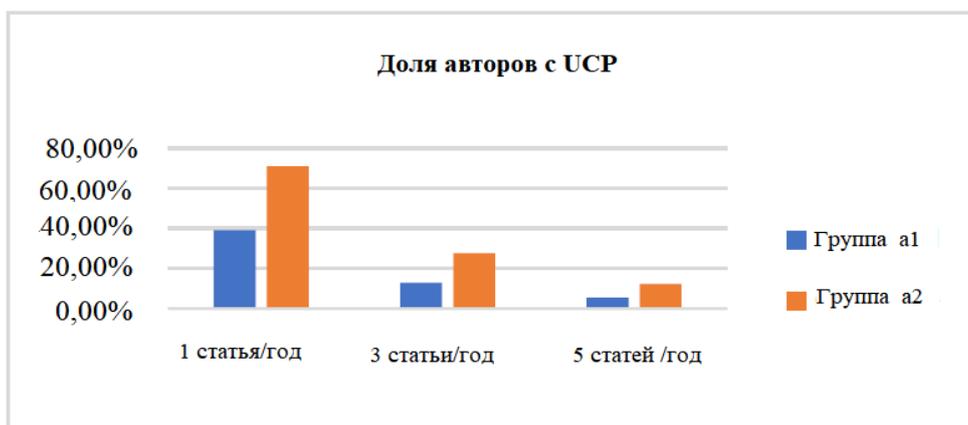


Рис. 5. Доля авторов с УСР в Группях а1 и а2

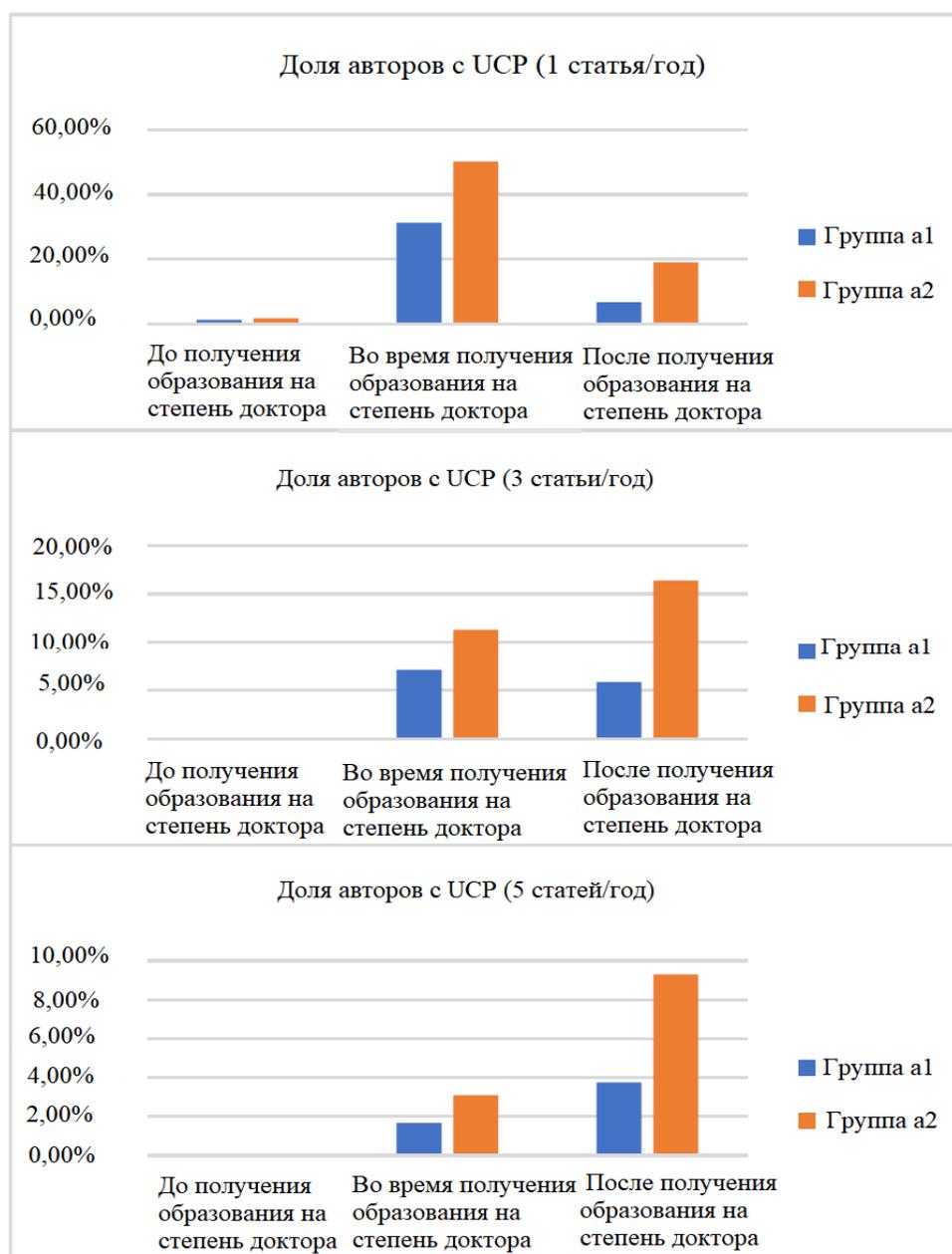


Рис. 6. Доля авторов с УСР в разные периоды с разными критериями: 1, 3 и 5 статей в год

Мы также можем изучить число сотрудничающих с автором лиц в каждой группе, как показано на рис. 3 слева (см. результаты тестов статистического значения в табл. 2, то же самое ниже), и увидим, что чем выше влияние автора, тем больше сотрудничающих лиц он имеет, что снова демонстрирует важность сотрудничества [3]. Авторы с высоким влиянием имеют вдвое больше среднего числа сотрудничающих лиц, чем авторы с низким влиянием.

Рис. 3 (справа) показывает, что авторы с высоким влиянием в этих двух группах имеют соответственно больше сотрудничающих с ними АА до завершения обучения на докторскую степень, чем после. Сотрудничество с АА дает молодым ученым шанс «стоять на плечах» гигантов [8, р. 307], гигантов, которые могут направлять их относительно того, как проводить исследование и как быть готовым к будущей работе. Однако совсем необязательно, что чем больше сотрудничества с АА, тем более успешны авторы, особенно в период от пяти до десяти лет после завершения образования на докторскую степень. Одна из интерпретаций, касающихся этого результата, состоит в том, что существует потенциал для «эффектов гало» авторитетных авторов «затмевать» молодых ученых и до некоторой степени сковывать их развитие [34]. Это открытие предупреждает нас о том, что сотрудничество с «гигантами», как мы можем думать, не всегда выгодно, особенно в начале карьеры. Рис. 4 (слева) отображает средний h-индекс сотрудничающих авторов Групп а1 и а2, чтобы изучить влияние сотрудничающих с ними лиц. Результат показывает, что средний h-индекс сотрудничающих снижается по мере развития карьеры, независимо от того, в какой группе находится автор. Этот результат подтверждает наши интуитивные предположения: в начале карьеры ученый, вероятно, является студентом, а основными сотрудничающими с ним лицами – куратор, завкафедрой или руководитель (наставник); студент будет стремиться иметь более высокий h-индекс, чем они; со временем они могут стать авторитетными исследователями и тогда сотрудничающие с ними лица будут включать больше студентов или младших научных сотрудников, снижая средний h-индекс их соавторов. Среднее число соавторов на статью, показанное на рис. 4 (справа), варьируется от двух до пяти. Крапф (2015 г.) указывает, что навыки и исследовательская способность ученых могут развиваться и повышаться посредством выполнения различных видов задач, таких как разработка исследования, проведение эмпирических исследований, обзор литературы, написание рукописей и анализ данных. Упражнение и опыт с такими отдельными задачами поможет им в достижении будущего успеха. Однако авторы с низким влиянием, независимо от того, к какой группе они относятся, готовы иметь больше соавторов на статью, предлагая предельную точку, когда слишком большое число соавторов перестает быть выгодным.

Рис. 5 показывает результаты УСР, основанные на различных критериях (одна, три и пять статей в год), в которых можно увидеть, что доля авторов с УСР в Группе а2 вдвое больше, чем в Группе а1. Характерно, что около 70%, 28% и 12% авторов Группы а2 могут поддерживать публикационный поток с УСР для одной, трех и пяти статей в год, в то время как только 39%, 12% и 5% авторов Группы а1 могут это делать. Такое высокое значение для УСР показывает, что Группа а2 имеет более последовательную (постоянную) научную деятельность на протяжении всей карьеры своих ученых,

подтверждая результаты работы [8]. Всех авторов, отвечающих критериям УСР в этих двух группах, мы относим к одной из трех категорий, основанных на том, получали ли они свой статус УСР до, во время или после завершения образования на докторскую степень. *До завершения образования на докторскую степень* – определяется как периоды до получения докторской степени, включая любые публикации, сделанные до того, как они получили свою докторскую степень. *После завершения образования на докторскую степень* – включает все статьи, опубликованные авторами после завершения образования на докторскую степень. Период *во время получения образования на докторскую степень* частично перекрывается с периодом до/после завершения образования на степень доктора. Можно сказать, что период *во время обучения на степень доктора* должен охватывать и год завершения образования. Например, если автор начинает обучение на докторскую степень в 2005 г. и заканчивает в 2010 г., тогда период *до завершения образования на докторскую степень* будет определен как любое время до 2010 г., *во время получения образования на степень доктора* будет определяться как любое время после 2010 г. Рис. 6 демонстрирует низкую долю авторов с УСР в период *до завершения образования на степень доктора* в обеих группах. В период *во время обучения на степень доктора* доля авторов с УСР в Группе а2 больше, чем в Группе а1, а после окончания обучения на докторскую степень пробел удваивается или даже утраивается. Это большое расхождение обозначает резкие изменения карьерных траекторий между группами. Считается, что исследователи без отметки о УСР, наиболее вероятно оставят академическую сферу и исследование [48]; следовательно, Группа а2 проявляет тенденцию быть успешной благодаря стабильной научной карьере.

ОБСУЖДЕНИЕ И ВЫВОДЫ

Роберт Мертон подробно описал «эффект гало» [48], который подразумевает, что престиж и уважение успешных ученых может стать помехой для близких партнеров, но наши результаты рисуют более сложный образ. Данная статья исследует измерение времени, касающееся лиц, сотрудничающих с авторитетными авторами (АА) или «гигантами», и изучает, как работа с гигантами на различных этапах научной карьеры может привести к определенным последствиям. Мы делим карьеру ученых на различные этапы и вычисляем ряд показателей для каждого этапа (например, число публикаций, нормализованное число ссылок, число сотрудничающих лиц, доля УСР и т.д.). Наши результаты показывают, что хотя работа с «гигантами» может иметь преимущества и призыв к честолюбивым молодым ученым, которые отдают предпочтение карьерным достижениям и продуктивности; едва ли это требование долгосрочного успеха. Взамен – трудная, требующая напряжения работа с целью стать признанным ученым и только позже привлечь к себе внимание «гигантов», чтобы работать вместе, учиться у них склонности к большему профессиональному успеху. Значимым определяющим фактором карьерного успеха, как кажется, является упорная, сложная работа и постоянный вклад в науку, о чем свидетельствуют и наша работа, и работа наших предшественников [49].

Эффект Матфея [48] подразумевает, что если молодой ученый мечтает об успехе и престиже, то он должен извлекать выгоду от создания кумулятивного преимуще-

ства на ранней стадии своей карьеры путем нахождения работы и с элитным, и с престижным «гигантом» в своей области. Наши результаты обеспечивают свидетельство, которое показывает, что это не совсем обязательный случай. Несмотря на то, что престиж имеет свои преимущества, внимательное руководство может быть более важным для молодого, только начинающего свою научную карьеру, ученого. У «гиганта» может быть мало времени для тщательного и внимательного руководства отдельными лицами, он помимо этого, вероятно, имеет множество профессиональных обязанностей, таких как чтение лекций, обслуживание сообществ профессиональных организаций, руководство кандидатами, стремление к преподаванию и другие служебные нагрузки [29, 31]. В таких обстоятельствах молодой ученый может почерпнуть пользу от более молодого наставника, который будет уделять больше времени индивидуальному руководству. Нам необходимо пересмотреть заново традиционную мудрость академической сферы и осознать, что существует правильное (подходящее) время для приобретения престижа и подходящее время для установления руководства. Работая с АА в подходящее время, когда еще молодой, но со сложившейся карьерой, ученый может распределить их (АА) для получения наибольшей выгоды от престижа и опыта такого сотрудничества. Выбор определенного времени «убедительно продвигает от инерции к динамичным и взрывным научным высотам» [13, р. 17]. Директора программы наставничества и организаторы конференций могут использовать эти результаты в целях планирования конференций, семинаров и других мероприятий, позволяющих начинающим, младшим и старшим научным сотрудникам коммуницировать и взаимодействовать таким способом, который приносит наибольшую выгоду. Что касается руководства и сотрудничества, то здесь не существует универсального рецепта; стратегии и программы скорее должны отражать тот факт, который необходим для получения лучших результатов – каждый должен делать правильную вещь в правильное время.

Данная статья также изучает меняющиеся стили сотрудничества между уровнями влияния и на протяжении всей карьеры. Мы обнаружили, что авторы с высоким влиянием склонны иметь небольшое число соавторов на статью по сравнению с авторами со средним и низким влиянием, независимо от этапа карьеры; поскольку сотрудничество является важным, то этот результат также показывает, что общее число совместных действий тоже имеет эффект. Слишком большое число участников в научном проекте может снижать эффективность команды, что отразится в появлении более слабой коммуникации. Увеличение числа соавторов в одной публикации также намекает на снижение соответствующих рабочих нагрузок для каждого автора [50]. Как отмечалось в работе [51], слишком большой состав команды может ограничивать производительность, что потом отразится на качестве статьи. Проведение исследования в больших командах и по многим проектам может привести членов команды к меньшей занятости в каждом отдельном проекте и потере эффективности из-за диспропорциональных временных затрат на организацию команды, а не на улучшение самого исследования [52]. Этот результат вносит дополнение в литературу по сотрудничеству и предлагает интуитивную точку зрения, что больше – это не всегда лучше, т.е. истина, которая, кажется, проходит через всю научную карьеру.

Настоящее исследование является одним из первых по изучению идеи «делать правильную вещь в правильное время» в академической сфере и, будучи пробным, страдает от ряда ограничений – большинство из которых вытекает из его объема. Определение сотрудничающих с АА лиц наряду с их основными этапами карьеры, например, такой этап как получение образования на степень доктора наук требует длительной работы и ограничивает размер выборки ученых всего несколькими сотнями. На такой небольшой выборке мы не можем проконтролировать многие, стоящие в тупик факторы в нашей работе. Например, ученые, определенные в Группе a1, те, кто начинает свою карьеру, публикуясь мало или не публикуясь вовсе, таким образом снижая средние групповые показатели. Наоборот, учеными Группы a2, т.е. теми, которые сотрудничают с гигантами в своей карьере позже, вероятно, будут те, кто остается в академической сфере и в достаточной степени занимается исследованиями, чтобы привлечь внимание этих гигантов и, следовательно, почти априори, по всей вероятности будут более сильными и более запоминающимися учеными. Различия в поддисциплине также могут играть свою роль; мы были вынуждены найти обходной путь, с помощью которого определить «гигантов», однако разные подобласти очень отличаются по числу ссылок и публикаций, которые они могут получить [52]. Порог h-индекса (40) может ограничить число «гигантов» в нашей выборке; в дальнейшем мы планируем сослаться на работу [53] и определять «гигантов» с помощью использования Top 1 (первого места в ранжировании авторов с двухпроцентильным h-индексом) и/или порог h-индекса (40) может ограничить источники «гигантов» небольшим числом высокопродуктивных или с высоким цитированием подобластей, которые могут иметь различные модели сотрудничества, руководства и карьерных траекторий после получения образования на докторскую степень; эти модели, вероятно, должны расходиться даже больше между дисциплинами, как в науках о Жизни, или в гуманитарных науках. Более того, поскольку мы стремимся анализировать сотрудничество с «гигантами» в различных видах карьеры, то данная статья не определяет основную линию тех авторов, которые никогда не имели никакой возможности работать с «гигантами». Также подсчет ссылок и h-индексов производится на базе записей в массиве данных, что может вызвать некоторые предубеждения. Эти ограничения ведут нас к тому, чтобы делать выводы с осторожностью; тогда как наши результаты предлагают свидетельство относительно различных направлений и нюансов в сотрудничестве в ходе карьеры, то необходимо провести больше работы для представления основательного понимания данной области.

Отсутствие возможности принять во внимание других соавторов является одним из ограничений настоящей статьи. Известно, что статьи пишутся на основе определенного исследования и АА тоже соавторствуют вместе с другими. Мы попытаемся рассмотреть эту проблему в ближайшем будущем. С другой стороны, игнорирование асимметрии между двумя соавторами также несовершенно. Потенциальный подход для борьбы с этим состоит в том, чтобы рассмотреть предпочитаемых определенным автором партнеров; это исследование Указателя родства (Affinity Index) и Указателя вероятностного родства (Probabilistic Affinity Index) [54]. Между тем, мы также надеемся работать над расширением и

разнообразием нашей выборки и контролем относительно смущающих факторов. Также мы надеемся увеличить круг дисциплин, чтобы понять, схожи ли модели сотрудничества, найденные нами в карьерах вычислительных ученых, с моделями в других областях. Кроме того, мы будем дополнять набор показателей, которые в настоящее время полезны для создания приблизительного понимания относительно достижения карьеры, но не охватывают нюансы научного влияния. Большое количество показателей также позволит нам изучить другие аспекты сотрудничества и этапы карьеры, такие как изменения в разнообразии влияний, темы исследований, разнообразие соавторов, степень постоянства [55] и стабильность [56], касающаяся научного сотрудничества. Эта работа закладывает фундамент дальнейшего изучения того, как сотрудничество и другие аспекты карьеры ученого могут влиять на будущий успех и определять не только то, как спланировать успешную карьеру, но и то, когда устанавливать приоритеты для разных задач и поведений.

Благодарность. Авторы выражают признательность трем анонимным рецензентам за улучшение качества данной статьи.

ЛИТЕРАТУРА

1. Kumar S., Ratnavelu K. Perceptions of scholars in the field of Economics on co-authorship associations: Evidence from an international survey// PLoS One. — Vol. 11, No. 6, e0157633.
2. Zhang C., Bu Y., Ding Y., Xu J. Understanding scientific collaboration: Homophily, transitivity, and preferential attachment// Journal of the Association for Information Science and Technology. — 2018. — Vol. 69, No. 1. — P. 72-86.
3. Wuchty S., Jones B. F., Uzzi B. The increasing dominance of teams in production of knowledge// Science. — 2007. — Vol. 316, No. 5827. — P. 1036-1039.
4. Camarinha-Matos L.M., Afsarmanesh H. Collaborative networks: A new scientific discipline// Journal of Intelligent Manufacturing. — 2005. — Vol. 16, No. (4-5). — P. 439-452.
5. Baker V.L., Pifer M.J. The role of relationships in the transition from doctoral student to independent scholar// Studies in Continuing Education. — 2011. — Vol. 33, No. 1. — P. 5-17.
6. Holgate S.A. How to collaborate//Science (Careers). — 2012. — DOI: 10.1126/science.caredit.a1200082.
7. Du J., Tang X. Perceptions of author order versus contribution among researchers with different professional ranks and the potential of harmonic counts for encouraging ethical co-authorship practices// Scientometrics. — 2013. — Vol. 96, No. 1. — P. 277-295.
8. Amjad T., Ding Y., Xu J., Zhang C., Daud A., Tang J., Song M. Standing on the shoulders of giants// Journal of Informetrics. — 2017. — Vol. 11, No. 1. — P. 307-323.
9. Merton R.K. On the shoulders of giants. — Chicago: University of Chicago Press, 1993.
10. Adegbola M. Scholarly tailgating defined: A diverse, giant network// The ABNF Journal: Official Journal of the Association of Black Nursing Faculty in Higher Education, Inc. — 2013. — Vol. 24, No. 1. — P. 17-20.
11. Kram K.E., Isabella L.A. Mentoring alternatives: The role of peer relationships in career development//Academy of Management Journal. — 1985. — Vol. 28, No. 1. — P. 110-132.
12. Quatman C., Chelladurai P. Social network theory and analysis: A complementary lens for inquiry//Journal of Sport Management. — 2008. — Vol. 22, No. 3. — P. 338-360.
13. Adegbola M. Nurses collaborating with cross disciplinary networks: Starting to integrate genomics into practice// Journal of National Black Nurses Association. — 2010. — Vol. 21, No. 1. — P. 46-49.
14. Subramanyam K. Bibliometric studies of research collaboration: A review// Journal of Information Science. — 1983. — Vol. 6, No. 1. — P. 33-38.
15. Hart R.L. Co-authorship in the academic library literature: A survey of attitudes and behaviors// Journal of Academic Librarianship. — 2000. — Vol. 26, No. 5. — P. 339-345.
16. Muschallik J., Pull K. Mentoring in higher education: Does it enhance mentees' research productivity? // Education Economics. — 2016. — Vol. 24, No. 2. — P. 210-223.
17. Packalen M., Bhattacharya J. Age and the trying out of new ideas. — National Bureau of Economic Research, 2015. — DOI: 10.3386/w20920.
18. DeCastro R., Sambuco D., Ubel P.A., Stewart A., Jaqsi R. Mentor networks in academic medicine: Moving beyond a dyadic conception of mentoring for junior faculty researchers// Journal of the Association of American Medical Colleges. — 2013. — Vol. 88, No. 4. — P. 488-496.
19. Costas R., Bordons M. Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective// Scientometrics. — 2011. — Vol. 88, No. 1. — P. 145-161.
20. Barbuto J.E. Jr., Story J.S., Fritz S.M., Schinstock J.L. Full range advising: Transforming the advisor-advisee experience// Journal of College Student Development. — 2011. — Vol. 52, No. 6. — P. 656-670.
21. Schlosser L.Z., Gelso C.J. Measuring the working alliance in advisor-advisee relationships in graduate school// Journal of Counseling Psychology. — 2001. — Vol. 48, No. 2. — P. 157-167.
22. Gulbradsen M., Smeby J.-C. Industry funding and university professors' research performance// Research Policy. — 2005. — Vol. 34, No. 6. — P. 932-950.
23. de Solla Price J., Cotten S.R. Teaching, research, and service: Expectations of assistant professors// The American Sociologist. — 2006. — Vol. 37, No. 1. — P. 5-21.
24. Raelin J. The clash of cultures. — Boston: Harvard Business School Press, 1985.
25. Jaeger A.J., Sandmann L.R., Kim J. Advising graduate students doing community-engaged dissertation research: The advisor-advisee relationship// Journal of Higher Education Outreach and Engagement. — 2011. — Vol. 15, No. 4. — P. 5-25.
26. Rowlands I., Nicholas D. The changing scholarly communication landscape: An international survey of senior researchers// Learned Publishing. — 2006. — Vol. 19, No. 1. — P. 31-55.
27. Gingras Y., Larivière V., Macaluso B., Robitaille J.-P. The effects of aging on researchers' publication and citation patterns // PLoS One. — 2008. — Vol. 3, No. 12, e0004048.
28. Kanda Y., Tomizawa H. Changes in the ratio of time spent on work activities by university and college faculty members: A comparison of results of the "Survey of full-time equivalency data at universities and colleges" of 2002, 2008 and 2013// NISTEP Re-search Material 236 (in Japanese). — 2015. — <http://data.nistep.go.jp/dspace/handle/11035/3027>.

29. *Barham B.L., Foltz J.D., Prager D.L.* Making time for science// *Research Policy*. — 2014. — Vol. 43, No. 1. — P. 21-31.
30. *Kawaguchi D., Kondo A., Saito K.* Researchers' career transitions over the life cycle// *Scientometrics*. — 2016. — Vol. 109, No. 3. — P. 1435-1454.
31. *Taylor W., Fender B., Burke K.* Unraveling the academic productivity of economists: The opportunity costs of teaching and service// *Southern Economic Journal*. — 2006. — Vol. 72, No. 4. — P. 846-859.
32. *van Eck Peluchette J., Jeanquart S.* Professionals' use of different mentor sources at various career stages: Implications for career success// *The Journal of Social Psychology*. — 2000. — Vol. 140, No. 5. — P. 549-564.
33. *Hamermesh D.S.* Age, cohort and co-authorship. — Cambridge, MA: National Bureau of Economic Research, 2015.
34. *Bu Y., Ding Y., Xu J., Liang X., Gao G., Zhao Y.* Understanding success through the diversity of collaborators and the milestone of career// *Journal of the Association for Information Science and Technology*. — 2018. — Vol. 69, No. 1. — P. 87-97.
35. *Bhopal R., Rankin, J., ..., Rodgers H.* (1997). The vexed question of authorship: Views of researchers in a British medical faculty // *British Medical Journal*. — 1997. — Vol. 314, No. 7086. — P. 1009-1012.
36. *Bennett D.M., Taylor D.M.* Unethical practices in authorship of scientific papers// *Emergency Medicine Australasia*. — 2003. — Vol. 15, No. 3. — P. 263-270.
37. *Ebadi A., Schiffauerova A.* How to receive more funding for your research? Get connected to the right people! — // *PLoS One*. — 2015. — Vol. 10, No. 7, e0133061.
38. *Tang J., Zhang J., Ya L., Li J., Zhang L., Su Z.* Arnet-Miner: Extraction and mining of academic social networks// *Proceedings of the fourteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.990-998, August 24-27, 2008, Las Vegas, NV., U.S.A.— 2008.
39. *Dong Y., Johnson R.A., Chawla N.V.* Can scientific impact be predicted? *IEEE Transactions on Big Data*. — 2016. — Vol. 2, No. 1. — P. 18–30.
40. *Tang J., Fong A.C.M., Wang B., Zhang J.* A unified probabilistic framework for name disambiguation in digital library// *IEEE Transaction on Knowledge and Data Engineering*. — 2012. — Vol. 24, No. 6. — P. 975-987.
41. *Hirsch J.E.* An index to quantify an individual's scientific research output// *Proceedings of the National Academy of Sciences of the United States of America*. — 2005. — Vol. 102, No. 46. — P. 16569-16572.
42. *Lacasse J.R., Hodge D.R., Bean K.F.* Evaluating the productivity of social work scholars using the h-index// *Research on Social Work Practice*. — 2011. — Vol. 21, No. 5. — P. 599-607.
43. *Ciriminna R., Pagliaro M.* On the use of the h-index in evaluating chemical research// *Chemistry Central Journal*. — 2013. — Vol. 7, No. 1. — P. 132.
44. *University of California Berkeley Career Center.* The transition from graduate student to assistant professor. — 2017. — <https://career.berkeley.edu/PhDs/PhDtransition>.
45. *The Princeton Overview.* A day in the life of a professor. — 2017. — <https://www.princetonreview.com/careers/127/professor>.
46. *de Solla Price J.* A general theory of bibliometric and other cumulative advantage processes// *Journal of the American Society for Information Science*. — 1976. — Vol. 27, No. 5. — P. 292-306.
47. *Ioannidis J.P.A., Boyack K.W., Klavans R.* Estimates of the continuously publishing core in the scientific workforce// *PLoS One*. — 2014. — Vol. 9, No. 7, e101698.
48. *Merton R.K.* The Matthew Effect in science// *Science*. — 1968. — Vol. 159, No. 3810. — P. 56-63.
49. *Schneider M.E., Doye B.* The everything guide to writing a book proposal: Insider advice on how to get your work published. — San Diego, California: Adams Media Corporation, 2005.
50. *Curral L.A., Forrester R.H., Dawson J.F., West M.A.* It's what you do and the way that you do it: Team task, team size, and innovation-related group processes// *European Journal of Work and Organizational Psychology*. — 2001. — Vol. 10, No. 2. — P. 187-204.
51. *Vrabie A.* Productivity and team size: Less is more. — 2013. — <http://blog.sandglaz.com/productivity-and-team-size-less-is-more/>
52. *Qian Y., Rong W., Jiang N., Tang J., Xiong Z.* Citation regression analysis of computer science publications in different ranking categories and subfields// *Scientometrics*. — 2017. — Vol. 110, No. 3. — P. 1351-1374.
53. *Hobberger J.* Does it pay to stand on the shoulders of giants? An analysis of the inventions of star inventors in the biotechnology sector// *Research Policy*. — 2016. — Vol. 45, No. 3. — P. 682-698.
54. *Chinchilla-Rodríguez Z., Bu Y., Robinson-García N., Costas R., Sugimoto C. R.* Travel bans and scientific mobility: Utility of asymmetry and affinity indexes to inform science policy// *Scientometrics*. — 2018.
55. *Bu Y., Ding Y., Liang X., Murray D.S.* Understanding persistent scientific collaboration// *Journal of the Association for Information Science and Technology*. — 2018. — Vol. 69, No. 3. — P. 438-448.
56. *Bu Y., Murray D.S., Ding Y., Huang Y., Zhao Y.* Measuring the stability of scientific collaboration// *Scientometrics*. — 2018. — Vol. 114, No. 2. — P. 463-479.
57. *Koch C.G.* The science of success: How market-based management built the world's largest private company. — Hoboken, NJ: Wiley, 2007.

Опосредованная искусственным интеллектом коммуникация: определение, повестка исследований и этические размышления*

Джеффри Т. ХЕНКОК
(Jeffrey T. HANCOCK)

Отделение коммуникации,
Станфордский университет, г. Станфорд,
шт. Калифорния, США

Мор НААМАН
(Mor NAAMAN)

Отделение информатики, Корнеллский
университет, г. Итака, шт Нью-Йорк, США

Карен ЛИВИ
(Karen LEVY)

Корнеллский технологический факультет,
Корнеллский университет, г. Нью-Йорк,
шт. Нью-Йорк, США

Опосредованная искусственным интеллектом коммуникация (Artificial Intelligence – Mediated Communication, AI-МС) определяется как межличностная коммуникация, в которой интеллектуальный агент действует от лица коммуниканта путем изменения, расширения или создания сообщений с тем, чтобы достичь коммуникационных целей. Недавнее появление AI-МС ставит новые вопросы относительно того, как технология может преобразовать коммуникацию, и требует переоценки и даже расширения многих базовых теорий, подходов и наблюдений о компьютерно-опосредованной коммуникации (computer-mediated communication, СМС). Повестка исследований относительно AI-МС должна рассматривать разработку этих технологий, а также психологические, лингвистические, относительные, политические и этические аспекты введения искусственного интеллекта в коммуникации между людьми. Статья нацелена на четкую формулировку подобной повестки.

Появление компьютерно-опосредованной коммуникации (computer-mediated communication, СМС) революционизировало межличностную коммуникацию, предоставляя индивидуумам множество форматов и каналов для отправки сообщений и взаимодействия с другими лицами во времени и пространстве [1]. В понимании классической социологии СМС (например, [2]) агент (посредник) и его качества играют важную роль в моделировании того, как субъекты используют технологии для достижения межличностных целей. Действие остается за коммуникантом: создание сообщения и управление впечатлением понимаются широко, чтобы

удовлетворять целям коммуниканта. Точно так же, получатель сообщения, как предполагается, должен понять и принять это действие.

Введение искусственного интеллекта в межличностную коммуникацию имеет потенциал снова трансформировать способ коммуникации людей, перевернуть предположения о действии и посредничестве, и поставить новые этические вопросы. СМС сегодня расширяется путем включения опосредованной искусственным интеллектом коммуникации (Artificial Intelligence – Mediated Communication, AI-МС): межличностная коммуникация, не просто передаваемая технологией, а изменяющаяся, расширяющаяся или даже создаваемая под действием интеллектуального агента с тем, чтобы достичь коммуникационных целей.

* Перевод Hancock J. T., Naaman M., Levy K. AI-Mediated communication: Definition, research agenda, and ethical considerations // Journal of Computer-mediated communication. — 2020. — P. 1-12. — <https://sml.stanford.edu/ml/2020/01/hancock-jcmc-ai-mediated-communication.pdf>

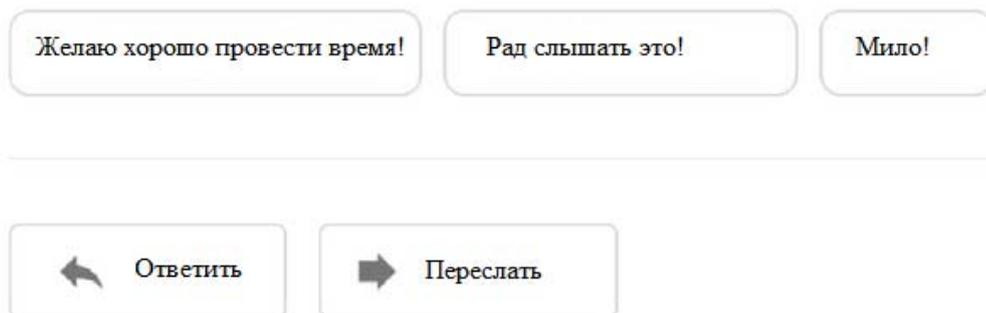


Рис. Предлагаемые ответы, созданные AI почтового сервера Gmail, часто считающиеся положительными.

КОНЦЕПТУАЛИЗАЦИЯ AI-МС

Наше определение AI-МС базируется на описании AI Расселом и Норвигом как самообучающегося «интеллектуального агента», действующего с учетом исходных данных (ощущений) для достижения наилучшего ожидаемого результата [3, с. 4]. Это определение оценивает AI в рамках поведения агента и не касается его обоснований агентом. Придерживаясь данного определения, будем применять AI довольно широко и относить к нему вычислительные системы, включая алгоритмы, методы машинного обучения, обработку естественного языка, а также другие средства, действующие от лица индивидуума в целях улучшения результата коммуникации. Интеллектуальный агент может анализировать исходные данные, включая сообщения, созданные человеком, историю коммуникации, личную информацию или любой другой источник данных. Затем агент может предложить, расширить, изменить или создать сообщения, чтобы достичь ожидаемого результата.

Исследование СМС может быть определено как анализ социальных эффектов коммуникации, происходящей между людьми с использованием сетевых устройств цифровой связи для обмена сообщениями (например, обработка сообщений электронной почты и текста, взаимодействия внутри сайта социальной сети, видео-конференц-связь) [4]. Будем придерживаться ориентира авторов [2, с. 530] на социальные научные подходы к динамике межличностной коммуникации между людьми через технологии.

Мы объединим эти концептуализации AI и СМС, чтобы определить AI-МС следующим образом: *опосредованная коммуникация между людьми, в которой интеллектуальный агент действует от лица коммуниканта, изменяя, расширяя или создавая сообщения, чтобы завершить коммуникацию или достичь межличностных целей.*

Мы уже начали рассматривать примеры AI из многих типов СМС. В вербальных каналах использование AI улучшает коммуникацию на основе текста – от автокоррекции, предсказания текста и исправления грамматики [5] до интеллектуальных ответов, автозаполнения и автоответов (как в Gmail и мобильных телефонах). Например, в интеллектуальных ответах Gmail получатель сообщения может выбрать один из нескольких созданных AI вариантов (рис.). Эта тенденция одинаково, если не больше, усиливается для невербальной СМС, такой как автоматическая вставка эмоций. Обычно вовлечение

AI не раскрывается, партнер, вероятно, думает, что сообщение создано отправителем.

Будущие системы AI-МС могут пойти даже дальше, оптимизируя сообщения для межличностных последствий, подобных сообщению высокого статуса [6] или попытке казаться заслуживающим доверия [7]. Более того, появляющиеся технологии могут оптимизировать сообщения для отдельного получателя; опираясь на профиль получателя в социальных медиа, искусственный интеллект Crystal сообщает отправителю о том, «как бы получатель хотел, чтобы с ним разговаривали» [8]; система Responsible [9] использует AI, чтобы посоветовать тем, кто пишет сообщения, как «выбрать правильный тон при создании сообщения электронной почты вашему боссу». Будущие системы AI-МС смогут делать больше, чем предлагать тексты. Поскольку алгоритмы создания естественного языка улучшаются [10], технологии AI будут способны полностью создавать сообщения от лица отправителя – включая создание онлайн профилей или даже появление сообщений в синхронном общении [11]. С ростом обеспокоенности относительно вовлеченности AI в коммуникацию людей (например, появление странных фейков, в которых AI используется в целях создания неправильного представления о том, что индивидуум говорит или делает в видеоролике или аудиозаписи [12; 13]), чрезвычайно важна потребность в понимании данного воздействия на межличностную коммуникацию.

ВЕЛИЧИНЫ AI-МС

Приведенные примеры иллюстрируют несколько величин, которые могут быть полезны для характеристики взаимодействий AI-МС (табл.). Первой является *магниту* вовлеченности со стороны AI агента. Эти изменения можно также ранжировать от минимальных предложений (например, изменение слова) до создания полного контента. Другим ключевым измерением является *тип медиа*, используемый в коммуникации, такой как текст, аудио и видео. Именуемые сегодня инструменты автоматического редактирования и фильтры – например, методы выявления и фиксации изображений, в которых сюжет размыт [14], – являются первыми шагами в сторону технологии, делающей желающего более привлекательным [15], заслуживающим большего доверия или более похожим на получателя [16].

Величины AI-МС

Величина	Определение	Примеры
<i>Мagnitude</i>	Степень изменения, с которой AI воздействует на сообщения	Проверка правописания или создание полностью новых сообщений
<i>Тип медиа</i>	Медиа, в которых действует AI (например, текст, видео, аудио)	Предложение текстовых ответов вместо изменения чьего-либо появления на видео
<i>Цель оптимизации</i>	Цель, ради которой AI оптимизирует сообщения	Стать привлекательным, заслуживающим доверия, с чувством юмора, доминирующим и т. п.
<i>Автономия</i>	Степень, с которой AI может воздействовать на сообщения без наблюдения отправителя	Отправитель выбирает между предложенным AI сообщением и вовлечением AI в коммуникацию с минимальными исходными данными от отправителя
<i>Ориентация ролей</i>	Роль, с которой AI действует от чьего-либо лица (например, отправитель или получатель)	Отправитель: предложение сообщений, улучшающих эффективность ответа; получатель: оценка на предмет лжи со стороны отправителя

В аудио можно манипулировать, чтобы заставить говорящего звучать более спокойно или авторитарно [17] или синтезировать полностью новую речь для говорящего [18]. Другие инструменты позволят реалистично изобразить вводящие в заблуждение сигналы в видео – скажем, поднятие тяжелого предмета или исполнение танца, как если бы это делалось профессионалом [19].

Синхронность – другое основное измерение систем AI-МС. Синхронные формы СМС вероятнее всего станут опосредованными AI, особенно в связи с недавним улучшением в манипуляции аудио или видео в реальном времени [12]. «Фильтры» видео в реальном времени позволяют индивидуумам изменять свою внешность [20] и выражения лиц в живой опосредованной компьютером беседе [13].

AI-МС можно также категоризировать по *цели оптимизации*, для которой она создается. Цели межличностной коммуникации множественны, поэтому уже существуют целевые таксономии для их описания [21]. Например, цели самопрезентации – казаться привлекательным, притягательным, или знающим – важные межличностные функции, помогающие AI стать, возможно, обученным для оптимизации коммуникации.

Автономия, данная AI для того, чтобы действовать от лица коммуниканта, – другое основное измерение. Она похожа на отношение администратор – агент [22], в котором индивидуум-коммуникант является главным лицом, делегирующим некоторый авторитет и автономно агенту AI. Например, в интеллектуальных ответах администратор реально сохраняет важную роль: он выбирает, какое из предложенных сообщений использовать, а какое – отклонить, и может также изменять сообщение. Будущие системы AI-МС могут получить даже большую автономию относительно вовлеченности в задачи коммуникации без наблюдения (напоминания) со стороны администратора – от автоматических и персонализированных пожеланий с днем рождения до автоматического составления расписания [11] или онлайн-сообщений о предстоящих встречах (беседах).

Наконец, важна *ориентация роли* агента AI. Большинство имеющихся примеров AI в средствах коммуникации ориентировано на отправителя, а мы полагаем, что получатели будут все больше пользоваться средствами AI. Переводчик Google позволяет отправителю и получателю взаимодействовать, используя AI в качестве посредника.

Можно представить себе и другие средства, ориентированные на помощь получателю, например, предложение извлечь социальные команды вызова [23] или выявить эмоции, обман и ложь [24] в речи в режиме реального времени. По мере расширения исследования AI-МС ожидается рост этой первоначальной группы величин.

ГРАНИЦЫ AI-МС

Чтобы внести ясность, полезно выделить некоторые примеры того, что мы, в соответствии с данным нами выше определением, не будем считать AI-МС. Самой тесно связанной является растущая область взаимодействия AI-человек или исследование взаимодействий человека с сущностями и другими интеллектуальными виртуальными агентами, *не* представляющими других индивидуумов, такими как поисковик Apple – Siri и его аналог от Amazon – Alexa. Эта область, иногда называемая коммуникацией человек-машина, перекрывает сферу AI-МС, хотя в нашей формулировке интересные вопросы включают введение AI, действующего в *коммуникации между людьми*. Например, устройства AI-МС можно применять для увеличения числа тех людей, с которыми индивидуум может поговорить межличностно (например, поиск офиса политической партии /политика или использование устройства с целью поговорить с как можно большим числом избирателей).

AI-МС можно концептуализировать еще шире, чтобы включать в нее все алгоритмы, опосредующие коммуникацию людей, подобные «Newsfeed» от Facebook и иному ранжированию контента, алгоритмы рекомендации и классификации (например, фильтры электронной почты, предложения от друзей), которые используют алгоритмы для поддержки коммуникации людей. Этот тип концептуализации является слишком большим, чтобы служить полезным научным подходом, поэтому мы исключаем такого рода алгоритмические средства, косвенно поддерживающие коммуникацию людей.

ПОВЕСТКА ИССЛЕДОВАНИЙ

Введение AI-МС имеет потенциал перевернуть и расширить существующее знание СМС, теории и научную сферу. Использование AI в межличностной коммуникации требует от СМС принятия на себя действия и посредничества и может разрушить существующую соци-

альную эвристику. Здесь мы стремимся заложить основу изучения появляющейся области AI-МС, обеспечивая повестку исследований, которая пересматривает основные темы СМС, начиная от лингвистических и коммуникативных процессов по отношению к связанной и межличностной динамике и до этических, культурных и политических применений. Предлагаем научные темы, усложняющиеся от низкого к высокому уровням анализа: от разработки AI-МС и ее непосредственного влияния на индивидуумов, взаимодействующих с этими системами, до относительных аспектов AI-МС и ее потенциального одновременного социального и культурного влияния. Наконец, поскольку системы AI-МС распространяются на более межличностные сферы, их общественные и этические применения все больше становятся заметными. Опишем эти области ниже.

Разработка AI-МС

Первая группа научных вопросов касается того, как люди *взаимодействуют и понимают* AI-МС. С точки зрения отправителей, как выбор дизайна влияет на их использование предложений AI-МС, таких как интеллектуальные ответы [25], или каково преимущество AI-МС в помощи в создании профиля? Что касается получателей, то осознают ли они тот факт, что AI-МС вовлечена в коммуникацию, и если да, то в чем проявляется эта осознанность? Недавняя работа предполагает, что смешение профилей, созданных AI и человеком, может оказать негативное влияние на рейтинг доверия *сомнительных или названных* как AI профилей [26]. Будущие исследования могут рассматривать тип и эффективность раскрытия вовлечения AI, одновременно анализировать факторы, которые заставляют людей предполагать, что AI вовлечен в коммуникацию.

Выбор дизайна также воздействует на интерпретацию действий с AI-МС, особенно с учетом того, что люди принимают исходные данные от AI иначе, чем исходные данные от людей [27; 28]. При каких условиях дизайна вовлечение AI будет восприниматься усиливающим действие отправителя, а когда считаться незаконным (узурпирующим)? Предыдущее исследование показывает, что, когда обусловленные задачами коммуникации беседы были безуспешными, участники приписывали значительно меньшую ответственность своим партнерам при использовании посредничества AI (интеллектуальные ответы) [29]. Исследование требуется, чтобы понять, когда AI-МС рассматривается как *фильтр* для представления людей, как независимый социальный агент, а когда и нечто среднее между ними.

Вопрос действия осложняется контекстом использования. Например, посредничество AI широко принято, когда используется для уточнения, такого как автокоррекция или машинный перевод между языками в коммуникации на основе текста [30]. Менее ясно то, как будет приниматься роль AI в оптимизации целей самопрезентации. Следовательно, достаточно изучить контексты, в которых применяется AI-МС, тип целей, используемых устройством для выполнения задачи, и как она (коммуникация) воспринимается в каждом контексте отправителем и получателем.

Влияние на язык

Из множества исследований СМС вопросы относительно того, как технология влияет на способ, с помощью которого люди пишут и интерпретируют сообще-

ния, представляют особый интерес [31]. AI-МС ставит эти вопросы по-новому и вводит много новых запросов.

Важным вопросом является потенциал AI-МС в формировании языка и мыслей человека. В соответствии с интерактивной моделью выравнивания [32] создание языка и понимание тесно переплетаются в процессе, который может помочь создать лингвистическое выравнивание между говорящими. Таким образом, когда созданный AI текст вставляется в диалог, то, вероятнее всего, он повлияет на это выравнивание с возможностью изменения не только выбора слов говорящим, но также и его партнеров. Например, функция интеллектуального ответа Gmail (рис. 1) предоставляет три опции для ответа, заполняющие синтаксический и семантический контент независимо от того, выберет ли их отправитель для использования или нет.

Исследование предложений «интеллектуальных ответов» в текстовом сообщении выявило, что они слишком позитивные (звучит возвышенно!) [25]. Данный избыток позитивного языка может заставить отправителя и получателя также использовать более позитивные высказывания в последующих сообщениях. Эти «интеллектуальные ответы», сейчас это часть популярной службы Google Gmail, а также «интеллектуальные составляющие», в которых Gmail предлагает автоматическое завершение слов и предложений, вводят AI-МС в миллионы коммуникаций по всему миру, однако мы мало знаем об их влиянии на использование динамики языка на индивидуальном или общественном уровне. С учетом ее (коммуникации) масштаба чрезмерно позитивные языковые предложения Gmail способны сдвинуть языковые нормы и ожидания, даже когда коммуниканты не пользуются этими инструментами, а также порождать долговременное изменение языка во времени.

Будущее влияние AI-МС на язык может состоять в том, как люди адаптируются друг к другу в межличностной коммуникации. Межличностная адаптация или изменение чьего-либо поведения, чтобы приспособить коммуникацию к партнеру, является фундаментом для социального взаимодействия [33]. Этот процесс может быть нарушен или усилен при вовлечении AI-МС. Что произойдет, если AI-МС будет играть роль в предложении или создании сообщений от лица одного из коммуникантов? Необходимо исследование, чтобы изучить влияние этих предложений и других языковых направлений AI-МС – от прагматических явлений (например, вежливость или грубость) до зарегистрированных явлений (гендерный стиль, региональные диалекты и внутригрупповые языки, используемые в онлайн сообществах) – все это может подвергаться влиянию AI.

Внедрение систем AI-МС может в дальнейшем «балканизировать» или гомогенизировать язык и языковые практики – разрешая или запрещая – личные или групповые изменения. В организационном контексте AI-МС может изменять рассеяние общей социальной семантики. Например, в процессе *вынужденного поиска* авторы используют понятия из ограниченной (например, политической или организационной) семантической популяции [34]. Если AI предложения исходят из этой популяции, то эта тенденция будет расширена, приводя к дроблению терминологии и понятий. Такие трудноуловимые решения, приводимые здесь в действие почти невидимо через простые механизмы AI-МС, такие как предлагаемые тексты, могут оказывать глубокое влияние на наш общий язык и терминологию.

Межличностная динамика

Исследование СМС часто изучает, как коммуникация через различные технологии может изменять наши сообщения, как они создаются и как воспринимаются. AI-МС поднимает важные вопросы относительно этих процессов. Рассмотрим, например, основные аспекты сверхчувствительной модели межличностной коммуникации [35]: отправитель, получатель, канал и обратная связь. Как будет трансформирована избранная самопрезентация отправителя, если инструменты AI можно использовать для его изменения сообщений, чтобы произвести желаемое впечатление? Будут ли получатели менять подход к тщательному изучению общения, узнав, что сообщение оптимизировано машиной?

Обработка социальной информации [35] предполагает, что характеристики языкового контента и стиля являются первичными условиями межличностной информации в СМС, но что произойдет с этими сигналами, если сообщения не только трансформируются через технологию, но и изменяются или создаются AI? Одна из возможностей состоит в том, что эти сигналы потеряют свою диагностику в межличностной информации, так как получатели поймут, что AI изменяет язык и стиль контента. Однако другая возможность – воспринятое действие останется у отправителя, так как получатели продолжают нести перед отправителем ответственность за ошибки написания, а не за неудачи системы проверки написания [36].

Самопрезентация, формирование впечатления и доверия

Устройство AI-МС вероятно должно влиять и на то, как люди представляют самих себя онлайн, и на то, как они оценивают других. Ожидается, что теории и подходы, связанные с сигналами и доверием, такие как Профиль как Обещание [37], и Теория гарантии [38] и сигналов [39], будут нуждаться в обновлении для AI-МС, где использование инструментов AI может подрывать или улучшать надежность профилей и сообщений. Например, в СМС оценивающие должны интерпретировать сигналы, представленные онлайн, чтобы вывести характеристики индивидуума, а осознанность того, что партнер по коммуникации использует AI, может повлиять на восприятие онлайн профилей. В соответствии с Теорией гарантии порождаемая AI информация может восприниматься как более гарантированная [38] с учетом того, что сообщения и профили можно рассматривать более объективно, когда они создаются компьютером [40].

Наоборот, посредничество AI в самоописаниях может нарушать ожидания и вызвать беспокойство относительно обмана и манипуляции. Недавняя работа продемонстрировала сильное влияние на оценку доверия, когда AI вовлечен в самопрезентацию в контексте Airbnb [26], предполагая, что в системе, смешивающей написанные людьми и AI профили, доверие к воспринимаемым профилям, написанным AI, будет снижаться. Это влияние, удваивающее «эффект отвечающего», демонстрирует сложную роль, которую AI-МС может играть в теориях относительно онлайн самопрезентации.

AI-МС также должна влиять на то, как описывают себя индивидуумы и как раскрывают информацию онлайн, потенциально требующую пересмотра подходов и теорий СМС, связанных с самораскрытием. Например, центральный принцип подхода Профиль как Обеща-

ние [37] состоит в ожидании людей, что онлайн профиль будет только приближенным отражением их офлайн идентичности с учетом ограничений представления самого себя онлайн – но что он может быть изменен присутствием AI, приводя к разным динамикам самопрезентации и самораскрытия. И для получателя, и для отправителя эти вопросы не ограничиваются одним текстом, а распространяются на презентацию через изображения и видео.

Обратная связь и отношения

Эффекты AI-МС могут выходить за рамки краткой коммуникации и первого впечатления в целях введения новых эффектов (влияний) на себя, партнеров и отношения. Смещение идентичности, например, предполагает, что избранная самопрезентация в СМС приводит к соответствующим переменам в том, как мы сами воспринимаем себя в нашем реальном мире [41]. Однако неясно, когда и как произойдет смещение идентичности, если машина изменяет поведение одного коммуниканта. Если AI изменяет сообщения отправителя в сторону большего позитива, веселого настроения или экстравертности, то сместится ли самовосприятие отправителя в эту же сторону? Есть четкие виды поведения, вызывающие более тревожащие последствия, такие как использование AI в целях манипуляции или обмана других, смещая идентичность в сторону рассмотрения самого себя управляемым или обманутым.

AI-МС может также влиять на то, как мы поддерживаем и развиваем важные отношения с другими людьми, следуя влиянию СМС относительно близости, привлекательности и поддержки отношений [42]. Количество усилий, которое партнер тратит на отношения, играет важную роль в поддержании отношений. С вводом AI в процесс коммуникации, такой как автоматические пожелания по случаю дня рождения на сайтах социальных сетей, будет ли подрываться относительная ценность выражения близости или благодарности, когда вступает в свою роль AI? С помощью научных усилий можно оценивать, когда AI-МС подрывает восприятие усилий в отношениях, и выходит ли этот эффект за рамки осязаемой близости.

Политика, культура и этика

По мере развития систем AI-МС представляется важным оценить их социальные и этические применения. Системы AI-МС уже действуют масштабно и могут широко распространять социальное влияние, которое диспропорционально воздействует на множество ранимых людей. Исследование должно изучить, как системы AI-МС сохраняют баланс нормативных ценностей, какие факторы необходимо учесть принимающим решения относительно управления AI-МС людям и какие практики могут использовать разработчики, чтобы снизить это беспокойство. Ниже опишем некоторые внушающие беспокойство сферы.

Предвзятость и справедливость

Системы AI-МС обучаются на существующей среди людей коммуникации: сохранившиеся предвзятости в отношении распознавания данных, вероятно, должны быть продолжены и усилены AI. Эффект от таких систем в масштабе может состоять в том, чтобы ужесточить структуры коммуникативной способности, что проявля-

ется в нормах общения. Когда AI-МС изменяет коммуникацию, чтобы стать социально влиятельным (например, звучать более авторитарно), можно спросить о применении нормализации определенных стилей взаимодействия при включении в категорию других лиц. Например, расширение Chrome *Just Not Sorry* предупреждает пользователей, когда они используют качественные слова, такие как «извини», и «Я думаю» так, что эти слова могут подорвать сообщение пишущего лица [43]; данное расширение предназначено для женщин и, поощряя их делать такие изменения, нормализует грубый язык как «правильный» способ общения.

В других контекстах AI-МС может помочь уменьшить (смягчить) межличностный уклон. Если AI-МС подает сигнал доверия между участниками социального обмена на основе равенства, эти компенсирующие сигналы могут нейтрализовать стереотипы, которые в противном случае будут препятствовать транзакции [44]. Альтернативно AI-МС может непосредственно поощрять пользователей изменять коммуникацию, что несет оттенок предубеждения, например, использование предупреждающего диалога, когда индивидуум пытается отправить отрицательный отзыв.

Прозрачность: должна ли AI-МС раскрывать себя?

Еще одну озабоченность вызывает то, как и когда и должны ли вообще системы AI-МС раскрывать свое существование и функциональность перед пользователями (отправителями и получателями). Иногда AI-МС могут быть столь общими и непротиворечивыми, что необходимо их раскрытие: например, кажется абсурдным требовать, чтобы сообщение несло декларацию, что оно было исправлено автоматически. В других контекстах можно требовать раскрытия информации в целях помочь в собственной интерпретации сообщения и намерении отправителя. Например, текст, автоматически переведенный с помощью Google Translate, часто несет отказ от ответственности, признающий это; это раскрытие поможет читателю лучше интерпретировать текст с ошибками как результат несовершенства AI, а не как намерение пишущего лица.

Имеются и другие этические и социальные причины для прозрачности. В мае 2018 г. Google продемонстрировала жизнеспособность помощника AI Duplex на конференции путем назначения встречи по телефону. Устная презентация Duplex была полна междометий типа «ums» и «mm-hmm's» и не раскрыла, что это был робот. Google обзала Duplex взаимодействовать с людьми естественно, но критика разнесла эту технологию в пух и прах за отказ раскрыть получателю ее истинный характер [11].

Но раскрытие осложняется. Недавняя законодательная инициатива шт. Калифорния (the B.O.T. Act of 2018) обязует подтвердить «любое лицо, [использующее] социальную сущность для коммуникации или взаимодействия с человеком онлайн [в целях раскрытия], что данная сущность не является человеком» [45]. Эта мера, кажется, полагается на предположение, что правовая оговорка будет лучше оснащать получателей для проведения оценки намерения и правдивости сообщения. В дополнение к эмпирической неуверенности этого предположения [26; 28] ученые-юристы отмечают, что масштабные законы раскрытия AI-МС могут иметь нежелательные политические последствия: например угроза правам Первой поправки

относительно анонимности речи или перегрузка отправителей, имеющих потребности в доступе (например, пользователи преобразования текста в речь).

Даже при гарантии прозрачности есть важные вопросы относительно того, какая прозрачность требуется и каким целям она служит в разных контекстах. Достаточно ли ее раскрытия для использования системы AI-МС, или следует выяснить больше информации об ее объективной функции? Насколько гранулированным должно быть объяснение? Касаются ли отправителей и получателей одни и те же требования прозрачности? Эти и другие вопросы, вероятно, возникнут на уровне публичного дискурса и связанной с AI-МС политики.

Неправильное представление и манипуляция

В той степени, в какой все техники самопрезентации могут быть поняты как стратегически разработанные, чтобы произвести впечатление на других определенным пониманием коммуникативных сторон и ситуации [37]. Но некоторые техники вынуждают нас пересекать линию от презентации к *неправильной* презентации, от убеждения к манипуляции. Манипулятивные AI-МС могут стремиться существенно ввести партнеров в заблуждение вероятно путем использования когнитивных уязвимостей или навязывания ложных убеждений [46].

Соответствующие линии между допустимой *неправильной* презентацией и неэтичной манипуляцией строго не фиксируются и не зависят от уязвимостей сторон, цели коммуникации и последствий действия. Ученые компании Facebook готовят системы машинного обучения к распознаванию фотографий, на которых субъект изображен нечетко, и помещать похожие, но четкие [14]. Хотя осуществление этого технически искажает реальность, многие могут посчитать это социально и этически приемлемым. Фактически неудачу отретушировать чьи-либо непривлекательные фотографии можно рассматривать как невежливое действие.

На другой стороне этого спектра располагаются *серьезные фейки*, синтезированные AI видео предполагают обман или манипуляцию. Модель Sundar's MAIN [40] включает «эвристику реализма», предсказывающую, что люди охотнее поверят аудиовизуальной модальности из-за ее высокого сходства с реальным миром. Такие манипуляции могут иметь широкое влияние, например, на политические коммуникации с учетом того, что такие паттерны как привлекательность могут предсказать результаты выборов [47], а изменение высоты голосового тона может улучшить восприятие лидера [17]. Возможности AI позволяют придать лицам кандидатов трансформированный вид на фотографиях или в реальном времени в целях соответствия предпочтениям избирателей.

Между этими двумя полюсами социальные и этические применения являются более мутными и контекстуально зависимыми. Мы не будем возражать против использования AI-МС в бизнес-обмене, где, например, можно ожидать, что самопрезентация будет высоко стратегической или даже желательной; но мы можем рассматривать ее относительно безотказной в тесных личных обменах, отмеченных более высокой степенью ожидания аутоидентификации (например, между романтическими партнерами). В ответ на доступность новых технологических норм принятия скорее всего будут эволюционировать.

ЗАКЛЮЧЕНИЕ

AI-МС быстро продвигается с потенциально важным влиянием на области, начиная от межличностных отношений и до принятия политических решений. В самом худшем случае эффекты могут быть весьма смазанными. Если манипулятивное и ложное сообщение гораздо легче создать и труднее выявить, то AI может толковать почти все СМС как ненадежные, подрывая доверие в любом взаимодействии, кроме тех, которые происходят лицом к лицу. Но AI также имеет потенциал совершенствовать коммуникацию людей за счет усиления нашей способности коммуницировать друг с другом и улучшать доступность таких взаимодействий в СМС каналах.

Остается открытым вопрос о том, как средства AI должны использоваться в целях оптимизации коммуникаций и достижения межличностных целей. Новая аппаратура, программное обеспечение, законы и политика сформируют роль, которую AI будет играть в коммуникации в предстоящие годы. С должной опорой на очевидность можно результативно изучать новые проблемы, возникающие в AI-МС. Чтобы разработать, внедрить и регулировать эти системы ответственно, мы должны сформировать основополагающее эмпирическое понимание их влияния на разнообразие поведений, включая управление впечатлением, доверие, обман, использование языка, взаимосвязи и другие основные факторы. Мы призываем сообщество присоединиться к этому исследованию и расширить рассмотрение AI-МС как важной научно-исследовательской темы.

Благодарность. Этот материал основан на работе, созданной при поддержке гранта № CHS 1901151/1901329 Национального научного фонда. Мы хотим выразить благодарность Сю Ма, чья работа по онлайн доверию легла в основу формирования подхода к AI-МС, Сю также нашел решение для фазы AI-МС, и другим коллегам – Морису Джакешу, Мальту Джунгу, Меган Френч, Джесс Хохенштейн, Санни Лью, а также Станфордской лаборатории социальных медиа за идеи и отзывы о более ранних вариантах статьи.

ЛИТЕРАТУРА

1. *Herring S. C.* Computer-mediated communication on the Internet// Annual Review of Information Science and Technology. — 2002. — Vol. 36, No. 1. — P. 109–168.
2. *Walther J. B., Parks M. R.* Cues filtered out, cues filtered in. Handbook of interpersonal communication, 2002. — P. 529–563.
3. *Russell S. J., Norvig P.* Artificial intelligence: A modern approach. — Pearson: Third Ed., 2010.
4. *Thurlow C., Lengel L., Tomic A.* Computer mediated communication: Social interaction and the internet. — Thousand Oaks, CA: Sage, 2009.
5. *Grammarly.* Free grammar checker - Grammarly. — 2018. — <https://www.grammarly.com/>.
6. *Pavlick E., Tetreault J.* An empirical analysis of formality in online communication// Transactions of the Association for Computational Linguistics. — 2016. — Vol. 4. — P. 61–74.
7. *Ma X., Hancock J.T., Lim Mingjie K., Naaman M.* Self-disclosure and perceived trustworthiness of Airbnb host profiles// Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW'17 (pp. 2397–2409). — New York: ACM, 2017.

8. *Zeide E.* Algorithms can be lousy fortune tellers. — Slate, 2018. — http://www.slate.com/articles/technology/future_tense/2015/05/crystal_app_algorithmic_fortunetelling_for_employers_and_potential_customers.html.
9. *Boomerang.* Responsible: Write better email. — 2018. — <https://www.boomerangmail.com/responsible/>.
10. *Graves A.* Generating sequences with recurrent neural networks. — [arXiv preprint. — 2013.— arXiv: 1308.0850.]
11. *Statt N.* Google now says controversial AI voice calling system will identify itself to humans. —The Verge, 2018, May 10. — <https://www.theverge.com/2018/5/10/17342414/google-duplex-ai-assistant-voice-calling-identify-itself-update>.
12. *Sumajanakorn S., Seitz S. M., Kemelmacher-Shlizerman I.* Synthesizing Obama: Learning lip sync from audio// ACM Transactions on Graphics (TOG). — 2017. — Vol. 36. — P. 95.
13. *Thies J., Zollhofer M., Stamminger M., Theobalt C., Niessner M.* Face2Face: Real-time face capture and reenactment of RGB videos// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — P. 2387–2395.
14. *Emerson S.* Facebook wants to use AI to replace your eyeballs in photos where you blinked// Motherboard. — 2018, June 18. — https://motherboard.vice.com/en_us/article/a3a8n8/facebook-wants-to-use-ai-to-replace-your-eyeballs-in-photos-where-you-blinked-artificial-intelligence.
15. *Leyvand T., Cohen-Or D., Dror G., Lischinski D.* Data-driven enhancement of facial attractiveness// ACM SIGGRAPH 2008 Papers, SIGGRAPH'08. — 2008, (pp. 381–389). — New York: ACM, 2008.
16. *Todorov A., Dotsch R., Porter J. M., Oosterhof N.N., Falvello V. B.* Validation of data-driven computational models of social perception of faces// Emotion. — 2013. — Vol. 13. — P. 724.
17. *Klofstad C. A., Anderson R. C., Peters S.* Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women// Proceedings of the Royal Society B: Biological Sciences. — 2012. — Vol. 279, No. 1738. — P. 2698–2704.
18. *Vincent J.* Lyrebird claims it can recreate any voice using just one minute of sample audi. —The Verge, 2017, Apr. 24. — <https://www.theverge.com/2017/4/24/15406882/ai-voicesynthesis-copy-human-speech-lyrebird>.
19. *Chan C., Ginosar S., Zbou T., Efros A.A.* Everybody dance now. — [arXiv preprint. — 2018. — arXiv:1808.07371.]
20. *Shah D., Allen K.* Chinese vlogger who used filter to look younger caught in live-stream glitch. — BBC News, 2019, July 30. — <https://www.bbc.com/news/blogs-trending-49151042>.
21. *Chulef A. S., Read S. J., Walsh D. A.* A hierarchical taxonomy of human goals// Motivation and Emotion. — 2001. — Vol. 25, No. 3. — P. 191–232.
22. *Sappington D. E.* Incentives in principal-agent relationships// Journal of Economic Perspectives. — 1991. — Vol. 5, No. 2. — P. 45–66.
23. *Razavi S. Z., Ali M. R., Smith T. H., Schubert L. K., Hoque M. E.* The LISSA virtual human and ASD teens: An overview of initial experiments// International Conference on Intelligent Virtual Agents (pp. 460–463). — Cham: Springer, 2016.
24. *Sen T., Hasan M. K., Tran M., Levin M., Yang Y., Hoque M.E.* Say CHEESE: Common Human Emotional Expression Set Encoder and its Application to Analyze Deceptive Communication// 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018) (pp. 357–364). — IEEE, 2018.

25. *Hobenstein J., Jung M.* AI-supported messaging: An investigation of human-human text conversation with AI support// CHI EA '18: Extended abstracts of the 2018 CHI conference on human factors in computing systems (pp. LBW089:1–LBW089:6). — New York: ACM, 2018.
26. *Jakesch M., French M., Ma X., Hancock J.T., Naaman M.* AI-mediated communication: How the perception that profile text was written by AI affects trustworthiness// CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. — 2019. — P. 1–13.
27. *Dietvorst B. J., Simmons J. P., Massey C.* Algorithm aversion: People erroneously avoid algorithms after seeing them err// Journal of Experimental Psychology: General. — 2013. — Vol. 144, No. 1. — P. 114.
28. *Waddell T.* A robot wrote this? How perceived machine authorship affects news credibility// Digital Journalism. — 2018. — Vol. 6, No. 2. — P. 236–255.
29. *Hobenstein J., Jung M.* AI as a moral crumple zone: The effects of AI-mediated communication on attribution of responsibility and perception of trust. Computers in Human Behavior. — 2019, in press. — doi: 10.1016/j.chb.2019.106190.
30. *Xu B., Gao G., Fussell S. R., Cosley D.* Improving machine translation by showing two outputs// Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, CHI14, 3743–3746. — New York: ACM, 2013.
31. *Herring S.C.* Language and the internet//W. Donsbach (Ed.), The international encyclopedia of communication (pp. 2640–2645). — 2008. — doi:10.1002/9781405186407.wbiecl005
32. *Pickering M. J., Garrod S.* An integrated theory of language production and comprehension// Behavioral and Brain Sciences. — 2013. — Vol. 36, No. 4. — P. 329–347.
33. *Toma C. L.* Towards conceptual convergence: An examination of interpersonal adaptation// Communication Quarterly. — 2014. — Vol. 62. — P. 155–178.
34. *Margolin D. B., Monge P.* Conceptual retention in epistemic communities: Examining the role of social structure// P.Moy (Ed.), Communication and community (pp. 1–24). — New York: Hampton Press, 2013.
35. *Walther J. B.* Theories of computer-mediated communication and interpersonal relations//M. L. Knapp & J. A. Daly (Eds.), The Sage Handbook of Interpersonal Communication. (pp. 443–479). — Sage Publications: Thousand Oaks, CA, 2011.
36. *Figueredo L., Varnhagen, C. K.* Didn't you run the spell checker? Effects of type of spelling error and use of a spell checker on perceptions of the author// Reading Psychology. — 2005. — Vol. 26, No. (4–5). — P. 441–458.
37. *Ellison N. B., Hancock J. T., Toma C. L.* Profile as promise: A framework for conceptualizing veracity in online dating self-presentations// New Media & Society. — 2012. — Vol. 14, No. 1. — P. 45–62.
38. *DeAndrea D. C.* Advancing warranting theory// Communication Theory. — 2014. — Vol. 24, No. 2. — P. 186–204.
39. *Donath J.* Signals in social supernets// Journal of Computer-Mediated Communication. — 2007. — Vol. 13, No. 1. — P. 231–251.
40. *Sundar S. S.* The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility// M. J. Metzger & A. J. Flanagin (Eds.), Digital Media, Youth, and Credibility. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning (Vol. 2008, pp. 73–100). — Cambridge, MA: The MIT Press, 2008.
41. *Gonzales A. L., Hancock J. T.* Identity shift in computer-mediated environments// Media Psychology. — 2008. — Vol. 11. — P. 167–185.
42. *Tong S., Walther J. B.* Relational maintenance and CMC // K. B. Wright & L. M. Webb (Eds.), Computer-Mediated Communication in Personal Relationships (pp. 98–118). — Bern, Switzerland: Peter Lang US, 2011.
43. *Canterucci C.* New Chrome app helps women stop saying “just” and “sorry” in emails. — Slate, 2015, Dec. 29. — http://www.slate.com/blogs/xx_factor/2015/12/29/new_chrome_app_helps_women_stop_saying_just_and_sorry_in_emails.html.
44. *Levy K., Barocas S.* Designing against discrimination in online markets// Berkeley Technology Law Journal. — 2018. — Vol. 32, No. 2. — P. 1–57.
45. *Williams J.* Should AI always identify itself? It's more complicated than you might think. — Electronic Frontier Foundation, 2018. — <https://www.eff.org/deeplinks/2018/05/should-ai-always-identify-itself-its-more-complicated-you-might-think>.
46. *Susser D., Roesler B., Nissenbaum H.* Online Manipulation. Working paper presented at Privacy Law Scholars Conference, 2018. — 2018.
47. *Mattes K., Spezio M., Kim H., Todorov A., Adolphs R., Alvarez R. M.* Predicting election outcomes from positive and negative trait assessments of candidate images// Political Psychology. — 2010. — Vol. 31, No. 1. — P. 41–58.

Построение матрицы активности метаданных: концептуализация и обоснование связей между жизненным циклом исследований и метаданными*

Соня ПАСКУА
(Sonia PASCUA),

Кай ЛИ
(Kai LI)

Университет Дрекселя, шт. Пенсильвания,
США

Нельзя недооценивать роль метаданных в поддержке исследований; более того, трудно развить системное понимание активности метаданных в ходе исследовательского процесса. В статье предварительно анализируется, как активности метаданных встроены в жизненные циклы исследования и данных. В частности, выделяются некоторые основные активности метаданных, связанные с компонентами общего исследовательского процесса, от формулирования гипотезы до распространения результатов и управления данными. В ходе исследования возникают эпистемологические вопросы относительно присутствия активности метаданных во время проведения исследования и управления данными. Эта работа концептуализирует и обосновывает связь между метаданными и жизненными циклами исследований и данных, а также представляет четкое отображение, выявляющее профиль этих активностей, и создает впечатление относительно ценности метаданных в области научных исследований и управления данными.

ВВЕДЕНИЕ

Метаданные играют ключевую роль в современной науке, в частности, в исследовании, включающем большие массивы цифровых данных. Рост числа данных вводит новые эпистемологические связи между данными и научной деятельностью [1]. Эти эпистемологические связи ведут к росту положения данных в современном научном предприятии, и задача управления данными становится базовой научной деятельностью [2]. Метаданные служат платформой для инициации совместности определений и согласия в терминологиях между учеными. С учетом роста интенсивности использования данных в науке представляется важным оценить значимость метаданных в проведении исследования, в частности с использованием количественных измерений.

Понятие капитала метаданных, введенного Уиллисом и др. [3], предложило средства измерения важности и количественной ценности метаданных. Кроме того, оно выяснило практики и цели метаданных, способные внести вклад в совокупность знания по поддержке имеющей больше возможностей взаимодействовать среды, т.е. открытой для полидисциплинарной и трансдисциплинарной науки. Что касается континуума, то в данном исследовании говорится о системном анализе схем метаданных, а наша статья фокусируется на другой стороне спектра и изучает различные модели исследований и данных сквозь призму метаданных. Мы установили четкую связь между исследованием, данными и метаданными через прямой подсчет активностей метаданных среди научной деятельности и данных для убедительного измерения капитала метаданных.

Понятие *капитала метаданных* обеспечивает различный взгляд на изучение ценности метаданных путем измерения затрат и выгод метаданных в процессах, в которых создаются и повторно используются метаданные [4,5].

* Перевод Pascua S., Li K. Toward a metadata activity matrix: Conceptualizing and grounding the research lifecycle and metadata connections // Proc. Int'l conf. on Dublin Core and Metadata Applications 2019. — P. 42-48. — <https://dcpapers.dublincore.org/pubs/article/view/4235/2429>

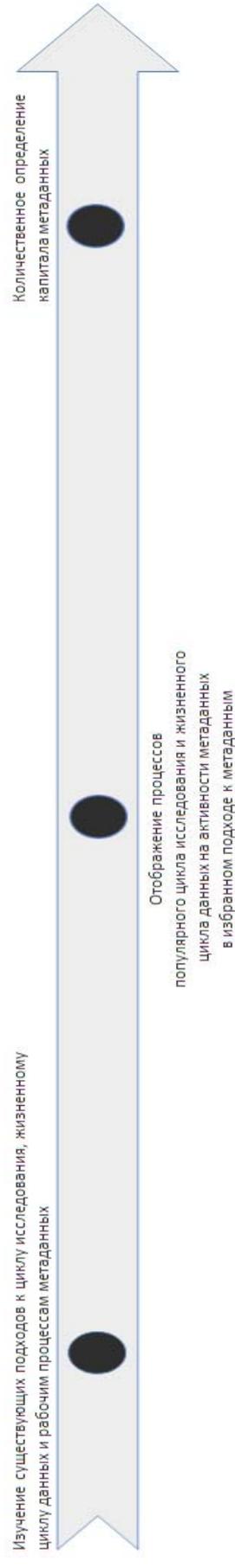
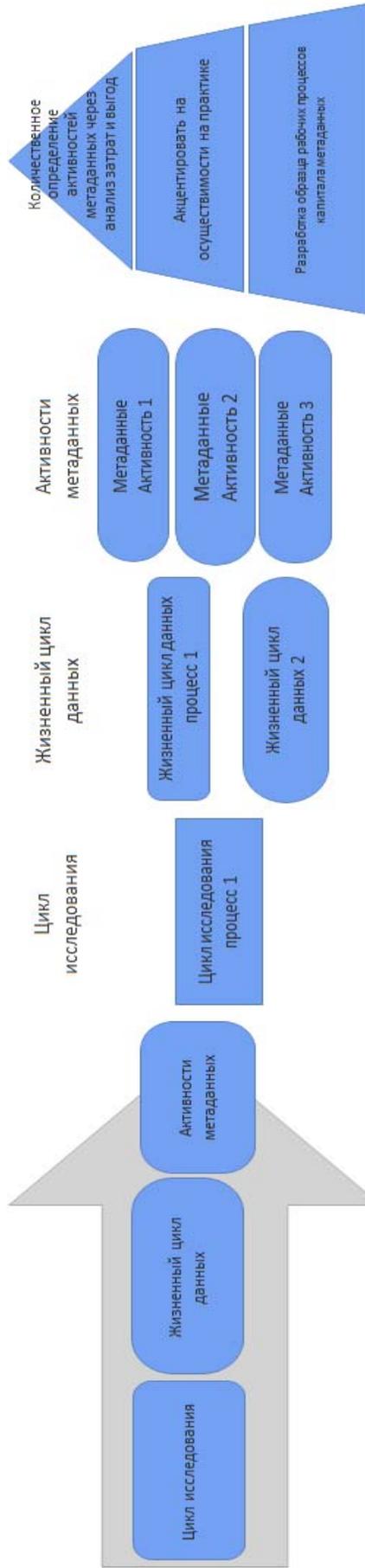


Рис. Концептуальный подход Проекта капитала метаданных

Исследование капитала метаданных также проводится для оценки степени его увеличения в ходе лабораторного процесса the Viral Vector Core Laboratory Национального института наук о состоянии окружающей среды [6]. В статье анализируется потребность дальнейшего изучения взаимосвязи между капиталом метаданных и исследовательскими процессами для улучшения нашего понимания как метаданные способствуют научной практике.

Одним из подходов к достижению этой цели исследования является анализ того, как активности метаданных и результаты встроены в модели жизненного цикла исследования. Хотя критика относительно того факта, что модели жизненных циклов не могут полностью раскрыть тонкости действующих процедур [7], усиливается, оба типа моделей можно использовать в качестве значимых представителей исследовательских установок в реальном мире, чтобы охарактеризовать более контекстуализированное знание о метаданных [8].

В качестве первого шага нашего анализа оценивались различные модели жизненных циклов исследований и данных и выяснилось, что одна выбранная модель подходит для целей данного исследования и отображает эти модели по нескольким высокоуровневым активностям метаданных с вероятностью формулировки матрицы метаданных. Хотя модели – абстракции общих компонентов жизненного цикла исследования, они полезны в представлении общих процессов и шагов встраиваемости метаданных.

Цели этого исследования следующие: 1. установить наиболее соответствующие жизненному циклу исследования модели в контексте изучения активностей метаданных; 2. получить высокоуровневые активности исследования и метаданных в выбранных моделях; 3. понять масштаб и расширение метаданных путем отслеживания активности метаданных в научно-исследовательской деятельности высокого уровня и оценить значимость метаданных для капитала метаданных.

МЕТОДОЛОГИЯ

На начальном этапе изучился ряд моделей в области исследования и были отобраны наиболее популярные модели его жизненного цикла и данных, которые принимают в расчет некоторые теоретически обусловленные соображения. Отобранные модели обсуждались и применялись в исследовании для построения связей — как метаданные способствуют научно-техническому исследованию и данным. Оценивалось, как они служат в качестве механизма активности метаданных. Матрица активностей жизненного цикла исследования и данных — в отличие от активностей метаданных — была сформулирована, чтобы отобразить существование метаданных в проведении исследования.

Понятие капитала метаданных было предложено с целью заполнить пробел в значимости метаданных путем измерения их затрат и выгод наряду с процессами, в которых используются метаданные [4, 5].

Рисунок отражает концептуальный подход Проекта капитала метаданных, частью которого является наше исследование. Он (подход) представляет концептуализацию, подчеркивающую связи метаданных с моделью жизненного цикла исследования и активностями метаданных, с последующим количественным определением

событий метаданных в их ситуационных контекстах. Он также предлагает разработку и формулирует образец поведения капитала метаданных на основе результатов и выводов этого исследования. Знание процесса метаданных в различных дисциплинах варьируется, хотя имеются явные различия в фазах; выбор этого подхода, соблюдение (исполнение) активности метаданных и наоборот неисполнение, были предметом количественного определения затрат и выгод непосредственно ведущих к капиталу метаданных. Что касается части этого проекта относительно цели исследования, то проводился эмпирический анализ и использовался качественный аудит этого подхода.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Этот раздел представляет наши предварительные результаты на основе отобранных моделей знания и соответствующие суждения относительно отображения по этим моделям.

Модель жизненного цикла исследования

Вопрос о том, как проводится исследование, был популярной темой в философии науки. Один, заслуживающий особого внимания спор вокруг этой темы (является ли или должно ли быть индуктивным или дедуктивным научное обоснование) продолжался очень долго [9]. Утверждалось, что научное исследование все больше становится дедуктивным после появления научной парадигмы, обусловленной данными [10]. Однако имеются также контраргументы относительно того, что большое число данных само по себе не может полностью влиять на традиционную научную эпистемологию; научное исследование в форме обусловленной данными все еще нуждается во многих традиционных индуктивных элементах, таких как существовавшие до этого научные законы и теории [11].

Несмотря на различие в стилях обоснования, существует относительно субъективная функциональная единица в процессах исследования, которая приводит к окончательным научным результатам. Эти единицы — важный маркер родственных служб (таких как библиотечная служба и /или служба данных), чтобы идти в ногу с помощью внутриситуационных научных контекстов [12;13]. Более того, эти единицы часто суммируются в модели жизненного цикла исследования в подобие линейных подходов.

На сетевом сайте Национальной сети медицинских библиотек (nml.gov) пять моделей показаны как связанные ресурсы, обсуждающие жизненные циклы исследования. Перечислим их ниже:

1. JISC Research Lifecycle Model
2. Oregon State University Library— Research Lifecycle
3. Research Lifecycle at the University of Central Florida Libraries
4. University of Virginia Library – Steps in the Research Lifecycle
5. University of Western Australia – Research Data Management Toolkit: Research Lifecycle Subject

Принимая во внимание первые три верхние модели, доступные для публичного пользования (без доступа и доверенных лиц), включая жизненный цикл исследования [14], табл. 1 показывает сравнительную матрицу их активностей для понимания метаданных.

Оценка некоторых популярных моделей жизненного цикла исследования

Модель жизненного цикла исследования/активности	Высший уровень
Предшествующая	Гипотеза
	Исследование
	Интерпретация
	Синтез
	Публикация
Oregon State University Library	Организация файлов
	Присвоение точных ролей для обеспечения качества/контроля качества
	Документация контекста массива данных
	Документация и управление версиями файлов
	Определение подходящего формата файла
The UCF Libraries RLC Committee model	Планирование
	Менеджмент проекта
	Публикация и представление
	Сохранение и распространение
	Престиж, влияние и открытие

Модель, предложенная автором работы [14, с. 6], — модель, удовлетворяющая обоим критериям, управляемым скорее функционально, чем логически, и тому, как объекты данных по возможности вовлечены в исследовательские процессы, несмотря на учет дедуктивного предположения. Она определяет следующие шесть частей циклического подхода: 1. Гипотеза, 2. Исследование, 3. Интерпретация, 4. Синтез, 5. Публикация, 6. Повторное использование. В частности, она учитывает точку зрения, основанную на центральности данных, т.е. все исследовательские шаги располагаются вокруг существования объектов научных данных. Это подчеркивает важность управления данными и активностями на каждом шаге научного исследования. Даже если автор не намерен дальше совершенствовать эту модель, то, как отмечалось, имеет место когерентность между этой моделью жизненного цикла исследования и данными, которые безоговорочно рассматривались для отслеживания активностей метаданных.

Модель жизненного цикла данных

Большинство моделей жизненного цикла данных включено в авторитетный обзор существующих моделей Болла [12]. Ниже приводятся 9 различных моделей, описанных и рассмотренных в Манифесте обзора моделей управления жизненным циклом данных, опубликованным 13 февраля 2012 г. Центром инновационного дизайна и производства исследований.

1. DCC Curation Lifecycle Model
2. I2S2 Idealised Scientific Research Activity Lifecycle Model
3. DDI Combined Life Cycle Model
4. DDI Combined Life Cycle Model
5. ANDS Data Sharing Verbs
6. DataONE Data Lifecycle
7. UK Data Archive Data Lifecycle
8. Research 360 Institutional Research Lifecycle
9. Capability Maturity Model for Scientific Data Management

Среди всех включенных модель Capability Maturity Model for Scientific Data Management (CMMSDM) была

признана как самая используемая [15], так как она предлагает точное описание, связанное с практикой управления объектами данных в контексте научной практики, и рассматривает масштабный функциональный подход к жизненному циклу данных. Она является также моделью, используемой в этой статье для формулировки матрицы метаданных.

Эта модель определила 4 этапа областей процесса обработки данных, включающих 1) получение, обработку и гарантию качества данных, 2) описание и представление данных, 3) распространение данных и 4) услуга хранилища и хранения. Так как модель входит в активные исследовательские процессы, ее этапы демонстрируют сильную когерентность с этапами исследований, обсуждавшимися ранее: хотя различная природа публикаций и объектов данных выдвигает разные требования к тому, как эти объекты следует обрабатывать, они (этапы модели), тем не менее, следуют очень похожей дорогой, где этапы модели и этапы исследований глубоко вовлечены в совместное производство друг друга [16].

Отображение активностей метаданных жизненных циклов исследований и данных строящейся матрицы метаданных

Вышеупомянутые отобранные модели жизненного цикла исследования и данных и идентификация активностей этих метаданных, а также их концептуальная карта представлены в табл. 2. Хотя эта карта не предполагает описания временных взаимосвязей между жизненными циклами исследования и данных, она отражает похожие процедуры обеих активностей жизненного цикла исследования и данных, включающие активности метаданных. Обе модели олицетворяют активности, присущие метаданным 1) понятие и экспозиция, 2) формулировка метаданных и построение, 3) управление метаданными, 4) использование метаданных и манипуляция и 4) ценность метаданных и значение. Можно также дедуктивно вывести, что имелись активности и в оценённых моделях, которые не распознают проведения активностей метаданных. Это в дальнейшем следует подтвердить как еще один успех данного исследования.

**Построение матрицы метаданных: отображение активностей
жизненных циклов исследования и данных**

Жизненный цикл исследования (предшествующий)	Жизненный цикл данных (CMMSDM)	Активности метаданных
Гипотеза		
Исследование	Получение данных, обработка и гарантия качества	✓
Интерпретация	Описание данных и презентация	✓
Синтез		
Публикация	Распространение данных и сохранение	✓
Повторное использование		

Оно (отображение активностей метаданных) предлагает целостный подход для наших следующих шагов исследования, кратко рассматриваемых в разделе «Заключение».

ЗАКЛЮЧЕНИЕ

Это изучение предлагает предварительный анализ взаимосвязей между исследованием, жизненным циклом данных и тем, как активности метаданных вовлечены в эти процессы. Это первый шаг количественного определения Проекта капитала метаданных, который ориентирован на формулировку подхода к оценке затрат и выгод работы и результатов метаданных в исследовательском процессе с использованием количественных методов.

Несколько выводов, информирующих о следующем шаге нашего проекта, можно сделать из обсуждения. Во-первых, наш анализ жизненных циклов исследования и данных демонстрирует, что эти две модели вместе предлагают когерентный подход к научному обусловленному данным исследованию, который в дальнейшем может быть применен в оценке того, как метаданные дополняют значимость исследования. Более того, абстрактный характер этих моделей облегчает использование нашего подхода в различных контекстах исследований и выдает результаты для сравнения. Во-вторых, ограничение модели жизненного цикла может также легко наблюдаться из наших результатов и таким образом его не следует игнорировать. Линейные и циклические метафоры, примененные по отношению к таким моделям жизненных циклов, могут предложить только краткую иллюстрацию этой исследовательской процедуры [7], неспособной охватить сценарии, в которых имеющийся процесс не вписывается в линейный подход [17].

На основе этих наблюдений предпринимается следующий шаг проекта - использовать как качественные, так и контекстуализированные методы, такие как интервью и наблюдение, а также количественные методы для достижения цели обсуждавшейся выше исследовательской проблемы. И наблюдательные, и методы контент анализа будут применяться для понимания того, какие активности данных происходят в концептуальном подходе капитала метаданных и каковы их качественные ценности. Хотя эти работы еще не проводились, некоторые приводимые ниже соображения относительно

этих задач обсуждаются в этом разделе на основе работ, представленных в ходе подготовки данной статьи.

Во-первых, полностью признаем напряжение, происходящее между конструктивистским взглядом на науку, сфокусированным на ситуационной научной процедуре [18], которая лежит в основе наблюдаемых исследований, и альтернативным, нормативным взглядом, присущим количественному определению активностей научных метаданных [6]. Предложенный выше подход ориентирован на изучение данного напряжения. Особым аспектом этого напряжения служит тот факт, что не все исследования следуют одной и той же процедуре. Социологи науки демонстрируют, что научные процедуры весьма условны: ученые принимают решения после события с целью изучать ситуации, происходящие в локальных научных контекстах [19].

С учетом этих соображений нам следует полностью признать различия между отдельными научными проектами. Это, как мы считаем, является самым важным подходом для наведения моста между двумя крайне необходимыми потребностями, чтобы понимать как обусловленные эмпирически виды научной деятельности, так и возможности получать всестороннее знание из проектов со многими незначительными различиями.

Во-вторых, хотя наблюдения в исследовательских лабораториях будут самым прямым научным методом решения обсуждаемой в этой работе проблемы, важно также сделать обзор эмпирических исследований по управлению научными данными, которые проводились в различных научных областях, сфокусированных на разных аспектах практики данных [20-23]. Большинство этих исследований, адаптируя методы наблюдения, создали богатое знание относительно того, как работа данных облегчается активностью метаданных в исследовательских процессах, и таким образом могли значительно поддержать наш собственный массив и интерпретацию первичных данных.

Что касается дальнейшей работы по этому исследованию, то должен быть подсчитан капитал метаданных. Надеемся, что наша работа не только будет способствовать лучшему знанию того, как работают метаданные, но и внесет вклад в реальный мир прикладных средств, а также поможет разработке новых методологий для исследования метаданных в целях их использования другими членами научного сообщества.

ЛІТЕРАТУРА

1. *Leonelli S.* Data-centric biology: A philosophical study. — Chicago: The University of Chicago Press, 2016.
2. *Mayernik M. S., Batcheller A. L., Borgman C. L.* How institutional factors influence the creation of scientific metadata// Proceedings of the 2011 iConference. (417–425). — New York, USA: ACM Press, 2011. — <https://doi.org/10.1145/1940761.1940818>.
3. *Willis C., Greenberg J., White H.* Analysis and synthesis of metadata goals for scientific data// Journal of the American Society for Information Science and Technology. — 2012. — Vol. 63, No. 8. — P. 1505-1520. — doi:10.1002/asi.22683
4. *Greenberg J.* Metadata capital: {Raising} awareness, exploring a new concept// Bulletin of the American Society for Information Science and Technology. — 2014. — Vol. 40, No. 4. — P. 30–33. — <https://doi.org/10.1002/bult.2014.1720400412>.
5. *Greenberg J.* Big metadata, smart metadata, and metadata capital: Toward greater synergy between data science and metadata// Journal of Data and Information Science. — 2017. — Vol. 2, No. 3. — P. 19–36.
6. *Greenberg J., Murillo A., Ogletree A., Boyles R., Martin N., Romeo C.* Metadata capital: Automating metadata workflows in the nih viral vector core laboratory// Research Conference on Metadata and Semantics Research, (1–13). — Springer, 2014.
7. *Cox A. M., Tam W. W. T.* A critical analysis of lifecycle models of the research process and research data management// Aslib Journal of Information Management. — 2018. — Vol. 70, No. 2. — P. 142–157.
8. *Gil I. S., Hutchison V., Frame M., Palanisamy G.* Metadata activities in biology// Journal of Library Metadata. — 2010. — Vol. 10, No. (2–3). — P. 99–118.
9. *Musgrave A.* Popper and hypothetico-deductivism// Handbook of the History of Logic. Vol. 10, 205–234. — Elsevier, 2011.
10. *Kitchin R.* Big data, new epistemologies and paradigm shifts // Big Data & Society.— 2014. — Vol. 1, No. 1, 205395171452848. — <https://doi.org/10.1177/2053951714528481>.
11. *Pietsch W.* Causation, probability, and all that: Data science as a novel inductive paradigm// Frontiers in Data Science. — 2017. — P. 329.
12. *Ball A.* Review of data management lifecycle models. — University of Bath: Bath, 2012. — <http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf>.
13. *Vaughan K. T., Hayes B. E., Lerner R. C., McElfresh K. R., Pavlech L., Romito D., ... Morris E. N.* Development of the research lifecycle model for library services// Journal of the Medical Library Association: JMLA. — 2013. — Vol. 101, No. 4. — P. 310–314.— doi:10.3163/1536-5050.101.4.013
14. *Pryor G.* Managing research data. — London, UK: Facet Publishing, 2012. ISBN: 978-1844047562.
15. *Cronston K., Qin J.* A capability maturity model for scientific data management: Evidence from the literature// Proceedings of the American Society for Information Science and Technology. — 2011. — Vol. 48, No. 1. — P. 1–9. — <https://doi.org/10.1002/meet.2011.14504801036>.
16. *Harris F., Lyon F.* Transdisciplinary environmental research: A review of approaches to knowledge coproduction// Nexus Network Think Piece Series. — 2014. — Paper 2.
17. *Li K., Greenberg J., Dunic J.* Data objects and documenting scientific processes: An analysis of data events in biodiversity data papers. — [ArXiv Preprint ArXiv:1903.06215. — 2019].
18. *Latour B.* Science in action: How to follow scientists and engineers through society. — Harvard University Press, 1987.
19. *Cetina K. K.* Laboratory studies: The cultural approach to the study of science/ Handbook of Science and Technology Studies, 140-167.— 1995.
20. *Edwards P. N., Mayernik M. S., Batcheller A. L., Bowker G. C., Borgman C. L.* Science friction: Data, metadata, and collaboration// Soc Stud Sci. — 2011. — Vol. 41, No. 5. — P. 667–690. — <https://doi.org/10.1177/0306312711413314>.
21. *Faniel I. M., Jacobson T. E.* Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data// Computer Supported Cooperative Work(CSCW). — 2010. — P. 355-374.
22. *Jirotko M., Lee C. P., Olson G. M.* Supporting scientific collaboration: Methods, tools and concepts// Computer Supported Cooperative Work (CSCW). — 2013. — Vol. 22, No. (4-6). — P. 667-715.
23. *Zimmerman A.* Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse// International Journal on Digital Libraries. — 2007. — Vol. 7, No. (1-2). —P. 5-16.
24. *Gray J., Liu D. T., Nieto-Santesteban M., Szalay A., DeWitt D. J., Heber, G.* Scientific data management in the coming decade//Acm Sigmod Record. — 2005. — Vol. 34, No. 4. — P. 34–41.
25. *Greenberg J.* Metadata Generation: Processes, People and Tools / Greenberg. — 2003// Bulletin of the American Society for Information Science and Technology. — Wiley Online Library. — (2005, January 31). <https://asistdl.pericles-prod.literatumonline.com/doi/full/10.1002/bult.269>
26. *Harman G.* The metadata lifecycle. — [blog post, 2015]. — <https://bigr.io/the-metadata-lifecycle-gh/>.
27. *Library of Congress.* Metadata standards & applications: Cataloging for the 21st century. — 2008. — <http://www.loc.gov/catworkshop/courses/metadastandards/pdf/MSTraineeManual.pdf>.
28. *Lytas M.D., Sicilia M.-A.* Where is the value in metadata? //International Journal of Metadata, Semantics, and Ontologies. — 2007. — Vol. 2. — P. 235–241.
29. Overview: Research Lifecycle. — (n.d.). — [https://library.ucf.edu/about/departments/scholarly-communication/overview-research-lifecycle/University of Central Florida Libraries](https://library.ucf.edu/about/departments/scholarly-communication/overview-research-lifecycle/University%20of%20Central%20Florida%20Libraries)
30. *Matthews B. M.* Metadata for information management in large-scale science. — 2008. — <https://epubs.stfc.ac.uk/work/50499>.
31. *Mayernik M. S.* Metadata realities for cyberinfrastructure: Data authors as metadata creators. — 2011. SSRN 2042653. — <http://dx.doi.org/10.2139/ssrn.2042653>.
32. *Willoughby C., Bird C. L., Coles S. J., Frey J. G.* Creating context for the experiment record. User-defined metadata: investigations into metadata usage in the LabTrove ELN // Journal of Chemical Information and Modeling. — 2014. — Vol. 54, No. 12. — P. 3268–3283.
33. *Whyte A., Tedds J.* Making the Case for Research Data Management/ DCC Briefing Papers. — Edinburgh: Digital Curation Centre, 2011. — <http://www.dcc.ac.uk/resources/briefing-papers>

Почему необходимы мягкие и нетехнические навыки для успеха в профессии библиотекаря данных*

Маргарет ХЕНДЕРСОН
(Margaret HENDERSON)

Университет г. Сан-Диего, г. Сан-Диего,
шт. Калифорния, США

Для библиотекарей и ученых имеется много курсов, обучающих управлению данными в науке. Эти курсы могут помочь в освоении технических навыков, таких как программирование или статистика, а также знание практики жизненного цикла данных или политики обмена данными, кроме них существуют «мягкие навыки» и нетехнические навыки, востребованные для успешного начала и управления службами данных. Хотя существует множество важных характеристик относительно того, что такое осведомленный библиотекарь данных, навыки библиографа, построения отношений, сотрудничества, умения слушать и слышать и поддержки — это лишь некоторые из них, которые представляют особую важность. Изучение этих навыков поможет библиотекарю данных в его многогранной работе.

ВВЕДЕНИЕ

«Многие современные вопросы, с которыми мои коллеги и я сталкиваемся чаще всего, не обсуждаются в литературе. Эти вопросы включают стратегии построения отношений с занятыми учеными, демонстрацию важности перспектив библиотеки в этой повестке дня, защиту улучшенных практик исследования, а также создание стратегий по вовлечению администрации высшего уровня в решение этих вопросов».

Хезер Коутес [1, с.52]

В предоставлении услуг по обработке научных данных необходимо гораздо большее, чем владение техническими навыками, востребованными в управлении данными. Об-

зор предложений о работе библиотекарем данных показывает, что в них фигурирует множество других навыков. Эклевия и др. [2] провели исследование 104 предложений о работе и обнаружили, что межличностные и коммуникативные навыки, а также способность работать самостоятельно и в коллективе, встречаются в предложениях о работе чаще, чем знание о курировании и управлении данными в науке. Кан и Дю [3] в похожем исследовании предложений о работе выяснили, что в них навыков научной поддержки и критического мышления/подхода к решению проблем было больше, чем знаний об источниках данных или статистики пакетного программного обеспечения. Федерер [4] провел анализ библиотекарей, занимающихся работой, связанной с данными, и обнаружил, что коммуникативные навыки, способность строить отношения с учеными и другие «мягкие» навыки были особенно релевантными в их работе.

«Мягкие» или межличностные навыки, и иные нетехнические навыки весьма важны при организации служб управления научными данными и остаются востребованными в поддержании работы служб. Навыки библиографа, построения отношений, сотрудничества, умения слушать и слышать, а также облегчение доступа к децентрализованным ресурсам, — все они важны для успеха службы данных. Сочетание этих навыков вместе с

* Перевод Henderson M. Why you need soft and non-technical skills for successful data librarianship // Journal of eScience Librarianship. — 2020. Vol. 9, No. 1, с1183. — P. 1-7. — <https://escholarship.umassmed.edu/jeslib/vol9/iss1/2>

Библиотекарь данных — общее понятие профессии библиотекаря, в функции которого входит как предоставление данных (сведений), так и составление каталогов, учет и контроль использования данных, хранящихся на электронных носителях. — прим. ред.

пониманием жизненного цикла знания, метаданных, а также некоторого знания кодирования и статистики, поставит библиотекарей и библиотеки в хорошую позицию по поддержке исследования в учреждениях. Если у вас есть желание работать с данными, вам необходимо быть администратором и предпринимателем, а также библиографом и обучающим библиотекарем, поэтому знание науки о данных – это только одна сторона работы.

ОБСЛУЖИВАНИЕ ЧИТАТЕЛЕЙ

Я всегда настаиваю, что навыки библиографа являются преимуществом библиотекаря. Выяснение того, что действительно необходимо пользователям, составляет большую часть того, чем мы занимаемся. Например, я работала на практикуме с людьми в университете, находящемся в корте университетов по биоинформатике, и у меня был шанс обучиться поиску электронных записей, касающихся здоровья. В то время клиницисты не могли найти свои собственные данные, они их запрашивали через программиста, работающего с базой данных. Он проводил поиск на основе их запроса и отправлял им набор данных, но он (набор данных) возвращался обратно, поскольку не содержал в достаточной степени того, что было необходимо. Программист мог бы попытаться снова осуществить поиск, а набор данных мог быть опрашен обратно с несколькими комментариями и т.д. Хорошее интервью в целях справки сэкономило бы время и сократило обмен электронными сообщениями. Фактически, когда в университетской больнице проводилось исследование нового массива данных по электронным записям о здоровье, запрос предложений продолжился включением данных персонального ассистента, чтобы помочь людям найти правильные данные.

Ученый в области данных Уильям Коэрсэн [5] подтверждает важность навыков библиографа в недавней статье о нетехнических навыках, востребованных в науке о данных. Его первый шаг – постановка правильного вопроса или вопросов. Он считает, что способность задать правильный вопрос скорее подразумевает поиск соответствующих проекту данных, а не погружение в массу данных.

ОТНОШЕНИЯ

Наши навыки библиографа естественным образом ведут к способности развивать хорошие отношения с другими группами в кампусе. Мы задаем вопросы, чтобы выяснить, чем они занимаются и что им необходимо для работы. И мы действуем в целях поддержки исследования, обучения или администрирования в тех областях, в которых они работают. Меня представили администраторам грантов из школ и колледжей моей организации, и затем мы поговорили с некоторыми из них на других мероприятиях в нашем кампусе. Вскоре после этого, со мной на контакт вышли отдельные сотрудники профессорско-преподавательского состава, чтобы помочь в написании планов по управлению данными, они сослались на своего администратора гранта. Путем построения отношений люди чувствуют себя уверенно, делясь своим опытом.

Поскольку построение отношений было неотъемлемой частью того, что я делала как директор по службам научных данных, я осознала, что создание карт отношений – превосходный способ показать прогресс при оценке служб научных данных. Эти карты могут количественно выражать силу отношений с различной протяженностью линий или цветов, чтобы показать типы контактов, встреч, сообщений электронной почты, консультаций и общее количество каждого типа. Рисунок отражает состояние отношений в первый год работы библиотечной службы по предоставлению данных с охватом спонсируемых программ и управления в форме электронных сообщений одному или двум людям без дальнейших взаимодействий. Но информационные технологии учреждения, научно-исследовательский офис и исследователь представлены более темными линиями с двумя стрелками, показывающими взаимодействие с помощью сообщений электронной почты и встреч с множеством людей. С помощью должного охвата и, может быть, анимации — при содействии программы отображения отношений, такой как NodeXL — можно показать изменения в отношениях со временем, придавая значение росту и силе вашей службы.



Рис. Пример карты отношений

СОТРУДНИЧЕСТВО

Библиотекари данных нуждаются в сотрудничестве с людьми как внутри, так и вне стен библиотеки. Более опытные в предметной области, знании статистики и т.д. коллеги могут помочь через проведение семинаров. Ученые, школы и отделения, а также другие организации, обладающие ресурсами, выходящими за рамки библиотечного бюджета, часто имеют основные службы для сотрудников своей организации. Я сотрудничала со службой статистики и анализа отделения статистической науки и исследования операций с целью проведения научного тренинга. Сотрудники службы представили ведущего для тренинга, а я организовывала классы в зале библиотеки, вела регистрацию и обеспечивала питание. Совместно со спонсируемыми программами университета я помогла организовать семинары по составлению грантов, которые включали сессию, по написанию планов по управлению данными для грантов, ее руководителем была я. Также я сотрудничала с библиотекарями данных из других организаций, чтобы иметь информацию по тренингу, используя удаленные образовательные ресурсы в режиме реального времени для обмена опытом и преподавательскими обязанностями.

Перед встречей с потенциальным партнером следует подумать о том, что вы хотите в результате получить. Раскрывая личности перед людьми или группами, в которых вы заинтересованы, и используя этих людей для определения, как подойти к данному лицу или группе, наилучшим способом было бы убедиться, что у вас есть, что сказать перед встречей. У Брайаны Маршал [6] и участников симпозиума the Midwest Data Librarians Symposium (2015) родилась идея создать длинный список потенциальных коллег или партнеров и подготовить несколько лиц и краткий рассказ о концепции идеи, которые помогут вам подумать, какой подход выбрать по отношению к людям и группам в кампусе. Бетси Хамфрис [7] рекомендует вам озвучить вопросы и идеи для совместных действий при первой встрече с группой и найти что-то конкретное и приемлемое для начала сотрудничества.

УМЕНИЕ СЛУШАТЬ И СЛЫШАТЬ

Умение выслушать коллег – основа построения отношений и усиления сотрудничества. С целью понять, что нужно ученому, необходимо выслушать сообщение целиком, проверить свое понимание и задать вопросы [8]. Вам необходимо избегать появления предположений. Когда я работала со студентом старших курсов по вопросу развития формы записи данных в рамках одного эксперимента, я его выслушала один раз в ходе беседы, таким образом мне стало понятно, чем он занимается. Затем повторив процедуру беседы, я смогла задать вопросы, уточняющие этапы эксперимента, типы собираемых данных и задействованные измерения. Пару раз я допустила определенное измерение объектов изучения, но поняла, что когда вводила в созданную нами форму информацию, то вместе со студентом пересмотрела свое понимание, и он мог убедиться в правильности моих измерений.

Умение слушать и слышать – самая трудновыполнимая вещь, особенно когда вы располагаете некоторыми службами данных на месте. Можно брать интервью у ученого или просто обменяться мнениями с членом профессор-

ско-преподавательского состава при встрече в холле. Можно услышать жалобы на то, что они не могут получить в университете или библиотеке, и в каком объеме им нужна помощь со средством DMP. Ваше желание их поправить и сказать о доступности средства DMP (<https://dmptool.org/>), а также то, что информационные технологии имеют хранилище, а у университета есть своя политика – все это не поможет. Как писала Аманда Рейнхарт [9], применение техник активного умения слушать и слышать может принести успех в подобных ситуациях. Позвольте им закончить жаловаться, проявите сочувствие, а затем попытайтесь перефразировать сказанное ими, чтобы они узнали, что их поняли. Потом можно предложить имеющиеся у вас решения, а также свою помощь. Если будете придерживаться оборонительной позиции и прерывать их, никогда не узнаете о их потребностях, а без проявления симпатии с вашей стороны они останутся разочарованными. Помните, люди действительно не замечают, что предлагает библиотека до тех пор, пока им реально нужно что-то сделать, поэтому поддерживайте маркетинговые действия и проявляйте сочувствие по отношению к занятым ученым и студентам.

ПОМОЩЬ

В течение более 30 с лишним лет проведенных в библиотеке, одним из дел, регулярно выполняемых мной за справочным столом при читальном зале, было оказание помощи людям в нахождении того, в чем они нуждались, но, что отсутствовало среди ресурсов библиотеки. Помощь студентам в поисках обучающих услуг или поиск для профессорско-преподавательского состава требуемого оборудования, или помощь посетителю кампуса найти нужное здание – все это составляющие процесса облегчения доступа к услугам. Везде, где я работала, первое, что делала, было выяснение того, чем можно помочь другим отделам в библиотеке, и затем, чем могут помочь мне конкретные отделения организации, а также кому я могу помочь за справочным столом или с помощью консультаций. Я также пытаюсь установить отношения с людьми, работающими в других областях, таким образом я легко могу к ним обратиться. Однажды я работала с исследовательской группой отдела по чрезвычайным ситуациям, которой требовалась помощь по проекту массива больших данных, охватывающих множество электронных таблиц программы Excel. Они не думали, что в этом поможет программа REDCap, хотя я была почти полностью уверена, что она поможет. Но я в этом не эксперт, поэтому соединила их с администраторами REDCap, которые смогли зафиксировать вещи в более простом формате, чем первоначальный Excel. Хотя я сделала немного, но я была рада оказать людям хорошую услугу.

В анализе обзора управления данными в науке Гобен и Гриффин [10 с. 319] обнаружили, что хранение, обмен и долговременный доступ к данным вызывали наибольшее беспокойство ученых, а некоторые библиотеки имеют возможность предоставлять технические средства для оказания таких услуг. Поэтому нам необходимо развивать отношения с институциональными вычислительными службами, чтобы знать, куда направлять ученых. Как директор служб научных данных, я встречалась с менеджерами разных университетских и школьных служб информационных технологий, чтобы узнать, ка-

кого рода сети сотрудничества, а также резервное и долговременное хранение доступны ученым. Хотя некоторые службы имели связь только с соответствующим факультетом, другие являлись основными службами для всех. Имея на руках список ресурсов и контактов, я могла знакомить ученых с доступными вычислительными возможностями.

ЗАКЛЮЧЕНИЕ

Эти пять видов навыков служат лишь началом необходимых для библиотекаря данных нетехнических навыков. В отчете [11] Федерер и др. включили навыки обучения в течение всей жизни, такие как гибкость и приспособляемость, разработка программ и услуг, готовность столкнуться с неудачей, а также другие традиционные библиотечные навыки, такие как оценка и определение качества. Все они лежат в основе развития действенных служб научных данных.

Я не поддерживаю игнорирование науки о данных и технических навыках. Но полагаю, если отбросить часть нашей профессии, дифференцирующей нас от науки о данных, то мы недооцениваем себя. Мы уподобляемся людям, считающим профессию библиотекаря несостоятельной, потому что она, как правило, является ориентированной на женщин областью, а вычислительную науку и науку о данных – прерогативой мужчин, поэтому и очень ценной [12]. Наше сочетание мягких навыков и вычислительных с навыками науки о данных делает нас более полезными, чем программист или статистик, которые не могут задавать правильные вопросы, чтобы выяснить, что необходимо ученому. С нашим идеалистическим взглядом на весь процесс исследования мы можем оказывать большую помощь, чем администратор гранта, который беспокоится только о соответствии политике. Мы знаем, как работать со всеми группами, вовлеченными в управление научными данными, и объединяем их в службу, непосредственно удовлетворяющую потребности ученых.

ЛИТЕРАТУРА

1. Coates H. L. Building data services from the ground up: Strategies and resources// Journal of eScience Librarianship. — 2014. — Vol. 3, No. 1, e1063. — <http://dx.doi.org/10.7191/jeslib.2014.1063>
2. Eclavia M. R., Fredeluces J. C.L.T., Maestro R. S., Eclavia C. L. Jr. What makes a data librarian? An analysis of job descriptions and specifications for data librarians //

Qualitative And Quantitative Methods In Libraries. — 2019. — Vol. 8, No. 3. — P. 273-290. — <http://qqml-journal.net/index.php/qqml/article/view/541>

3. Khan H. R., Du Y. What is a Data Librarian?: A content analysis of job advertisements for data librarians in the United States academic libraries// IFLA WLIC 2018 – Transform libraries, Transform societies in session 139 – Education and Training, Science and Technology, and Continuing Professional Development and Workplace Learning, Kuala Lumpur, Malaysia, 2018. — <http://library.ifla.org/id/eprint/2255>

4. Federer L. Defining data librarianship: A survey of competencies, skills, and training// Journal of the Medical Library Association. — 2018. — Vol. 106, No. 3. — P. 294–303. — <https://doi.org/10.5195/jmla.2018.306>

5. Koehrsen W. How to get the right data? Try asking for it//Medium: Towards Data Science, July 26, 2018. — <https://towardsdatascience.com/how-to-get-the-right-data-why-not-ask-for-it-d26ced1bbd46>

6. Marshall B., MDLS15 participants. Outline of Partnerships Session: Topic 3: Forming Partnerships on Campus// Midwest Data Librarians Symposium, University of Wisconsin, Milwaukee. — UWM Digital Commons, 2015. — <https://dc.uwm.edu/mdls/2015/partnerships/8>

7. Humphreys B. L. How to earn a reputation as a great Partner//Journal of the Medical Library Association. — 2018. — Vol. 106, No.4. — P. 521-528. — <https://doi.org/10.5195/jmla.2018.504>

8. Ross C. S., Dendney P. Communicating professionally: A how-to-do-it. Manual for Library Applications. — New York: Neal-Schuman, 1989.

9. Rinehart A. K. Getting emotional about data: The soft side of data management services// College & Research Libraries News. — 2015. — Vol. 76, No. 8. — P. 437-440. — <https://doi.org/10.5860/crln.76.8.9364>

10. Goben A., Griffin T. In aggregate: Trends, needs, and opportunities from Research Data Management Surveys// College & Research Libraries. — 2019. — Vol. 80, No. 7. — P. 903-924. — <https://doi.org/10.5860/crl.80.7.903>

11. Federer L., Sarah C. C., Zaringhalam M. Developing the librarian workforce for Data Science and Open Science // OSF Preprints January 16.— 2020. — <https://doi.org/10.31219/osf.io/uycax>

12. Harris R. M. Librarianship: The erosion of a woman's profession. — Ablex: Norwood, NJ, 1992.

АВТОРСКИЙ УКАЗАТЕЛЬ к МФИ, 2020, т. 45

	№	Стр.		№	Стр.		№	Стр.
Агарвал Н. К.	2	3	Кобзева Л. В.	4	3	Реес-Лильо Д.	3	3
Ай П.	4	5	Коркимяки Т.	2	14	Сантеликес-Верчес К.	3	3
Арельяно-Рохас П.	3	3	Ли К.	4	27	Сюй Ц.	4	5
Бу И.	4	5	Ливи К.	4	19	Тели С.	3	16
Бьёрк Б.-К.	2	14	Майерник М.	1	22	Томпсон К. М.	3	3
	2	25	Миллайон А. Дж.	1	3	Хадсон-Витале С.	1	3
Гаррисон К.	3	3	Мулесон Сэнди Х.	1	3	Хендерсон М.	4	33
Дин И.	4	5	Мюррей Д. С.	4	5	Хенкок Д. Т.	4	19
Добрески Б.	1	13	Нааман М.	4	19	Шэнь Ц.	4	5
Дугта Б.	3	16	Овченкова Е. А.	4	3	Ян Ф.	4	5
Ислам М. А.	2	3	Паскуа С.	4	27			

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

Ответственный за выпуск *Л. В. Кобзева*

Компьютерная верстка *М. А. Филимонова*

ИД № 04689 от 28.04.2001 г.

Подписано в печать 16.12.2020 г.

Бумага офсетная. Формат 60x84 1/8. Гарн. литер. Печать цифровая

Усл. печ. л 5,00 Уч.-изд. л. 5,38 Тираж 33 экз.

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

Тел. (499) 155-44-95

УВАЖАЕМЫЕ КОЛЛЕГИ!

ВИНИТИ РАН предлагает Вашему вниманию Реферативный Журнал в электронной форме

РЖ в электронной форме (ЭлРЖ) выпускается по всем разделам естественных, технических и точных наук.

Каждый номер ЭлРЖ является полным аналогом печатного номера РЖ по составу описаний документов, их оформлению и расположению. Он сопровождается оглавлением, указателями.

ЭлРЖ представляет собой информационную систему, снабженную поисковым аппаратом и позволяющую пользователю на персональном компьютере:

- читать номер РЖ, последовательно листая рефераты;
- просматривать рефераты отдельных разделов по оглавлению;
- обращаться к рефератам по указателям авторов, источников, ключевых слов;
- проводить поиск документов по словам и словосочетаниям;
- выводить текст описаний документов во внешний файл.

ЭлРЖ в версии Windows Вы можете получить за текущий год с любого номера, а также за предыдущие годы.

Подробную информацию Вы можете получить:

Адрес: 125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефон 499-155-42-85 499-151-78-61

E-mail: Contact@viniti.ru, Feo@viniti.ru

ВНИМАНИЮ ЧИТАТЕЛЕЙ!

ИЗДАНИЕ УДК

УНИВЕРСАЛЬНАЯ ДЕСЯТИЧНАЯ КЛАССИФИКАЦИЯ
АЛФАВИТНО-ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ
в 2-х томах

Алфавитно-предметный указатель (АПУ) к 4-му полному изданию УДК на русском языке:

Том I содержит АПУ от буквы А до Н;

Том II содержит АПУ от буквы М до Я и указатель латинских наименований к классам УДК 56 Палеонтология, 57 Биологические науки, 58 Ботаника, 49 Зоология, 61 Медицинские науки.

АПУ содержит около 100 000 понятий, представленных в полных таблицах УДК.

При его составлении были учтены изменения, опубликованные в Выпусках № 1 – 6 «Изменения и дополнения к УДК»

Для подписки необходимо направить заявку для оформления счета по адресу:

125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефоны: 499 155-42-85, 499 151-78-61

E-mail: feo@viniti.ru

<http://www.udcc.ru>