

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 7

Москва 2013

ОБЩИЙ РАЗДЕЛ

УДК 002:004

И.М. Зацман

Информационно-компьютерная наука: предпосылки становления

Сопоставляются подходы С. Горна и Ю. Шрейдера к интеграции информационной и компьютерной наук в интересах формирования единой области знания «информационно-компьютерная наука». Приводится описание характерных черт этой области знания, которая позиционируется как научная основа создания информационных технологий будущих поколений. Рассматриваются концептуальные и программные документы 7-й Рамочной программы Евросоюза, в которых обосновывается необходимость переосмысления теоретических основ создания новых информационных технологий. Предлагается один из возможных вариантов их переосмысления на основе идей Горна и Шрейдера.

Ключевые слова: информационная наука, компьютерная наука, информатика как информационно-компьютерная наука, информационные технологии будущих поколений

ВВЕДЕНИЕ

Фундаментальное переосмысление теоретических основ создания информационных и конвергентных технологий следующих поколений в XXI в. стало одной из ключевых предпосылок становления информационно-компьютерной науки как общенаучной области знания. Анализ опыта подготовки специалистов для

создания новых технологий, проведенный Марком Сниром, показал, что программа их обучения должна интегрировать дисциплины информационной и компьютерной наук [1]. Во второй половине XX в. необходимость их интеграции предвидели Сол Горн [2-7] и Юлий Шрейдер [8, 9]. Идеи Горна и Шрейдера в наше время оказались весьма востребованными. Об этом свидетельствуют рассматриваемые в настоящей статье

концептуальные и программные документы по созданию информационно-коммуникационных технологий следующих поколений [10-15], а также проблематика конвергентных технологий, включенная в 2011 г. в конкурс Российского фонда фундаментальных исследований (РФФИ) по ориентированным фундаментальным исследованиям. В настоящей статье предлагается один из возможных вариантов развития идей Горна и Шрейдера.

В апреле 2013 г. исполнилось 50 лет, как была опубликована статья Сола Горна «Информационно-компьютерные науки как новая фундаментальная область знаний» [2]. В течение 20 лет (1963-1983гг.) Горн опубликовал серию статей, в которых обосновывал необходимость преподавания комплекса дисциплин, относящихся к новой области знания, интегрирующей информационную и компьютерную науки. В 1963 г. Горн использовал множественное число для названия этой области знания - информационно-компьютерные науки. Единственное число появилось в 1982 г. в названии его статьи «Информатика как информационно-компьютерная наука: ее идеология, методология и социальные аспекты», т.е. через 20 лет после публикации первой работы из этой серии [6]. Пятьдесят лет назад Горн рассматривал результат интеграции информационной и компьютерной наук как новую фундаментальную науку, что нашло отражение в названии его первой работы – «a new basic discipline» [2].

Другой подход к определению информатики как информационно-компьютерной науки и единой области знания был предложен Ю.А. Шрейдером [8, 9]. Его результаты были получены независимо от Горна, но идейно их подходы во многом весьма близки. В статье Шрейдера «Информация и знание» говорится, что не существует двух информатик (информационная наука и компьютерная наука), а есть два облика информатики. Первый из них (информационная наука) дополнительно нагружен представлениями о традиционном информационном обслуживании специалистов-ученых и инженеров в области их профессиональных интересов. Второй облик (компьютерная наука) неправомерно искажен чисто программистскими проблемами, не специфичными для информатики. Специфические же проблемы информатики оказываются там, где возникают задачи информационного представления знаний в форме, удобной для обработки, передачи и творческого реконструирования знаний в результате усилий пользователя [9, с. 51].

Идеи Горна и Шрейдера, краткому обзору которых посвящен второй раздел настоящей статьи, были практически полностью преданы забвению. Более подробное описание их подходов можно найти в работах [16, 17], материалы которых использованы в нашей статье. Идеи Горна и Шрейдера стали возрождаться в начале XXI в. Аналогичные подходы стали заново формулироваться в редуцированной форме и предлагаться как теоретическая основа для разработки новых информационных и конвергентных технологий. При этом работы основоположников фундаментальных подходов к описанию информатики как информационно-компьютерной науки, как правило,

не упоминаются. Третий раздел нашей статьи содержит пример редуцированного подхода к интеграции компьютерной и информационной наук.

В четвертом разделе статьи рассматривается актуальная проблематика целенаправленной генерации новых систем знаний. Анализируются результаты моделирования и экспериментов по созданию информационных технологий, обеспечивающих генерацию новых систем экспертных знаний. Эти результаты были получены на основе подходов Горна и Шрейдера. Модели, разработанные на основе их идей, дали возможность описать процессы генерации и эволюции экспертных знаний в динамике их формирования [18].

ГОРН И ШРЕЙДЕР ОБ ИНФОРМАТИКЕ КАК ИНФОРМАЦИОННО-КОМПЬЮТЕРНОЙ НАУКЕ

Уже в 1963 г. Горн перечислил ряд вопросов, которые относятся к информационно-компьютерной науке: «Примерами основных вопросов исследования в этой области могут быть системы программирования, проектирование компьютерных систем, искусственный интеллект, информационный поиск и т.д. Вероятностная теория информации Шеннона определенно принадлежит к этой области знания, но помимо нее существует теория компьютерной информации и процессов ее обработки, которую также необходимо включить в предметную область этой науки. Одним из центральных вопросов этой новой науки, скорее всего, станет анализ и синтез машинных языков и процессоров их обработки» [2, с. 150].

После перечисления направлений исследований, Горн рассматривает информационно-компьютерную науку как учебную дисциплину. Он сопоставляет новую область знаний с другими дисциплинами с точки зрения образовательного процесса и формулирует следующие вопросы. Каким образом студент может узнать, относится ли сфера его интересов именно к этой новой области знания, а не к одной из уже устоявшихся дисциплин? Какое ему необходимо образование для того, чтобы углубиться в эту новую область знания? Горн рассматривает вопросы преподавания, отталкиваясь в своих рассуждениях от прагматических аспектов исследований:

«Студент, изучающий численный анализ, в процессе разработки или анализа какого-либо алгоритма позиционирует себя как математик в том случае, если его единственный прагматический интерес заключается в доказательстве существования алгоритма или определения его теоретической точности. Если он позиционирует себя как будущий специалист в области информационно-компьютерной науки, то он рассматривает этот алгоритм с точки зрения его реализации и обработки процессором, т.е. он интересуется эффективностью работы этого алгоритма, временными затратами, распределением памяти и т.д.

Аналогично, студент, изучающий процедуру адаптивного управления, описывающую поведение животного в некоторой ситуации, позиционирует себя как биолог в том случае, если его главной задачей является выяснение того, обладает ли он хорошей

моделью поведения этого животного. Если его интересует проблема искусственного интеллекта, как одного из направлений информационно-компьютерной науки, то он интересуется применимостью этой процедуры независимо от того, является ли она моделью поведения животного или не является.

Студент, занимающийся грамматикой, мыслит себя как лингвист в том случае, если его больше всего интересует, действительно ли изучаемый естественный язык работает так, а не иначе. Однако он думает как будущий ученый в области информационно-компьютерной науки в том случае, если его занимает вопрос, каким образом можно использовать эту грамматику в информационных системах. Лингвист может рассматривать механизм стековой памяти, но с глубиной не более семи уровней из-за ограниченных возможностей локальной памяти человека, но для решения информационно-компьютерных задач такой глубины может оказаться недостаточно» [2, с. 154].

Рассмотрев эти вопросы, Горн приводит перечень дисциплин, которые должны преподаваться студентам, изучающим информационно-компьютерную науку, включая математику, физику, философию, лингвистику, психологию, вычислительную технику и компьютерное программирование. Предложенный им подход к изучению информационно-компьютерной науки уже тогда начал опробоваться в Пенсильванском университете. Через двадцать лет, в 1983 г., когда уже накопился опыт ее преподавания, Горн пишет, что его понимание концепции информационно-компьютерной науки заключается в том, что эта область знания не является ветвью математики, так как она должна соотносить себя с прагматическими вопросами, от которых математика не должна зависеть [7, с. 137].

Статья Горна 1983 г. начинается с фразы: «Позвольте мне, прежде всего, выбрать более короткое название, чем информационно-компьютерная наука. Я выбираю термин «информатика», созвучный французскому *Informatique* и немецкому *Informatik*. Он несет в себе идею информации, а оканчивается так же, как и математика, подразумевая формализованную теорию. Плохо то, что при использовании слова «информатика» теряется компьютерная составляющая в названии, и кроме того, оно не вызывает ассоциаций с какой-либо экспериментальной основой» [7, с. 121].

Используя в 1983 г. термин «информатика», Горн подразумевает под ним именно информационно-компьютерную науку. Прежде чем сделать свой главный вывод о необходимости единой области знания, он обращается к ее истокам:

«Все, что я до сих пор говорил о вычислениях, является ориентированным на практическую деятельность и связано с компьютером. Но сама теория вычислений уже сформировалась и существовала к тому времени, когда появились цифровые компьютеры. /.../ Специалисты в области символической логики уже исследовали логические пределы вычислений; была описана универсальная машина Тьюринга и доказана неразрешимость проблемы останковки; Гедель продемонстрировал пределы формализма при помощи своих теорем о неразрешимости; Черч, Клини и Карри проанализировали вычисления в теории рекурсивных функций и комбинаторной логике; Туэ и

Пост, а в более позднее время Марков, рассмотрели вычисления с синтаксической точки зрения. /.../ когда появились компьютеры, обсуждение лингвистических вопросов Ноамом Хомским происходило в компьютерных терминах /.../ В результате этих новых разработок появились описания лингвистических аспектов программирования, математическая теория автоматов и формальные языки. Эти результаты, в свою очередь, повлияли на разработки языков программирования и проектирование вычислительных машин. /.../ Теперь под информатикой мы понимаем нечто, связанное с синтезом и анализом символьных выражений, а также синтез и анализ процессоров, которые интерпретируют, транслируют и обрабатывают такие выражения. Если говорить более прозаично, то информатика занимается изучением, проектированием и использованием структур данных и их обработкой» [7, с. 131].

Главный вывод Горна, которым он завершает статью 1983 г., состоит в следующем: «...нам не следует отделять компьютерную науку от информационной науки, а следует пытаться отстаивать единую область знания – информатику. Любая попытка поощрить разделение /.../ повлечет за собой отделение практики от знаний /.../ Такое разделение будет причиной прекращения деятельного кипения, которое поддерживается сплавом знаний и практической деятельности ...» [7, с. 139-140].

В 1988 г., т.е. через пять лет после публикации 1983 г., которая завершает серию статей Горна, независимо от него идейно близкий подход был предложен Ю.А. Шрейдером. В своей статье [9] он формулирует ряд положений парадигмы информатики:

«...информация есть общественное достояние, она в принципе социальна, в то время как знание, вообще говоря, соотносено с конкретной личностью, с тем, кто им владеет и непосредственно пользуется. /.../ Информация должна пройти через "когнитивный экран" тех, для кого она представляет ценность. Так возникает необходимость считаться не только с существованием мира объективированного социализированного знания, т.е. информации как превращенной формы знания, но и с феноменом личностного знания. /.../ Тождественность информации и знания при этом исключается, но информация как превращенная форма знания сохраняет следы своего происхождения. /.../ Наиболее принципиальные вопросы информатики всегда возникали на стыке информации и знания, там, где речь шла о превращении одного в другое». Далее Шрейдер пишет о пропасти, разделяющей *информацию и знания как сущности разной природы* (выделено мной – ИЗ) [9, с. 50-51].

В этой же статье Шрейдер сопоставляет свой подход к определению понятий «информация» и «знания» с подходом Б. Брукса [19]. Он описал суть подхода Брукса следующим образом:

«Фактически на позиции признания *тождества информации и знания* (выделено мной. – И.З.) как сущностей стоит Б. Брукс, построивший основания традиционной информатики (информационной науки) на базе концепции "третьего мира" [Карла Поппера [20]]. Он определяет информацию как приращение знаний» [9, с. 49].

В этой же работе Ю.А. Шрейдер написал о том, как исторически сформировалась концепция А.И. Михайлова, А.И. Черного и Р.С. Гиляревского об информатике как науке, изучающей информацию, связанную со знаниями, а не с управляющими сигналами [21].

Через пятьдесят лет после публикации первой работы Горна [2] наблюдается следующий парадокс. С одной стороны, работы Горна, который предсказал необходимость становления новой области знания, редко упоминаются, что показано в таблице, которая с названиями статей Горна и числом их цитирований построена на основе данных Web of Science (<http://apps.webofknowledge.com>).

С другой стороны, в наши дни предлагаются редуцированные решения по интеграции информационной и компьютерной наук, являющиеся частными случаями подходов Горна и Шрейдера. Ссылки на их работы, как правило, отсутствуют, что говорит о забвении их идей. Пример предлагаемых редуцированных подходов и описание новых проблем, которые могут быть отнесены к новой области знания, даны в следующих двух разделах настоящей статьи.

НОВЫЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ

В наши дни появляются новые направления исследований, которые могут оказать существенное влияние на процессы интеграции информационной и компьютерной наук. Например, в начале XXI в. стало формироваться комплексное направление исследований, получившее название «когнитивная информатика» [22-25]. Ее предметная область включает актуальные проблемы на стыке когнитивных, компьютерной и информационной наук. Если включить эти проблемы в предметную область информационно-компьютерной науки, то это окажет существенное

влияние на результаты построения ее теоретических оснований. Поэтому сначала необходимо уточнить границы новой области знания. В процессе уточнения этих границ были использованы следующие концептуальные и программные документы:

- аналитический отчет по вопросам обеспечения конкурентоспособности США в XXI в., подготовленный Консультативным комитетом по информационным технологиям при Президенте США [26] (информацию о содержании этого отчета на русском языке можно найти в работе [27]);
- описание приоритетных направлений исследований и разработок по информационно-коммуникационным технологиям (ИКТ) 7-й Рамочной программы ЕС, принятой на период 2007-2013 гг.;
- описание ориентированных фундаментальных исследований конкурса РФФИ 2011 г. в части разработки фундаментальных основ создания конвергентных технологий.

Аналитический отчет

Основная идея аналитического отчета заключается в том, что конкурентоспособность США в XXI в. будет во многом определяться широкомасштабным применением новых информационных технологий во всех областях знания и сферах практической деятельности [26]. Согласно отчету для создания этих технологий требуется дальнейшее развитие вычислительной отрасли в целом и информационно-компьютерной науки, в частности. В сфере научного познания информационные технологии представлены в отчете как составляющая триады «научная теория – научный эксперимент – информационные технологии, обеспечивающие моделирование и проведение экспериментов».

Таблица

Название статьи Горна	Год публикации статьи	Цитируемость по Web of Science (по данным на март 2012 г.)
The computer and information sciences: a new basic discipline	1963	16
The individual and political life of information systems	1965	2
Computer and information sciences and the community of disciplines	1967	53
The identification of the computer and information sciences: their fundamental semiotic concepts and relationships	1968	3
Informatics (computer and information science): its ideology, methodology, and sociology	1982 (первая журнальная публикация), 1983 (вторая публикация)	12 (1 цитирование первой публикации, 11 - второй)

Предлагаемое в отчете определение вычислительной отрасли имеет следующий вид [26, с. 10]: «это быстро растущая мультидисциплинарная предметная область, в которой используются возможности передового компьютеринга (advanced computing) для понимания и решения сложных проблем и которая включает три компонента:

- алгоритмы (численные и нечисленные), программное обеспечение, разработанное для решения проблем естественных, гуманитарных и инженерных наук;

- *информационно-компьютерная наука* (выделено мной – ИЗ), которая разрабатывает и оптимизирует современные аппаратные, программные и сетевые средства, а также компоненты управления данными, которые необходимы для решения вычислительно сложных проблем;

- вычислительная инфраструктура, которая поддерживает решение научных и инженерных проблем, а также развитие информационно-компьютерной науки».

В этом определении используется словосочетание «информационно-компьютерная наука», с помощью которого идея интеграции была отражена С. Горном в серии его статей. В аналитическом отчете по вопросам обеспечения конкурентоспособности США информационно-компьютерная наука понимается как один из трех компонентов вычислительной отрасли. В отчете предметная область этой науки существенно редуцирована по сравнению с идеями Горна и Шрейдера. Ее предназначение определяется необходимостью решения вычислительно сложных проблем. Перефразируя Шрейдера, можно сказать, что в этом подходе облик информационно-компьютерной науки редуцирован из-за ориентации на вычислительно сложные проблемы.

Приоритетные направления исследований и разработок по информационно-коммуникационным технологиям

В документах 7-й Рамочной программы ЕС сформулировано восемь приоритетных направлений исследований и разработок, включая направление «Перспективные ИКТ» [10-15]. В этих документах констатируется необходимость фундаментального переосмысления теоретических оснований создания ИКТ следующих поколений. При этом термин «информационно-компьютерная наука» в явном виде не используется.

Цели проектов, финансируемых в рамках направления «Перспективные ИКТ», сформулированы следующим образом: «Своевременная идентификация и обоснование новых тематических направлений исследований и разработок, которые имеют большой научно-технический потенциал и могут стать основой для разработки ИКТ следующих поколений. Эти проекты должны включать междисциплинарные исследования новых и альтернативных подходов к разработке ИКТ будущего и быть нацеленными на *фундаментальное переосмысление системы теоретических, прикладных, методологических и технологических принципов, подходов и понятий* (выделено мной – И.З.), используемых сегодня в сфере ИКТ» [12, с. 54].

Цитируемый документ содержит описание нескольких конкурсных тем направления «Перспективные ИКТ», включая тему «ИКТ долговременного применения». В описании этой темы говорится о необходимости разработки новых подходов к генерации, представлению и сохранению личностных знаний человека, их интеграции в цифровой электронной среде и глобальному использованию, а также формулируются актуальные исследовательские проблемы [12, с. 57-63].

Приведем формулировки двух проблем, иллюстрирующих актуальность этих направлений исследований [12, с. 62-63]:

1) *разработать теоретические и прикладные основы создания долговечных систем*, способных к эволюции при минимизации затрат на их развитие в условиях многократной смены поколений программно-аппаратных и сетевых средств и/или форматов данных. Долговечные системы должны быть способными к сохранению своей первоначальной социально-значимой функциональности в течение долгого периода времени, а также должны быть способны изменяться в случае необходимости. Методы сохранения и изменения функциональных возможностей должны быть компьютерно-независимыми и обеспечивать устойчивую эволюцию долговечных систем;

2) *разработать новые подходы к представлению и сохранению знания*, ориентированные на долговременный и безотказный к ним доступ в условиях локальной генерации отдельных «квантов» знания, их интеграции, а также глобального использования систем представления и сохранения знания с учетом контекста и временной эволюции систем. Должна быть обеспечена долговременная устойчивость систем представления и сохранения знания в условиях многообразия их использования и эволюции семантики во времени.

Отметим, что в формулировке второй проблемы ключевыми терминами являются «знания» и «кванты знания», а термин «информация» отсутствует. Это отражает актуальность и прикладную востребованность в наши дни исследований процессов генерации и разработки методов компьютерного представления новых знаний в динамике их формирования. С точки зрения Шрейдера, эта проблематика относится к информатике как информационно-компьютерной науке, специфические проблемы которой оказываются там, где возникают задачи информационного представления знаний /.../ и творческого реконструирования знаний в результате усилий пользователя [9, с. 51].

Таким образом, в документах 7-й Рамочной программы ЕС сформулирована потребность в разработке теоретических оснований создания ИКТ будущих поколений, учитывающих следующие обстоятельства и ограничения:

- локальная генерация новых «квантов» знания, их интеграция и компьютерное представление новых знаний в динамике их формирования;

- глобальное использование систем представления и сохранения новых знаний с учетом контекста и временной эволюции [компьютерных] систем;

- эволюция семантики [форм и структур представления знаний человека] во времени.

Сформулируем ряд вопросов для того, чтобы детализировать перечисленные обстоятельства генерации новых знаний. На эти вопросы нужно ответить тем или иным образом в процессе построения теоретических оснований информационно-компьютерной науки. Ответы на эти вопросы необходимы для сравнительного анализа разных вариантов их построения.

Во-первых, допускают ли предлагаемые теоретические основания категоризацию ментальных знаний по степени их конвенциональности? Есть ли категория личностных ментальных знаний, которые их автором ни с кем не согласованы? Если ответ «да», то в результате каких процессов происходит преобразование личностных несогласованных ментальных знаний человека в коллективные (согласованные) и конвенциональные знания?

Во-вторых, определяются ли в явном виде источники и цели генерации новых и пополнения существующих систем знаний?

В-третьих, можно ли влиять на процессы генерации и эволюции систем знаний (сделать их целенаправленными)?

В-четвертых, допускают ли предлагаемые теоретические основания разделение ментальных знаний человека на имплицитные, которые не выражены в явном виде, и эксплицитованные знания?

Все четыре перечисленных вопроса являются ключевыми для описания теоретических оснований создания ИКТ следующих поколений, в частности, и парадигмы информационно-компьютерной науки, в целом.

Ориентированные фундаментальные исследования

В марте 2011 г. Российский фонд фундаментальных исследований (РФФИ) провел конкурс ориентированных фундаментальных исследований по 23 междисциплинарным темам, включая тему «Фундаментальные основы конвергентных технологий» [28]. По всем 23 темам поступило 1137 заявок, из которых 468 получили гранты на выполнение проектов. По теме «Фундаментальные основы конвергентных технологий» поступило чуть более 40 заявок, из которых почти половина получила гранты на выполнение проектов, включая следующие [29, 30]:

- Исследование и разработка нейроморфных средств и сложных антропоморфных технических систем (модели восприятия информации, накопления знаний о среде и поведении путем обучения в реальном времени).

- Исследование ритмических кодов мыслительной деятельности и создание на этой основе модели когнитивного пространства человека и интерфейса мозг-компьютер высокого уровня (разработка технологии определения типа совершаемых в уме мысленных действий по рисунку электроэнцефалограммы; распределение ритмов мозга для отдельных видов мышления: пространственного, образного, вербально-логического и смешанных форм; количественная оценка расстояния между разными видами мышле-

ния с определением их координат на модели когнитивного пространства).

- Когнитивные основы конвергентной технологии распознавания речи (построение обучающего множества паттернов для технологии распознавания речи).

Названия и краткие аннотации позволяют получить первое представление о проблематике этих проектов по теме «Фундаментальные основы конвергентных технологий». В них рассматриваются вопросы взаимодополняемости био-, нейро-, когнитивных технологий и их интеграции с ИКТ. В рамках этого конкурса РФФИ конвергенция технологий разных видов позиционируется как фактор формирования новой технологической базы цивилизации, основанной на воспроизведении систем и процессов живой природы в виде технических систем и технологических процессов [28-30], в том числе в виде информационно-компьютерных систем и технологий.

В XXI в. конвергенция технологий разных видов позиционируется как актуальная проблематика исследований. В частности, в ежегодных программных документах под названием «NSF Investments and Strategic Goals» Национального научного фонда (ННФ) США идея конвергенции технологий появилась в 2003 г. Важно отметить, что одно из приоритетных направлений исследований, финансируемых ННФ, имеет название «*Computer and Information Science and Engineering*» [31].

Исследования по конвергенции технологий направлены на решение крупных социальных задач, что отражено в самих названиях исследовательских программ, проектов, научных мероприятий и отчетных материалов. Так, отчет 2003 г. по проекту ННФ № CTS-0128860, подготовленный ведущими специалистами США по этой тематике, называется «Конвергенция технологий для улучшения работоспособности человека» [32], а аналитический отчет 2004 г., подготовленный ведущими специалистами Европейского союза – «Конвергенция технологий: построение будущего для европейских сообществ» [33].

РФФИ финансирует фундаментальные исследования в области интеграции ИКТ с био-, нейро- и когнитивными технологиями на конкурсной основе с 2011 г. Далее совокупность последних трех видов технологий кратко будем обозначать как когнитивные технологии. Создание фундаментальных основ их интеграции с ИКТ является сегодня актуальной тематикой. Как уже отмечалось, если эту тематику включить в предметную область информационно-компьютерной науки, то это окажет существенное влияние на построение теоретических оснований этой науки.

ЦЕЛЕНАПРАВЛЕННАЯ ГЕНЕРАЦИЯ СИСТЕМ ЗНАНИЙ

Исследования процессов генерации новых знаний традиционно относились к экономике, в рамках которой использовались и до сих пор используются качественные модели разных технологий «мозгового штурма» [34-40]. Необходимость в количественных моделях и функциональном симбиозе человека и компьютерной системы в интересах целенаправлен-

ной генерации новых знаний стали теми факторами, которые сделали доминирующей информационно-компьютерную компоненту этих исследований.

Роль и значение этой компоненты рассматривались приглашенными экспертами при подготовке 7-й Рамочной программы ЕС в рамках семинара «Knowledge Anywhere Anytime: “The Social Life of Knowledge”». Этот семинар состоялся 29–30 апреля 2004 г. в Брюсселе [15]. Материалы этого семинара использовались при формировании программ работ по ИКТ 7-й Рамочной программы ЕС [12–14].

В этих материалах отмечается, что исследование процессов, разработка методов и моделей формирования новых и развития существующих методов и моделей структуризации и представления знания является актуальной проблематикой в XXI в. Участники семинара определили следующие четыре актуальных направления этих исследований [15].

1. Формирование научного понимания того, как знание появляется, каким образом на этот процесс и его результаты влияет совместная деятельность, как формируется конвенциональное знание. Одна из задач этого направления заключается в том, чтобы создать методы и средства описания различий в личностном понимании участниками совместной деятельности смысла одного и того же текста, графика, диаграммы, изображения и т.д.

2. Исследование многообразия форм представления одних и тех же концептов как «квантов» ментального знания. Кроме форм представления конвенциональных и стабильных концептов, предметом исследования являются формы представления личностных и коллективных (согласованных) концептов. В рамках этого направления предполагается выполнение исследований процессов формирования конвенциональных концептов на основе личностных и коллективных концептов.

3. Создание нового поколения интеллектуальных информационных систем, которые должны обеспечить семантическую интероперабельность в процессе совместной работы пользователей этих систем.

В рамках третьего направления, помимо создания методов и средств поддержки семантической интероперабельности в интеллектуальных информационных системах, планируется исследовать вопросы выявления и экспликации стадий генерации и эволюции знания, представленного в виде классификационных систем, словарей, тезаурусов, онтологий и других моделей структуризации и представления знания. При этом не предполагается, что пользователи заранее будут владеть согласованной между ними системой терминов и единым пониманием принципов структуризации и представления знания.

Степень новизны интеллектуальных систем, поддерживающих процессы выявления и экспликации стадий генерации и эволюции нового знания, предлагается оценивать, сравнивая их с традиционными системами управления знаниями (Knowledge management systems - KMS), основанными на гипотезе стабильности ментального знания человека [41]. Согласно этой гипотезе, в процессе создания и применения KMS можно не учитывать эволюцию во времени ментального знания, представленного в KMS.

4. Исследование принципиальных возможностей и средств влияния на процессы генерации новых или эволюции существующих систем знаний в процессе совместной деятельности коллективов специалистов. Наиболее актуальные вопросы этого направления исследований связаны с пространственно-распределенными коллективами специалистов, совместная деятельность которых обеспечивается сетевыми технологиями. В случае решения задач целенаправленной генерации ментального знания предполагается, что имеет место ситуация его неполноты, эта неполнота специфицирована, и перед коллективом специалистов поставлена задача уменьшения степени лакуарности существующих систем знаний.

В рамках этого направления предлагается исследовать, какими видами перспективных ИКТ и до какой степени можно оказывать влияние на процессы формирования новых знаний, отвечающих социально значимым потребностям и необходимым для получения запланированных результатов. По мнению экспертов, приглашенных на семинар «Knowledge Anywhere Anytime: “The Social Life of Knowledge”», возможность оказывать влияние на эти процессы является характерной чертой общества, основанного на знаниях.

Приведенный перечень из четырех направлений исследований говорит о том, что участники этого семинара придают ИКТ ключевую роль в решении проблем формирования новых и развития существующих систем знаний. При этом они существенно расширили границы предметной области представления знаний. Это расширение произошло, в основном, за счет вопросов генерации новых знаний и влияния на процессы генерации и эволюции систем знаний средствами ИКТ. Анализ материалов этого семинара позволяет сделать вывод о том, что предметная область, очерченная приглашенными экспертами, включает, в дополнение к традиционным задачам представления знаний, следующие три актуальных направления исследований [15]:

1) генерация и компьютерное представление в цифровой электронной среде¹ личностных и согласованных концептов как «квантов» ментального знания, формируемых пространственно-распределенными коллективами специалистов;

2) анализ и оценивание степени релевантности систем формируемых концептов социальным, экономическим, технологическим и другим общественно значимым потребностям, в интересах удовлетворения которых они формируются;

3) целенаправленное влияние с помощью ИКТ и когнитивных технологий на генерацию и эволюцию систем формируемых концептов, необходимых для удовлетворения общественно значимых потребностей.

¹ Согласно ГОСТ Р 52292-2004, электронная среда – это среда технических устройств (аппаратных средств), функционирующих на основе физических законов и используемых в информационной технологии при обработке, хранении и передаче данных. В нашей статье цифровая электронная среда – это цифровые технические устройства (цифровые аппаратные средства) электронной среды.

В рамках первого направления исследований получены решения двух проблем:

- концептуальное индексирование геоизображений как форм представления личностного ментального знания специалиста о геообъектах [42];
- генерация систем новых экспертных знаний, формируемых пространственно-распределенными коллективами специалистов в интересах обеспечения мониторинга программно-целевой деятельности [18, 43].

В процессе решения этих проблем были разработаны количественные модели процессов генерации и методы компьютерного кодирования личностных и согласованных концептов как структурных элементов систем экспертных знаний. Теоретические основания для создания этих моделей были построены в процессе развития подходов Горна и Шрейдера к интеграции информационной и компьютерной наук.

Теоретические основания интеграции включают: деление предметной области информационно-компьютерной науки на среды; использование двух классов понятий (однородных по своей природе и двуединых); распределение однородных понятий по средам и двуединых понятий по границам между средами; использование двух взаимосвязанных стадий представления концептов в цифровой электронной среде в виде компьютерных кодов [42-44].

Приведем краткое описание перечисленных оснований интеграции информационной и компьютерной наук, используя систему терминов из работы [45]. В эту систему терминов включены два основных класса понятий: однородные по своей природе (знания, знаковая информация, цифровая информация, цифровые данные и компьютерные коды) и двуединые (семиотические знаки, формкоды и семокоды). В начале процесса построения теоретических оснований рассматривались следующие четыре среды (затем число сред было увеличено) и соответствующие им однородные понятия:

- *ментальная среда знаний человека*, которая включает выраженные и невыраженные (имплицитные) знания;
- *социально-коммуникационная среда*, к которой относятся отчужденные от человека сенсорно-воспринимаемые формы представления его знаний (знаковая информация);
- *материальная среда объектов и явлений*, на основе интерпретации которых человеком генерируются новые концепты знания, относящиеся к *ментальной среде*;
- *цифровая электронная среда*, к которой относятся цифровая информация, цифровые данные, компьютерные коды концептов и форм представления выраженных знаний.

Таким образом, каждое однородное по определению понятие было соотнесено только с одной из перечисленных четырех сред. Используя вышеупомянутое образное выражение Шрейдера [9, с. 50-51], можно сказать, что границы между этими средами являются той «пропастью», которая разделяет информацию, знания и другие однородные понятия.

После распределения однородных понятий по средам были описаны две последовательные стадии представления знаний в цифровой электронной среде

в виде компьютерных кодов. Первая стадия представления знаний относится к границе между ментальной и социально-коммуникационной средами. Семиотические знаки как двуединые по своей природе сущности принадлежат именно этой границе², на которой осуществляется ассоциативное соотнесение концептов и форм их представления с использованием семиотических знаков.

Рис. 1 иллюстрирует ситуацию, когда в системе знаний человека может быть выделено несколько планов содержания в зависимости от числа тех языков или знаковых систем, которыми он владеет и которые использует для представления своих знаний в отчужденной форме. В используемой системе терминов такие формы по определению являются знаковой информацией. На рис. 1 изображено два плана содержания в ментальной среде (обозначены как А и Б), соответствующие двум разным знаковым системам.

Например, для описания одних и тех же знаний могут использоваться русский язык (план содержания А) и английский язык (план содержания Б). Тогда каждому из этих двух планов содержания, полученных в результате членения одной и той же системы знаний, будет соответствовать свой план выражения в социально-коммуникационной среде. Каждый план выражения на рис. 1 обозначен двумя книгами (две книги на русском языке и две книги их переводов на английский). Граница между ментальной и социально-коммуникационной средами обозначена штриховой линией. На этой границе устанавливаются ассоциативные связи между концептами ментальных знаний и текстами как отчужденными формами представления концептов в социально-коммуникационной среде.

Вторая стадия представления концептов как «квантов» ментального знания относится к границе между социально-коммуникационной и цифровой электронной средами (рис. 2).

На этой границе используются таблицы компьютерного кодирования символов, необходимые для представления текстов в цифровой электронной среде. Таблицы принадлежат границе между этими средами (таблицы кодирования условно обозначены компьютером на рис. 2). Для вербализуемых знаний последовательное использование сначала языковых знаковых систем в процессе генерации текстов, а затем таблиц компьютерного кодирования является традиционным способом представления концептов в цифровой электронной среде.

Рис. 2 иллюстрирует ситуацию, когда текстам как вербальным формам представления знания ставятся в соответствие коды составляющих их символов. Важно отметить, что в процессе интеграции информационной науки с компьютерной наукой и формирования единой предметной области получились не две, а три границы между:

- ментальной средой знаний человека и социально-коммуникационной средой (на рис. 2 обозначена как интерфейс № 1);

² Здесь использована диадическая модель знака.



Рис. 1. Две среды, планы содержания и выражения, знаки

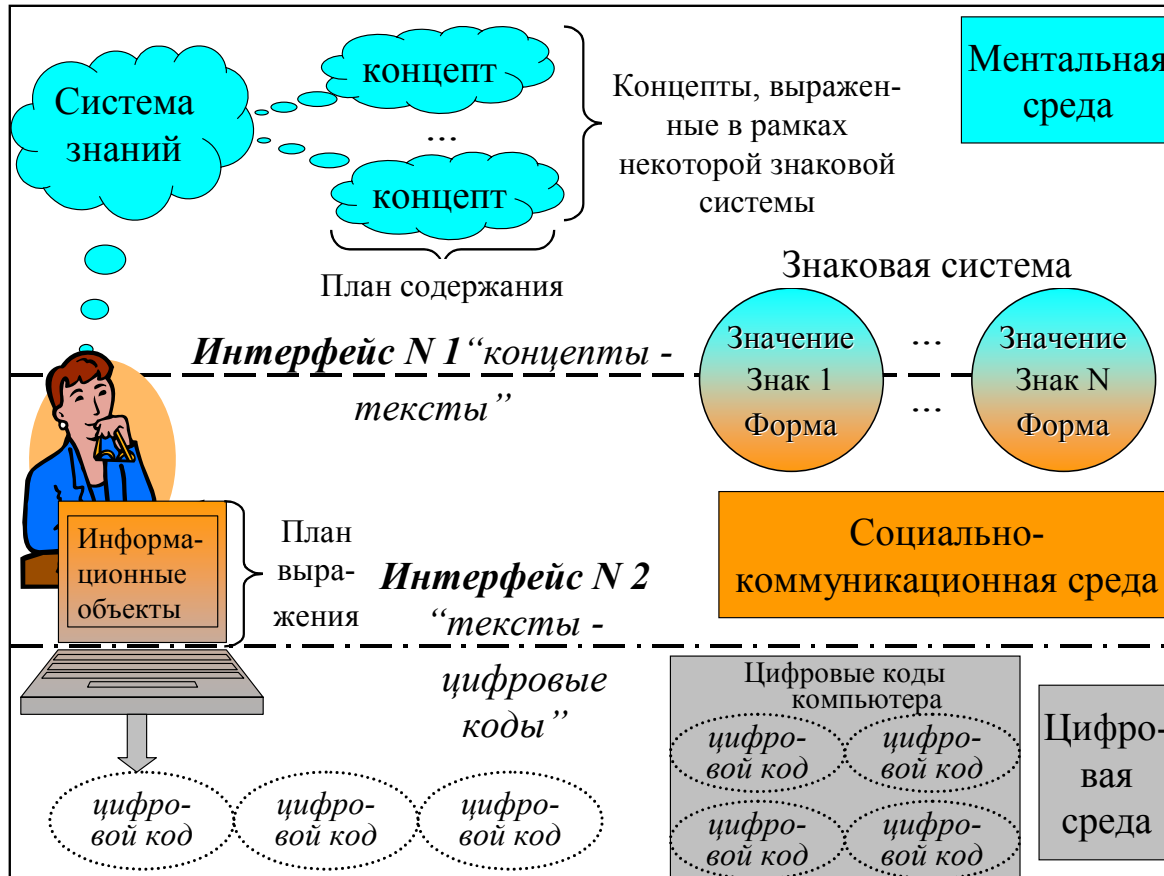


Рис. 2. Два интерфейса и две стадии представления концептов в виде цифровых кодов

- социально-коммуникационной средой и цифровой электронной средой (на рис. 2 обозначена как интерфейс № 2);

- ментальной средой знаний человека и цифровой электронной средой (эта граница на рис. 2 не показана).

Определение двуединых понятий на третьей границе, которые в используемой системе терминов названы семокодами, дало возможность предложить методы компьютерного кодирования тех концептов, которые не имеют вербальных форм представления или имеют несколько разных форм их представления из-за асимметрии знаковых систем [45].

Итак, решение двух названных нами проблем было найдено в результате использования и развития подходов Горна и Шрейдера к определению информационно-компьютерной науки как единой области знания. Отметим, что при решении первой проблемы (концептуальное индексирование геоизображений) использовались все четыре среды, а при решении второй (генерация систем экспертных знаний) – только три среды из четырех. Однако отсюда не следует, что для описания предметной области информационно-компьютерной науки всегда будет достаточно именно этих четырех сред.

Например, в задачах нейроинформатики необходимо включать в рассмотрение совокупность объектов других четырех сред: ментальной, социально-коммуникационной, цифровой электронной и нейрофизиологической [46] (нейробиологической [47]). Соответственно, объекты исследования предметной области нейроинформатики могут, в общем случае, принадлежать этим четырем средам. Кроме того, в процессе исследования могут использоваться понятия, принадлежащие следующим шести границам между:

- ментальной средой знаний человека и социально-коммуникационной средой;
- социально-коммуникационной и цифровой электронной средами;
- ментальной и цифровой электронной средами;
- нейрофизиологической и социально-коммуникационной средами;
- нейрофизиологической и цифровой электронной средами;
- ментальной и нейрофизиологической средами.

Для первой границы традиционным примером таких понятий являются семиотические знаки. Для второй и третьей границ аналогичные понятия – формокод и семокод – были определены в монографии [45]. В настоящее время актуальным является вопрос, как определить аналогичные двуединые понятия для оставшихся трех границ.

Если в задачах нейроинформатики дополнительно необходимо учитывать среду материальных объектов и явлений (внешних по отношению к человеку), то число сред ее предметной области возрастает до пяти. В настоящее время отсутствует общее описание средового деления предметной области информационно-компьютерной науки, учитывающее другие ее отраслевые составляющие: геоинформатику, когнитивную информатику, биоинформатику и т.д. Это не позволяет сегодня определить общее число сред, ко-

торые охватывает предметная область информационно-компьютерной науки. Однако, если к пяти упомянутым средам добавить аналоговую электронную среду, рассмотренную в монографии [45], то можно предположить, что общее число ее сред окажется не меньше шести.

ЗАКЛЮЧЕНИЕ

Фундаментальное переосмысление системы теоретических, прикладных, методологических и технологических принципов, подходов и понятий, используемых сегодня в сфере разработки ИКТ и их интеграции с когнитивными технологиями [31-33], стало ключевой предпосылкой становления информационно-компьютерной науки как общенаучной области знания. В настоящее время осознана и эксплицирована (с научной, технологической и образовательной позиций) потребность в построении ее теоретических оснований как фундаментальной науки [1, 10-15, 26]. Самые первые подходы к описанию информационно-компьютерной науки были предложены Горном и Шрейдером. В развитие их идей нами предложен один из возможных подходов к построению ее теоретических оснований [18, 42-45].

Предлагаемый подход включает: деление предметной области информационно-компьютерной науки на среды; использование двух классов понятий (однородных по своей природе и двуединых), распределение однородных понятий по средам и двуединых понятий по границам между средами; использование двух взаимосвязанных стадий представления концептов в цифровой электронной среде в виде компьютерных кодов.

В предлагаемом подходе однородные понятия могут принадлежать, как минимум, к шести разным средам: ментальной, социально-коммуникационной, цифровой электронной, аналоговой электронной, нейрофизиологической (нейробиологической), а также к среде материальных объектов и явлений, внешних по отношению к человеку. Предметные области других областей знания редко включают подобное разнообразие сред и, соответственно, разнообразие природы объектов исследования. Исключение составляет семиотика. В семиотическом треугольнике Фреге денотат является одной из его вершин и может быть любой природы (ментальной, социально-коммуникационной, цифровой электронной, нейрофизиологической и т.д.). Если задачей семиотики является изучение знаков и знаковых систем, используемых во всех областях знания, но *только на одной границе* между ментальной средой и социально-коммуникационной средой, то задачей информационно-компьютерной науки – создание теоретических основ разработки ИКТ для всех областей знания и широкого спектра сфер практической деятельности. При этом ИКТ будущих поколений смогут охватывать однородные сущности пяти и более сред, а также двуединые сущности на *нескольких границах* между ними.

Изучение столь широкого спектра однородных и двуединых сущностей позволяет говорить об особой роли информационно-компьютерной науки в системе научного познания. С одной стороны, трудно опре-

делить место для этой области знания в рамках классической парадигмы «двух культур» – естественных и гуманитарных наук [48]. С другой стороны, разнообразие сред и природы объектов исследования позволяет выдвинуть гипотезу о становлении информационно-компьютерной науки как «третьей культуры» [49], способствующей фундаментальному переосмыслению и гармонизации взаимосвязей естественных и гуманитарных наук.

Предлагаемое развитие идей Горна и Шрейдера является одним из возможных вариантов построения теоретических оснований информационно-компьютерной науки. Разумеется, могут быть предложены и другие варианты описания ее оснований. Разработка и сопоставительный анализ разных вариантов построения теоретических оснований представляют собой современную стадию становления информационно-компьютерной науки как общенаучной области знания, и, возможно, как «третьей культуры».

СПИСОК ЛИТЕРАТУРЫ

1. Snir M. Computer and information science and engineering: one discipline, many specialties // *Communication of the ACM*, March 2011. – P. 38–43.
2. Gorn S. The computer and information sciences: a new basic discipline // *SIAM Review*. – 1963. – Vol. 5, № 2. April, – P. 150–155.
3. Gorn S. The individual and political life of information systems. In: *Proc. Symposium on Education for Information Science*. – New York: Spartan Books, 1965. – P. 33–40.
4. Gorn S. Computer and information sciences and the community of disciplines // *Behavioral science*. – 1967. – Vol. 12, № 6. November. – P. 433–452.
5. Gorn S. The identification of the computer and information sciences: their fundamental semiotic concepts and relationships // *Foundations of language*. – 1968. – Vol. 4, № 4. November. – P. 339–372.
6. Gorn S. Informatics (computer and information science): its ideology, methodology, and sociology // *Knowledge: Creation, Diffusion, Utilization*. – 1982. – Vol. 4, № 2. – P. 173–198.
7. Gorn S. Informatics (computer and information science): its ideology, methodology, and sociology. In: *The studies of information: Interdisciplinary messages* / ed. by F. Machlup, U. Mansfield. – New York: Wiley, 1983. – P. 121–140.
8. Шрейдер Ю.А. ЭВМ как средство представления знаний // *Природа*. – 1986. – № 10. – С. 14–22.
9. Шрейдер Ю.А. Информация и знание. В кн. *Системная концепция информационных процессов*. – М.: ВНИИСИ, 1988. – С. 47–52.
10. Decision No 1982/2006/EC of the European Parliament and of the Council of 18 December 2006 concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007-2013) // *Official Journal of the European Union* L412 30.12.2006. – P. 1–41.
11. CORDIS ICT Programme Home. – URL: http://cordis.europa.eu/fp7/ict/programme/home_en.html (состояние страницы на 23.05.2011).
12. ICT FP7 Work Programme 2007-08. – URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2007-08_en.pdf (состояние файла на 23.05.2011).
13. ICT FP7 Work Programme 2009-10. – URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2009-10_en.pdf (состояние файла на 23.05.2011).
14. ICT FP7 Work Programme 2011-12. – URL: ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2011-12_en.pdf (состояние файла на 23.05.2011).
15. FP7 Exploratory Workshop 4 «Knowledge Anywhere Anytime». – URL: http://cordis.europa.eu/ist/directorate/f/f_ws4.htm (состояние страницы на 23.05.2011).
16. Зацман И.М., Кожунова О.С. Предпосылки конвергенции компьютерной и информационной наук // *Системы и средства информатики: спец. вып. Научно-методологические проблемы информатики*. – М.: ИПИ РАН, 2006. – С. 112–139.
17. Зацман И.М., Кожунова О.С. Предпосылки и факторы конвергенции информационной и компьютерной наук // *Информатика и ее применение*. – 2008. – Том 2, вып. 1. – С. 77–98.
18. Зацман И.М. Основы компьютерного представления экспертных знаний для мониторинга программно-целевой деятельности. Автореф. дис. ... докт. техн. наук. – М.: ИПИ РАН, 2011. – 50 с. – URL: http://www.ipiran.ru/announce/avto_2011_Zatsman.doc.
19. Brookes B.C. The foundations of information science. Part I. Philosophical aspects // *Journal of Information Science*. – 1980. – № 2. – P. 125–133.
20. Поппер К.Р. Объективное знание. Эволюционный подход / пер. с англ. Д.Г. Лахути; отв. ред. В.Н. Садовский. – М.: Эдиториал УРСС, 2002. – 384 с.
21. Михайлов А.И., Черный А.И., Гиляревский Р.С. Основы информатики. – М.: Наука, 1968.
22. Bryant A. Cognitive Informatics, Distributed Representation and Embodiment // *Brain and Mind*. – 2003. – Vol. 4, № 2. – P. 215–228.
23. Wang Y., Kinsner W. Recent Advances in Cognitive Informatics // *IEEE Transactions on Systems, Man, and Cybernetics (Part C)*. – 2006. – Vol. 36, № 2. – P. 121–123.
24. Wang Y. The theoretical framework of cognitive informatics // *International Journal of Cognitive Informatics and Natural Intelligence*. – 2007. – Vol. 1, № 1. – P. 1–27.
25. Wang Y., Widrow B., Zhang B., Kinsner W., Sugawara K., Sun F., Lu J., Weise T.,

- Zhang D. Perspectives on the Field of Cognitive Informatics and its Future Development // International Journal of Cognitive Informatics and Natural Intelligence. – 2011. – Vol. 5, № 1. – P. 1–17.
26. Computational Science: Ensuring America's Competitiveness. Report to the President. – Arlington, VA: National Coordination Office for Information Technology Research and Development, 2005.
 27. Колин К.К. Новая стратегическая компьютерная инициатива США и задачи России в области развития фундаментальной информатики // Информационные технологии. – 2006. – № 7. – С. 2–5.
 28. Аннотация и рубрикатор темы «Фундаментальные основы конвергентных технологий» в конкурсе РФФИ ориентированных фундаментальных исследований по актуальным междисциплинарным темам 2011 года. – URL: <http://www.rfbr.ru/rffi/getimage/20>. Фундаментальные основы конвергентных технологий.pdf?objectId=31173.
 29. Мысяков Д. Практика широкого формата (интервью с П. Пашининым) // Газета Поиск». – 2012. – № 12(1190). – С. 6–7.
 30. Список проектов конкурса "ОФИ-М-2011", получивших финансовую поддержку РФФИ. – URL: <http://www.rfbr.ru/rffi/getimage?objectId=38202>.
 31. The National Science Foundation's FY 2003 funding request to Congress. – URL: <http://www.nsf.gov/about/budget/fy2003/goals.htm>.
 32. Converging technologies for improving human performance: nanotechnology, biotechnology, information technology and cognitive science. U.S. National Science Foundation / eds. M. Roco, W.S. Bainbridge. – Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003. – 482 p.
 33. Converging technologies - shaping the future of European societies. A report from the high level expert group on "foresighting the new technology wave" / ed. By A. Nordmann. – Brussels: European Communities, 2004. – 68 p.
 34. Nonaka I. The knowledge-creating company // Harvard Business Review. – 1991. – Vol. 69, № 6. – P. 96–104.
 35. Nonaka I., Takeuchi H. The knowledge-creating company. – Oxford; N.Y.: Oxford University Press, 1995 (перевод на русский язык: Нонака И., Такеучи Х. Компания – создатель знания. – М.: ЗАО «Олимп-бизнес», 2003).
 36. Knowledge emergence / ed. by I. Nonaka and T. Nishiguchi. – Oxford; N.Y.: Oxford University Press, 2001.
 37. Wierzbicki A.P., Nakamori Y. Basic Dimensions of Creative Space. In: Creative space: Models of Creative Processes for Knowledge Civilization Age / eds by A.P. Wierzbicki, Y. Nakamori. – Springer Verlag: Berlin-Heidelberg, 2006. – P. 59-90.
 38. Wierzbicki A.P., Nakamori Y. Knowledge sciences: Some new developments // Zeitschrift für Betriebswirtschaft. – 2007. – Vol. 77, № 3. – P. 271–295.
 39. Wierzbicki A.P., Nakamori Y. The importance of multimedia principle and emergence principle for the knowledge civilisation age // Journal of Systems Science and Systems Engineering. – 2008. – Vol. 17, № 3. – P. 297–318.
 40. Yamashita Y., Nakamori Y., Wierzbicki A.P. Knowledge synthesis in technology development // Journal of Systems Science and Systems Engineering. – 2009. – Vol. 18, № 2. – P. 184–202.
 41. Мамардашвили М.К. Классический и неклассический идеалы рациональности. – М.: Издательство «Логос», 2004. – 240 с.
 42. Зацман И. М. Семиотическая модель взаимосвязей концептов, информационных объектов и компьютерных кодов // Информатика и ее применения. – 2009. – Том 3. Вып. 2. – С. 65–81.
 43. Зацман И. М. Нестационарная семиотическая модель компьютерного кодирования концептов, информационных объектов и денотатов // Информатика и ее применения. – 2009. – Том 3. Вып. 4. – С. 87–101.
 44. Зацман И.М., Косарик В.В., Курчавова О.А. Задачи представления личностных и коллективных концептов в цифровой среде // Информатика и ее применения. – 2008. – Том 2. Вып. 3. – С. 54-69.
 45. Зацман И.М. Концептуальный поиск и качество информации. – М.: Наука, 2003. – 271 с.
 46. Shepard R.N. Perceptual-Cognitive Universals as Reflections of the World // Behavioral and Brain Sciences. – 2001. – Vol. 24, № 3. – P. 581–601.
 47. Каменская М.А. Понятие «информация» в контексте молекулярно-клеточной биологии // Научная и техническая информация. Сер. 1. – 2012. – № 6. – С. 4-17.
 48. Сноу Ч.П. Две культуры. – М.: Прогресс, 1973.
 49. Snow C.P. The Two Cultures: A Second Look. – Cambridge: Cambridge University Press, 1963.

Материал поступил в редакцию 11.02.13.

Сведения об авторе

ЗАЦМАН Игорь Моисеевич - доктор технических наук, зав. отделом Института промышленной информатики РАН, Москва
E-mail: izatsman@yandex.ru

Исходные понятия энтропии, порядка, организации, информации, знания и смысла

Показано, что классическое понятие энтропии, как феноменологически, так и теоретически, не отражает сути второго закона термодинамики. В связи с чем, дается новое понятие энтропии – S_n , придающее ей конкретный физический смысл и возможность экспериментального определения ее величины. Это позволило уйти от многочисленных попыток определения феномена жизни через ее энтропийные характеристики и дать определения жизни и элементарной структуры, как самой жизни, так и моделирующих ее устройств (автоматов, компьютеров, роботов). Это, в свою очередь, позволило придать математическому характеру наук о них (теории информации, кибернетике, теории систем и т.п.) содержательный смысл и определить операционально исходные единичные понятия (информации, знания, смысла, управления), а также сформулировать операциональные понятия упорядоченности, с одной стороны, и организации и самоорганизации – с другой и осуществить новый подход к проблемам биологии и биофизики.

Ключевые слова: энтропия, порядок, организация, знание, смысл, информация

ВТОРОЙ ЗАКОН ТЕРМОДИНАМИКИ И ПОНЯТИЕ УПОРЯДОЧЕННОСТИ

Рассмотрим вначале результаты экспериментов, собранных в течение почти двух веков, в которых из хаоса возникали упорядоченные структуры и которые подвели ученых к идее создания синергетики. К ним относятся структурирование потоков при электролизе, отмеченное А. Фехнером (1828г.), образование сотообразных структур при испарении керосина из находящегося в его растворе шеллака, отмеченное Г. Спенсером (1850 г.), получение Р. Лизегангом структур в виде колец в растворах и порошках в результате химических реакций (1896 г.), получение А. Бенаром шестиугольных структур в воде, возникавших при определенных условиях в процессе ее подогрева (1900-1901 гг.), а также известная реакция Б. Белоусова – А. Жаботинского (1951-1969 гг.). Уже Г. Спенсер на основании своего и предшествующих ему экспериментов выдвинул идею, являющуюся, по его мнению, *принципом жизни, суть которого – постоянный переход гомогенной системы в гетерогенную*. Эту идею фактически взяли на вооружение другие авторы, возникшей впоследствии синергетики [1, с. 314-315]. Однако ни одна из перечисленных выше упорядоченных систем, возникших из хаоса, далее не эволюционировала, а динамичная система в реакции Б. Белоусова – А. Жаботинского совершала только колебательные движения, также без дальнейшей эволюции. Но, очевидно, что каждая возникшая подобным образом молекула упорядоченной системы отличается от белковых и нуклеотидных молекул миллионами этапов дальнейшей эволюции. Кроме того, процесс этой гипотетической эволюции должен

был бы иметь определенное направление. Достаточно сказать, что, как было показано, возможное число изомеров ДНК кишечной палочки равно $10^{1000000}$ [2, с. 21]. Число это невообразимо огромно, например, число атомов в видимой Вселенной составляет не более 10^{100} . Но кроме единственного изомера едва ли все остальные способны обеспечить полноценную жизнеспособность организма. Таким образом, идея Г. Спенсера, подвигнувшая других ученых на создание синергетики, фактически не дала никакого обоснования для последующей экстраполяции.

Столь же безосновательно экстраполируется предположение о многовариантности перехода хаоса в упорядоченную структуру в зависимости от той или иной бифуркации. Как справедливо отмечает французский математик и философ Р. Том, изучение субстрата эволюционирующей системы позволяет предвидеть все возможные типы траекторий, приводящих от хаоса к разным видам упорядоченности. *Флуктуации могут выступать только как фактор, развязывающий процесс самоорганизации, но не детерминирующий его* [3]. Рассмотрим два крайних случая бифуркаций, подтверждающих его тезис. Наиболее простой из них пример в виде переохлажденной жидкости. Практически любая бифуркация в виде слабого воздействия на эту жидкость не может привести ни к чему иному, как к ее кристаллизации. Аналогично, в другом, наиболее сложном случае, бифуркация в виде идеи западного мироустройства, обеспечивающего высокие жизненные стандарты, может быть реализована только при достаточном культурном уровне населения. В то же время авторы синергетики и их последователи утверждают, что на

ее основе будет создана общая теория эволюции любых систем, начиная от физико-химических и кончая психологическими и социальными [4-6]. Таким образом, очевидно, что *никакая единая физическая характеристика не может выступать в качестве фактора, обеспечивающего эволюцию всего многообразия систем.*

Однако идея Г. Спенсера была не только принята на вооружение создателями синергетики, но под нее упорно создается уже в течение полувека теоретическая физико-математическая база. Строится она на основе представления о том, что рост энтропии рассматривается в физике именно как общий показатель деградации и разрушения и что, наоборот, при убыли энтропии происходят процессы упорядочения, организации и даже самоорганизации. Именно на этом представлении строятся, в частности, в астрофизике сценарии гибели Вселенной. Это мнение стало столь распространенным, что вошло даже в культурный контекст.

Однако более глубокий анализ феноменологии входит в противоречие с этим представлением. Например, в случае изоляции трехфазной смеси лед-вода-пар, в зависимости от начальных параметров компонентов смеси (масс, температур, давлений), эта смесь может перейти в состояние пара, т.е. в состояние, близкое к хаосу, в другом случае – эта смесь может превратиться в упорядоченный ледовый кристалл. Этот пример относится не только к воде, но и к подавляющему большинству трехфазных систем. Но возможность образования порядка при росте энтропии относится не только к изолированным системам. Рассмотрим, например, термодинамику образования периодической таблицы химических элементов. С этой точки зрения она, с одной стороны, сходна, а с другой – принципиально различна для легких элементов до № 26 (железа) и для железа и элементов с большим, чем железо атомным весом. В обоих случаях происходило *упорядочение*, так как из хаоса более мелких ядер в процессе космической эволюции возникали более упорядоченные структуры более тяжелых ядер. Но в первом случае образование ядер шло с выделением тепла (подобно взрыву водородной бомбы), тогда как во втором – с поглощением его при взрывах сверхновых звезд. Таким образом, три четверти химических элементов связаны своим происхождением с ростом энтропии. Следовательно, упорядочение, как и хаотизация, как для изолированных, так и для открытых систем может идти *как с ростом, так и с убылью энтропии* [7-15].

Суть возникающего парадокса заключается в том, что выводы Больцмана выполнены методами статистической физики на модели хаотической, *бессвязной* структуры идеального газа и поэтому не подходят непосредственно для изучения реальных систем, специфика которых определяется *связями элементов*. При переходе системы из одного состояния в другое или преобразование одной системы в другую изменяются состояния связей между элементами, а в случае разрыва прежних связей образуются новые. От энергетического баланса этих процессов и зависят энергетический и энтропийный итоги процесса. При переходе от идеальной системы к реальным эти факторы в термодинамике не учитываются, как это сде-

лано, например, в аэродинамике и гидродинамике (движение с сопротивлением), в теории информации при переходе от идеального канала к реальному (выделение сигнала из шума) и т.п.

Первым, обратившим внимание на несоответствие выводов Л. Больцмана реальности, по крайней мере, в масштабах Космоса, был французский ученый А. Дюкрот, указавший, что благодаря гравитации происходит процесс упорядочивания при возникновении небесных тел из первичного хаоса [16, с. 106-122]. В свете открытий последних десятилетий высказанное А. Дюкротом положение усиливается противоположными, с точки зрения энтропийных оценок, эффектами «темной массы» и черных дыр, с одной стороны, и «темной энергии» – с другой, что говорит о том, что свойства Космоса еще недостаточно изучены, чтобы судить о том, каковы сценарии эволюции Вселенной.

ЭНТРОПИЯ – СМЫСЛ И ВЫРАЖЕНИЕ

Однако сказанным не исчерпываются «энтропийные» парадоксы. В основе другого парадокса лежит лишенное основания отождествление дифференциальной формы выражения работы $\delta A = P_i dx_i$ и тепла $\delta Q = T dS$, что едва ли допустимо без глубокого анализа, так как работа – процесс, происходящий на *макроуровне*, а теплообмен – процесс, происходящий на *микроуровне* [17, с. 40, с. 52]. Здесь P_i – потенциалы (точнее их разность), а x_i – координаты. Например, P – давление, v – объем газа, F – сила, перемещающая тело, l – путь, V – электрический потенциал, q – заряд, μ – химический потенциал, m – масса продуктов реакции и т.п. [18, с. 17-18, с. 22-23]. Отличием энтропии от координат рабочих процессов является неясность ее физического смысла, так как она «*на опыте не измеряется и не наблюдается*». Она может быть только вычислена через другие наблюдаемые на опыте величины» [18, с. 19]. Неслучайно Джон фон Нейман высказался с откровенностью, позволительной признанным классикам: «Никто не знает, что же такое в действительности энтропия» [19, с. 153-163]. Тем не менее, выражение энтропии используется для обоснования второго закона термодинамики в виде доказательства роста энтропии в процессе теплообмена между двумя телами, изолированными от внешней среды, в виде $\delta Q_1 = \delta Q_2$, откуда $T_1 dS_1 = T_2 dS_2$ и $dS_2 > dS_1$.

Чтобы выяснить физический смысл энтропии, исследуем исходный вывод ее формулы [17, с. 62-64], начиная с определения КПД – η цикла тепловой машины Карно (рис. 1):

$$\eta = \frac{Q - Q_0}{Q} = \frac{T - T_0}{T}, \quad (1)$$

где Q и Q_0 , T и T_0 , соответственно, тепло, полученное газом от нагревателя при температуре – T , и тепло, отнятое у него холодильником, при температуре – T_0 .

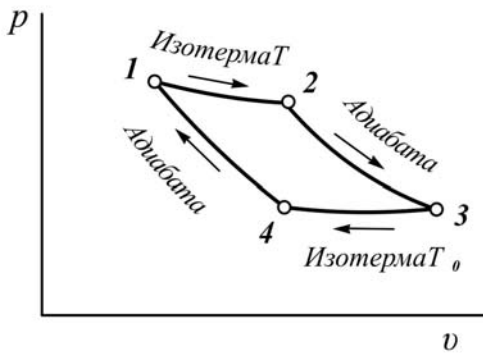


Рис. 1. Цикл Карно

Из (1) получаем величину ω , т.е. той части тепловой энергии, которая будет необратимо рассеяна и потенциал которой станет ниже исходного и далее:

$$\omega = 1 - \eta = \frac{Q_0}{Q} = \frac{T_0}{T}, \quad (2)$$

$$\frac{Q}{T} = \frac{Q_0}{T_0}, \quad (3)$$

$$\frac{\delta Q}{T} = \frac{\delta Q_0}{T_0}. \quad (4)$$

Итог этих преобразований выражения (4) назван «приведенными теплотами». Отмечается, что в итоге чисто математических преобразований (3) и (4) может быть потеря их физический смысл [17, с. 66]. Затем рассматривается произвольный процесс работы тепловой машины (рис. 2), где цикл разбит на равные отрезки параллельными адиабатами на микроциклы Карно [17, с. 65-68].

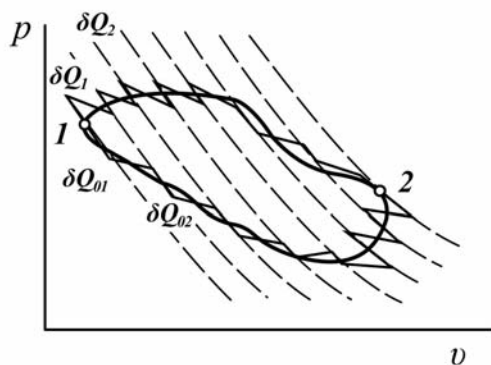


Рис. 2. Схема термодинамического цикла (к доказательству теоремы Карно)

В (5) дана сумма приведенных теплот – S , названная энтропией, которая согласно (4) будет

$$S = \int_1^2 \frac{\delta Q}{T}, \quad (5)$$

$$S = S_0, \quad (6)$$

из чего делается вывод, что энтропия – S не зависит от пути процесса и является, поэтому функцией состояния.

Но работа (A), которая в каждом из этих двух случаев измеряется величиной

$$A = \int_1^2 p dv, \quad (7)$$

(площадью под кривыми) по верхней кривой больше. Опираясь средними величинами температур, получим, согласно (2) (так как средняя температура по верхней кривой больше таковой по нижней), что *необратимость* – ω в первом случае меньше. Но величина энтропии, оставшаяся в обоих случаях неизменной, этого не отразила. Таким образом, энтропией не выполняются основные требования: во-первых, *второго закона термодинамики, говорящего о факте и мере необратимости процессов, и во-вторых, требования к координате, должной изменяться аналогично рабочим процессам одновременно с изменением потенциала* (в данном случае T) [18, с. 17-18, с. 22-23]. Действительно, изменение P_i должно однозначно привести к изменению x_i , например, с изменением давления P должен измениться объем газа – v , с изменением движущей силы – F должен измениться путь – l , проходимый телом, с изменением электрического потенциала – V , должен измениться заряд – q , с изменением химического потенциала μ – должна измениться масса продукта – m и т.п.

Для поиска координаты, сопряженной с тепловым потенциалом T воспользуемся следующим соображением. Как известно, отличие тепла от работы, которая может быть преобразована непосредственно в любую другую форму работы (механическую, электрическую и т.п.), тепло может быть преобразовано в работу только через посредство промежуточного (рабочего) тела, т.е. *через проявление теплового микропроцесса на макроуровне* [17, с. 60-61]. В цикле Карно функцию рабочего тела выполняет газ. Получив тепло от нагревателя (первое тело), газ (второе тело) может выполнить работу только через третье тело, например, через поршень, приводящий в движение локомотив, или термоду, в которой возникнет электрический ток. При этом могут совершаться и другие виды работ, например, тепловое расширение поршня и цилиндра, термоду и т.п. С учетом сказанного и исходя из закона сохранения энергии, а также оставляя за новым выражением координаты теплового потенциала прежнее название энтропии и добавив во избежание путаницы к ней индекс – n , получим:

$$TdS_n = \sum_k P_i dx_i, \quad (8)$$

$$dS_n = \sum_k P_i dx_i / T, \quad (9)$$

где k число работ, которые были совершены всеми телами в результате получения ими тепла от нагревателя. Таким образом, *энтропия приобретает реальный физический смысл и предстает в виде обычной, экспериментально определяемой координаты.*

Следовательно, фактором, реализующим второй закон термодинамики, является не рост энтропии, а стремление разности потенциалов к нулю, что позво-

ляет выразить второй закон термодинамики в простом и ясном виде

$$\sum_k P_i \mapsto 0, \quad (10)$$

а суммарная мера необратимости любого процесса выразится как

$$\omega = \sum_k \omega_i, \quad (11)$$

а мерой реализации этого закона, в частности в цикле Карно, его необратимость, выраженная в (2).

Статистическое выражение второго закона термодинамики:

$$S = k \ln W, \quad (12)$$

где W – термодинамическая вероятность, т.е. число микросостояний, реализующих данное макросостояние, с учетом того, что:

$$W_P \gg W_{HP}, \quad (13)$$

где P и HP означают соответственно равновесные и неравновесные состояния, то выражение (12) может быть представлено в виде, аналогичном по форме выражению (10)

$$W_{HP} \rightarrow W_P. \quad (14)$$

ВТОРОЙ ЗАКОН ТЕРМОДИНАМИКИ И БИОЛОГИЯ

В наших работах на основании предлагаемого подхода удастся устранить кажущийся парадокс физического плана феномена жизни [7-15]. Суть его заключается в том, что, согласно существующим представлениям об энтропии, растения, получая тепло от Солнца, увеличивают свою энтропию с одновременным повышением своей организации, а через трофические цепи и организацию всей биосферы. Многие ученые, по предложению Н.А. Умова (1901 г.), пытались объяснить это противоречие с позиций третьего, специфичного только для жизни закона термодинамики (противоположного ее второму закону) [20, с. 108-109], с позиций открытости систем – Л. фон Бергаланфи (1953 г.) [21], с позиций негэнтропийного принципа информации – Л. Бриллюэн (1956 г.) [22, с. 200-201], на основе синергетики – Г. Хакен (1969 г.) [23, с. 12]. На базе этих работ возникло целое направление их последователей, включая таких авторов, как М. Эйген [24] и И. Пригожин [25].

Однако, как представляется, *решение проблемы физической сущности жизни не связано непосредственно с ее энтропийными характеристиками*. Как было показано П.К. Анохиным и Н.А. Бернштейном, специфика жизни заключается в способности организмов с опережением адекватно реагировать на различные события до непосредственного контакта с ними, например, на звук, запах, световой импульс и т.п., еще задолго до контакта с порождающим их объектом (жертвой, хищником, водой, половым партнером и т.п.). Способность к опережающему реагированию придает жизни огромные преимущества по сравнению с другими объектами. Она позволяет организмам заранее уклоняться от губительных воздействий и энергично стремиться к условиям, обеспечивающим существование их и их вида. Но никто

не ставил при этом вопрос: *каким физическим условиям с необходимостью должны удовлетворять организмы, чтобы реагировать подобным образом?* Таких условий всего три.

1. *Термодинамическое*, т.е. организм должен обладать потенциальной энергией. При этом она должна сохраняться без рассеяния достаточно продолжительное время, чтобы обеспечивать работу, необходимую для существования организма. Это требование удовлетворяется широко распространенными в природе метастабильными состояниями, в которых энергия высокого потенциала защищена от выравнивания потенциальным барьером. Наиболее простым примером такого барьера может быть плотина гидроэлектростанции. Вода перед ней обладает запасом потенциальной энергии по сравнению с уровнем воды за плотиной. Именно потенциальным барьерам обязано своим существованием все разнообразие устойчивых изотопов периодической таблицы и разнообразие мира в целом. Иначе все элементы периодической таблицы скатились бы к ее центру за счет синтеза легких (взрыв водородной бомбы) и распада тяжелых (взрыв атомной). Роль потенциальных барьеров здесь играют многомиллионные температуры или необходимость бомбардировки частицами, разрушающими структуру атомного ядра для обеспечения переходов от элемента к элементу. В организме эта энергия сохраняется в основных его веществах – жирах, углеводах и белках, *а у бактерий может сохраняться и в виде неорганики (серы, мышьяка, соединений железа и т.п.)*.

2. *Информационное*, т.е. организм должен обладать веществами или структурами, регулирующими процесс освобождения этой энергии в ответ на *сигнал-информацию – слабый, но специфический энергетический импульс*. В примере с плотиной такой информацией будет воздействие на устройство, перемещающее заслонку, перекрывающую переток воды с верхнего уровня на нижний. Этому условию отвечают органические и минеральные катализаторы, которые могут быть введены или, наоборот, выведены из контакта с веществом за счет слабого воздействия, например механического, или их активации за счет добавки к ферменту кофермента – выключатели электросети, разнообразные триггеры и т.п. Подобные структуры (называемые нами стрейторами от английского straight – прямой) обладают способностью изменять состояния потенциального барьера. Они способны при получении сигнала снижать (в пределах устранять) или, наоборот, восстанавливать потенциальный барьер метастабильного состояния, т.е. *управлять процессом выделения энергии*. Стрейторные реакции обладают высокой селективностью, например, в то время как нагревание ускоряет множество химических реакций катализатор ускоряет одну или несколько.

3. *Преобразовательное*, т.е. организм должен обладать веществами или структурами, преобразующими выделившуюся энергию высокого потенциала в работу по сохранению организма. Например, это может быть молекула фермента, так как она выполняет не только каталитическую функцию, но иногда вместе с мембраной определяет направление реак-

ции, или кинематическая часть станка, преобразующая вращения двигателя в конкретную работу, *костно-связочный аппарат организма, преобразующий сокращения мышц в огромное число комбинаций – разнообразных движений*. Свойство преобразования энергии при взаимодействии с телами присуще вообще всем материальным объектам. В примере с гидроэлектростанцией таким преобразователем будет блок гидротурбина-электрогенератор.

Структура, удовлетворяющая трем перечисленным условиям, являет собой сигнальный элемент или сокращенно – *сиэл*. На понятии сиэла определяются основные исходные понятия, лежащие в основе теоретической биологии и наук, описывающих устройства, выполняющие функции, свойственные организмам: теории информации, кибернетики, общей теории систем и т.п.

Знание. Сиэл – это элементарная структура, которая «знает» на какой сигнал и как реагировать.

Смысл. В структуре сиэла заключен элементарный семантический, а не только физический аспект реакции, направленный конкретно на локальную операцию. Из многих триллионов этих элементарных смыслов складываются функции отдельных клеток, из них – функции органов, а из них – функции организма в целом.

Управление. Сиэл – являет собой элементарную структуру управления, в которой малая энергия сигнала-информации управляет существенно более мощными энергетическими потоками.

Программа. Программа – это структура, способная под воздействием энергетического потока порождать сигналы для данного организма или автомата. Например, текст – программа, порождающая информацию при воздействии потока света, магнитная лента – программа, порождающая информацию под воздействием магнитного потока и т.п. Программа, как и сигнал-информация, является понятием *относительным*. Так, неподвижный ландшафт, порождающий сигналы для человека и многих животных, не является программой для лягушки, способной замечать лишь движущиеся предметы. Примеры программ: ДНК и РНК, магнитные ленты и лазерные диски и т.п.

Мир по отношению к информационному устройству содержит в себе лишь программы. Например, геологическое обнажение, на котором виден тот или иной этап геологической эволюции, является программой. Для превращения ее в сигналы-информацию необходим энергетический поток (солнечный или электрический свет), и лишь тогда геолог сможет получить сигналы, из части которых он извлечет информацию о содержании обнажения. *Очевидно, что реальная информация всегда будет носить относительный характер*, в отличие от ошибочного понятия абсолютной информации Л. Бриллюэна [22, с. 200-201]. Действительно, текст будет источником информации для человека, но не для животного, геологическое обнажение будет источником гораздо большей информации для специалиста, чем для простого зрителя и т.п.

Организация. Очевидно, что организованными называются системы, существование которых обеспечивается, в частности, за счет содержащихся в них

сигнальных элементов, т.е. организмы, автоматы, компьютеры, роботы и их совокупности.

Самоорганизация. Сигнал по своей мощности может на много порядков уступать энергии реакции, возникающей при его получении. Иными словами, обеспечение реакции энергией и связанные с этой реакцией перестройки в организме осуществляются уже самим организмом, т.е. фактически происходит *самоорганизация*. Либо по договоренности можно назвать *самоорганизацией* необходимые перестройки в организме, связанные с необходимостью осуществления реакции уже не на внешние воздействия, а на информацию из внутренних программ организма, например, голода, исследовательской потребности и т.п.

Определение феномена жизни. Организмы должны проявлять активность для обеспечения питания, роста, размножения, освоения новых территорий и т.п. Таким образом, способность к информационным реакциям является необходимым условием жизни, а активность – достаточным. Отсюда *определение жизни как активной, сигнальной (информационной) формы существования систем*.

Порядок (упорядоченность). Упорядоченность и организация – принципиально разные понятия. Понимание существа организации позволяет рассмотреть уже не на интуитивном, а на конкретном уровне различие между ними. Определение упорядоченности принадлежит Дж. фон Нейману и сводится к тому, что *система является тем более упорядоченной, чем меньше требуется информации для ее описания*. Например, если бесконечное количество точек на плоскости распределено хаотично, то для описания их расположения потребуются бесконечное количество информации. Действительно, для этого необходимо будет записать координаты каждой точки. Если же их расположение закономерно, например, в виде прямой линии, то запись будет выглядеть как $y=kx$. Из этого следует, что если описание системы может быть дано в математической форме, то упорядоченность системы будет тем выше, чем меньше потребуются символы для ее выражения. Например, простота космических законов (ньютоновский закон всемирного тяготения, электростатического взаимодействия, законы Эйнштейна) говорит о высокой упорядоченности Космоса. Положение о минимуме символов, характеризующем степень упорядоченности, остается в силе и для текстового описания, например, алгоритма, ибо, чем более удастся упростить алгоритм, тем более упорядоченным становится описание действий по его реализации.

Совокупность приведенных понятий дает унифицированную, операциональную основу наукам о жизни и моделирующим ее устройствам и позволяет в принципе осуществлять исчисление и математический анализ исследуемых в них структур [7-15].

Может, однако, показаться малоперспективной попытка исчислять сложные системы на основе сиэлов, которые в единственной клетке исчисляются многими миллионами единиц. Но несомненно, что существуют закономерности, организующие их в некоторые ассоциации. Таковую представляет собой, например, линейная последовательность сиэлов, регулируемая накоплением конечного продукта — A_n (в химии — процесс ретроингибирования). В этой

последовательности постоянство концентрации A_n поддерживается тормозящим воздействием этой концентрации на первый стрейтор в цепи Φ_1

$$A_1 \xrightarrow{\Phi_1} A_2 \xrightarrow{\Phi_2} \dots A_i \xrightarrow{\Phi_i} \dots A_n \xrightarrow{\Phi_n} \Phi_1. \quad (17)$$

В принципе, по той же схеме работает и автоматическая линия, работа которой останавливается, когда в накопителе оказывается достаточное количество деталей. Тогда $A_i \xrightarrow{\Phi_i} -$ это сизл, где A_i – энергоресурс в метастабильном состоянии. Таким образом, оказывается возможным исчислить количество сизлов (n-1) в структуре следующего иерархического уровня. В [7-15] показано, что приведенная структура, как и структуры более высокого порядка, регулярно повторяются на разных уровнях – от молекулярного до организменного и государственного, что, наряду с быстро возрастающей мощностью компьютеров, делает этот подход реалистичным для анализа сложных систем [7-15]. Таким образом, совокупность приведенных понятий дает унифицированную операциональную основу биофизике, теоретической биологии, теории информации, кибернетике, общей теории систем и позволяет в принципе осуществлять исчисление и математический анализ исследуемых в них структур. Кроме того, способность компьютеров к интеллектуальным операциям при различии образующих их элементов (электронных лампах, ферритах, полупроводниках, криотронах) свидетельствует о наличии идеи жизни, которая может быть реализована на различных материалах и делает бесперспективными поиски в Космосе жизни, идентичной в материальном плане земной.

СПИСОК ЛИТЕРАТУРЫ

1. Серебровская К.Б. Сущность жизни. – М., 1994. – 400 с.
2. Эйген М., Винклер Р. Игра жизни. – М., 1979. – 93 с.
3. Сокулер З.А. Спор о детерминизме во французской философской литературе // Вопросы философии. – 1993. – № 2. – С. 21-23.
4. Ласло Э. Материалы международной конференции по синергетике // Вопросы философии. – 1998. – № 3. – С. 21-23.
5. Волькенштейн М.В. Биология и физика // Успехи физических наук. – 1973. – Т. 109, вып. 3. – С. 499-516.
6. Князева Е.Н., Курдюмов С.П. Синергетика как новое мировоззрение: диалог с И. Пригожиным // Вопросы философии. – 1992. – № 12. – С. 5-20.
7. Штеренберг М.И. Проблема Бергаланфи и определение жизни // Вопросы философии. – 1996. – № 2. – С. 51-56.
8. Штеренберг М.И. Синергетика и биология // Вопросы философии. – 1999. – № 2. – С. 95-108.
9. Штеренберг М.И. Принципы организации и самоорганизации // Биофизика. – 2000. – Т. 45, № 3. – 576 с. Депонент ВИНТИ за № 514-В2000 от 28 февраля 2000 г.

10. Штеренберг М.И. Сущность жизни: физический, системный и мистический планы // Научно-техническая информация. Сер. 2. – 1999. – № 6. – С. 7-11; Shterenberg M.I. The essence of life: the physical, of the system, and the mystic. // Automatic documentation and mathematical linguistics. Allertonpress, Inc. – 1999. – Vol. 33, № 3. – P. 19-24.
11. Штеренберг М.И. Место синергетики в науке // Научно-техническая информация. Сер. 1. – 2002. – № 8. – С. 1-9; Shterenberg M.I. The place of synergetics in science // Scientific and technical information processing. Allertonpress, Inc. – 2002. – Vol. 29, № 4. – P. 13-22.
12. Штеренберг М.И. Физическая сущность жизни и начала теории организованных систем. – М., 2003. – 164 с.
13. Штеренберг М.И. Энтропия в теории и в реальности // Вопросы философии. – 2003. – № 10. – С. 103-108.
14. Штеренберг М.И. Синергетика: наука? философия? псевдорелигия? – М.: Academia., 2007. – 176 с.
15. Штеренберг М.И. Физика жизни и термодинамика // Вестник Российского философского общества. – 2011. – № 2. – С. 93-95.
16. Дюкрок А. Физика кибернетики // Сб. Кибернетика ожидаемая и кибернетика неожиданная. – М., 1968. – С. 106-122.
17. Путилов К.А. Термодинамика. – М., 1971. – 375 с.
18. Леонова В.Ф. Термодинамика. – М., 1968. – 158 с.
19. Большаков Б.Е., Минин В.В. Взаимосвязь второго закона термодинамики, принципов устойчивого неравновесия и информации // В сб. Эрвин Бауэр и теоретическая биология. – Пушкино, 1993. – 256 с.
20. Кузнецов П.Г. К истории приложения термодинамики к биологии // Тринчер К.С. Биология и информация. – М.: Наука, 1965. – С 107-118.
21. Бергаланфи Л. фон. Системные исследования. – М., 1969.
22. Бриллюэн Л. Наука и теория информации. – М., 1960.
23. Хакен Г. Информация и самоорганизация. – М., 1991. – 240 с.
24. Эйген М. Молекулярная самоорганизация и ранние стадии эволюции // Успехи физических наук. – 1973. – Т. 109, вып. 3. – С. 545-580.
25. Пригожин И., Николис Ж. Биологический порядок, структура и неустойчивости // Успехи физических наук. – 1973. – Т. 109. – С. 517-544.

Материал поступил в редакцию 16.03.13.

Сведения об авторе

ШТЕРЕНБЕРГ Михаил Иосифович – кандидат технических наук, Москва
E-mail: shterenberg@inbox.ru

УДК 004.5: [550.34:001.103]

И. М. Алешин, В. Н. Корягин, О. В. Сухорослов, К. И. Холодков, А. Н. Шогин

Инверсия сейсмических данных: высокоуровневый веб-интерфейс к инструментарию Globus Toolkit*

Приведено описание грид-приложения решения обратной геофизической задачи. Использование веб-интерфейса к грид-системе не только значительно ускоряет вычисления за счет проведения параллельных вычислений, но и существенно упрощает доступ к использованию распределенных вычислительных ресурсов.

Ключевые слова: *распределенная вычислительная система, грид-приложение, веб-интерфейс*

Использование аппарата апостериорной функции распределения (АПФР) вероятности [1] сводит решение обратной задачи к табулированию функции многих переменных. Так как значение АПФР вычисляется в каждой точке независимо, то использование слабо связанных распределенных вычислительных систем позволяет существенно расширить круг задач, допускающих точное решение. Основой распределенных вычислительных систем являются программные продукты, называемые промежуточным программным обеспечением (*middleware*), в задачи которого входит организация многопользовательской работы в распределенных системах при максимальной защите от несанкционированного использования предоставляемых ресурсов. Распределенные системы, использующие промежуточное программное обеспечение для разделяемого использования вычислительных ресурсов, называются грид-системами, или просто грид. Такое программное обеспечение можно условно разделить на две группы по принципу используемых вычислительных ресурсов: простаивающие ресурсы персональных компьютеров (*BOINC*, <http://boinc.berkeley.edu>, *XtremWeb*, <http://www.xtremweb.net>, и др.) и ресурсы специально выделенных для этого высокопроизводительных систем (*Globus Toolkit*, <http://www.globus.org/toolkit>, *UNICORE*, <http://www.unicore.eu>, и др.).

Грид-системы, использующие программное обеспечение первой группы, часто называются волонтерскими или добровольными. В гридах, состоящих из специально выделенных высокопроизводительных компьютеров, применяется вторая группа промежуточного программного обеспечения. В большинстве реализаций программного обеспечения этой группы можно выделить пять основных компонентов: меха-

низм безопасности, инструментарий для запуска задач или программные интерфейсы, планировщик задач, средства доставки данных и средства мониторинга. Перечисленные компоненты не являются обязательными, могут различаться по функциональности и сложности, частично привлекать сторонние компоненты. В целом, состав компонент промежуточного программного обеспечения диктуется спецификой решаемых задач, аппаратной организацией, политикой доступа и другими обстоятельствами.

Для многих программ этого класса характерной является архитектура обеспечения аутентификации и авторизации пользователей, работа которой основана на инфраструктуре открытого ключа в соответствии со стандартом X.509 (подробное описание стандарта и другую документацию можно найти по ссылке <http://datatracker.ietf.org/wg/pkix>). Для доступа к ресурсам грид пользователь должен получить цифровой сертификат от удостоверяющего центра. Затем при запуске своей программы в грид необходимо также иметь так называемый сертификат-посредник (*proxy certificate*), который будет использоваться при всех дальнейших операциях в системе. Описание всего хода исполнения задания (порядок выполнения вычислений, входные параметры, пересылка входных данных и пр.) требует использования специфичных для каждого конкретного промежуточного программного обеспечения набора средств. Так, в выбранной нами системе *Globus Toolkit* применяется язык описания задач (*Job Description Language, JDL*), который предназначен для задания параметров запуска вычислительного кода, доставки исходных данных и других действий. Подобные языки нередко бывают плохо документированы, а также имеют тенденцию меняться со сменой версии промежуточного программного обеспечения. Таким образом, использование грид нельзя назвать простым, ввиду необходимости свободного владения набором специфиче-

* Работа выполнена при поддержке грантов РФФИ № 11-05-00988-а, 11-07-12045-офи-м.

ских приемов взаимодействия с распределенной вычислительной системой. Существует ряд программных решений, которые упрощают взаимодействие пользователя и грид: как инструментарию для создания грид-приложений [2], так и интерфейсы для решения конкретных ресурсоемких задач в грид. В качестве примера приведем портал *QuakeSim* (<http://www.quesim.org>), где организован доступ к коллекции веб-служб для доступа к источникам данных, вычислительных программ, средств визуализации и создания карт. Естественно, что такой способ использования грид-систем не допускает каких бы то ни было модификаций со стороны пользователей. Тем не менее, именно этот подход, по нашему мнению, предоставляет максимально удобный способ взаимодействия пользователя и грид.

Нами разработано грид-приложение, скрывающее большую часть технологических особенностей вычислений в грид. Оно состоит из веб-интерфейса и средств стыковки с промежуточным программным обеспечением. Веб-интерфейс предоставляет возможность редактировать параметры вычислений, оценивать время выполнения задания, отслеживать его выполнение, а также возможность проанализировать полученные данные. Помимо удобств, использование веб-интерфейса обеспечивает дополнительную безопасность ввиду невозможности выполнения произвольного кода пользователем. Это позволяет отказаться от аутентификации пользователей мерами промежуточного программного обеспечения для грид и скрыть от пользователя детали запуска вычислений в грид. Средства стыковки с грид-инфраструктурой состоят из набора программ, которые обеспечивают упаковку входных данных, запуск вычислений в грид, сбор результатов. Помимо этого, средства стыковки обеспечивают слежение за выполнением запущенной задачи, а также занимаются сбором результатов выполнения.

В качестве тестового примера нами была взята задача определения анизотропных сейсмических свойств коры и верхней мантии по волновым формам обменных волн. Подробности обработки данных и параметризации модели описаны в другом месте [3]. В нашем случае целевая функция

$$C(m) = \sum_{i < N} (d_i - s_i(m))^2$$

определена как среднеквадратичное отклонение N наблюдаемых сейсмограмм d_i от синтетических s_i , рассчитанных для совокупности параметров m . В наших расчетах была применена модель слоистого полупространства. В такой постановке задача уже была адаптирована для запуска в грид [4]. Для проведения пробных вычислений мы использовали одну из приведенных там моделей (см. таблицу в процитированной статье), однако теперь значения анизотропных параметров η второго и третьего слоев также включены в число искомым параметров (см. таблицу).

Эта модель была использована для построения волновых форм приемных функций, которые в дальнейшем рассматривались как данные наблюдений. Расчет целевой функции производился на равномерной прямоугольной сетке со следующими параметра-

ми: коэффициент анизотропии S -волн – в пределах от 0 до 0,05, с шагом 0,01, параметр η – в пределах 1,00-1,05, с шагом 0,01. Толщины слоев изменялись в пределах от 50 до 100 км с шагом 10 км, азимут оси анизотропии – от 0 до 175 градусов с шагом 5 градусов. Запуск тестовой задачи показал, что использование созданной нами грид-системы (описание системы приведено далее) ускоряет расчеты в 30-50 раз, сводя время вычислений до нескольких часов. При использовании одного компьютера для выполнения этого объема работ нам понадобилось бы больше недели.

Для взаимодействия с пользователем используется веб-интерфейс, в котором можно выделить три основных компонента: задание параметров вычислений, мониторинг выполнения задания, а также визуализацию полученных результатов. Редактирование параметров расчетов реализовано в виде динамической веб-страницы, содержащей ряд текстовых полей для заполнения. Помимо этого веб-страница содержит управляющие элементы, с помощью которых осуществляются сохранение введенных параметров и загрузка ранее сохраненных, загрузка данных, запуск выполнения задания, оценка времени выполнения расчетов.

Адаптация вычислительной программы под грид-приложение была реализована с помощью системы сборки пакетов-задач. Каждый такой пакет (сжатый *tar*-файл) состоит из исполняемого файла вычислительной задачи, исходных данных, файла конфигурации, а так же скриптов, необходимых для обеспечения запуска и мониторинга. Этот пакет передается для запуска на вычислительных узлах с помощью средств, предоставляемых используемым промежуточным программным обеспечением. В нашем случае это *GRAM (Grid Resource Allocation Manager)* – компонент инструментария *Globus Toolkit*, задача которого состоит в получении заданий на стороне ресурса грид, отправки их локальному менеджеру ресурсов на выполнение, мониторинга исполнения, а также, при необходимости, отмены задания.

Для удобства хранения, передачи и последующей обработки результаты расчета в процессе выполнения хранятся в реляционной базе данных. Значения многомерной функции в базе данных невозможно упорядочить простым индексированием, имея в виду, что при анализе данных нам приходится делать выборки вдоль различных направлений в пространстве аргументов. Для организации оперативного доступа к данным в такой базе необходимо использовать механизм многомерного индексирования и, следовательно, нам нужна система управления базами данных, поддерживающая такой механизм. В нашем проекте использовалась система управления базами данных *MySQL*. Следует отметить, что построение многомерных индексов может занимать достаточно много времени, в нашем случае – несколько часов. Файл базы данных с результатом расчетов можно загрузить на локальный компьютер для последующего анализа. Однако созданный интерфейс предоставляет возможность первичного просмотра результатов в виде набора графиков, отображающих одномерные и двумерные сечения целевой функции в точке абсолютного минимума.

Параметры модели для проведения проверочных расчетов

Номер слоя	V_P , км/с	V_S , км/с	ρ , г/см ³	Толщина H , км	Азимут ϕ , град.	Коэффициент анизотропии α_s	Коэффициент η
1	6,4	3,7	2,9	40	–	0,00	1,0
2	8,1	4,5	3,3	60	30	0,04	1,03
3	8,1	4,5	3,3	100	100	0,02	1,03
4	8,5	4,72	3,4	∞	–	0,00	1,0

Моделирование мы проводили на специально развернутой грид-инфраструктуре, реализованной на аппаратной базе Института физики Земли им. О.Ю. Шмидта РАН, ВИНТИ РАН и Геофизического центра (ГЦ) РАН. Кластер в ИФЗ представляет собой блэйд-систему на основе технологии *Intel Modular Server*. Каждый вычислительный модуль несет 2 процессора *Intel Xeon E5520* и 12 гигабайт оперативной памяти. Связь между узлами обеспечена посредством гигабитного *Ethernet*. Система хранения позволяет предоставить одновременный доступ к данным. Кластер в ВИНТИ являет собой пять стандартных серверов на базе процессоров *Intel Xeon E5430*. Система хранения с разделяемым доступом не предусмотрена аппаратно и реализована программно. Сервер в ГЦ РАН – стандартный сервер, выполняющий административные функции, включая доступ к интерфейсу. Сервер обеспечивает доступ к 1 терабайту памяти на жестких дисках с обеспечением надежного хранения данных (*RAID 1*).

Грид-система построена на основе промежуточного программного обеспечения *Globus Toolkit*. Функционирование настоящего проекта потребовало установки следующих компонент: удостоверяющего центра, хранилища реквизитов, файлового хранилища, локального менеджера вычислительных ресурсов, а также локального планировщика задач. Удостоверяющий центр представляет собой набор инструментов для поддержки инфраструктуры открытых ключей (*public key infrastructure*) и обеспечивает возможность аутентификации пользователей и сервисов в грид. В настоящей реализации используется набор инструментов от проекта *OpenSSL* (<http://www.openssl.org>).

Так как веб-интерфейс исключает взаимодействие пользователя с системами безопасности грид-инфраструктуры, то все вычисления реализуются от имени одного специального пользователя. Для этого потребовалось реализовать дополнительные механизмы взаимодействия между грид-приложением и промежуточным программным обеспечением. Это связано с тем, что инструментарий *Globus Toolkit* для процессов аутентификации пользователей и сервисов основан на так называемых сертификатах-посредниках, для создания которых может быть задействован сервис *MyProxy*. Для размещения данных и доступа к ним применяется протокол *GridFTP*, который является расширением протокола передачи

данных *FTP*, с переработанной системой контроля доступа и рядом других значительных улучшений. Безопасность *GridFTP* основана на средствах *Grid Security Infrastructure (GSI)*, которые, в свою очередь, опираются на сертификаты инфраструктуры открытого ключа.

Задачи на грид-узле выполняются и останавливаются по сигналам от локального менеджера вычислительных ресурсов. Это программное обеспечение работает только в пределах определённого кластера и обеспечивает непосредственный запуск и отслеживание вычислительных программных процессов. В тестовой грид-системе на всех вычислительных кластерах используется менеджер *TORQUE* (<http://www.adaptivecomputing.com/products/open-source/torque>).

Все службы грид-инфраструктуры (кроме *GridFTP*), веб-интерфейса и средства стыковки работают на виртуальных машинах — экземплярах операционной системы, запущенных параллельно с основной операционной системой на одном и том же оборудовании. Аппаратная поддержка виртуализации позволяет свести к минимуму накладные расходы на такую конфигурацию, что, в целом, позволяет экономить на оборудовании и более полно его использовать. Идея построения сервисов грид на виртуальных машинах может быть расширена в рамках механизма виртуальной частной сети. В этом случае появляется возможность переноса виртуальных машин средствами обеспечения распределённых вычислений не только внутри одного вычислительного кластера, но между грид-узлами, размещёнными в виртуальной частной сети.

На основе проведенных тестовых вычислений в рамках созданного приложения показано, что предложенный подход не только критически сокращает время выполнения расчетов за счет применения грид-технологий, но и существенно упрощает использование вычислительных ресурсов грид-инфраструктуры за счет реализации проблемно-ориентированного веб-интерфейса. Естественно, представленная реализация приложения не является оптимальной и может быть существенно улучшена с точки зрения как применения более эффективной техники построения решения, так и использования более легковесного промежуточного программного обеспечения.

СПИСОК ЛИТЕРАТУРЫ

1. Tarantola A. Inverse problem theory; methods for data fitting and model parameter estimation. – Elsevier, 1987. – 613 p.
2. Allen G., Davis K., Goodale T., Hutanu A., Kaiser H., Kielmann T., Merzky A., Nieuwpoort R. van, Reinefeld A., Schintke F., Schütt T., Seidel E., Ullmer B. The Grid Application Toolkit: Towards Generic and Easy Application Programming Interfaces for the Grid // Proceedings of the IEEE. – 2005. – Vol. 93(3). – P. 53–550.
3. Vinnik L. P., Aleshin I. M., Kiselev S. G., Kosarev G. L., Makeyeva L. I. Depth localized azimuthal anisotropy from SKS and P receiver functions: The Tien Shan // Geophys. J. Int. – 2007. – № 169. – P. 1289–1299 doi: 10.1111/j.1365-246X.2007.03394.x
4. Алешин И.М., Мишин Д.Ю., Жижин М.Н., Корягин В.Н., Медведев Д.П., Новиков А.М., Перегудов Д.В. Применение распределенных вычислительных систем при определении параметров сейсмической анизотропии коры и верхней мантии // Геофизические исследования. – 2009. – Т. 10, № 4. – С.34–47.

Материал поступил в редакцию 11.02.13.

Сведения об авторах

АЛЕШИН Игорь Михайлович – кандидат физико-математических наук, зав. сектором Института физики Земли им. О.Ю. Шмидта РАН, Москва

E-mail: ima@ifz.ru

КОРЯГИН Владимир Николаевич – кандидат экономических наук, старший научный сотрудник Института физики Земли им. О.Ю. Шмидта РАН, Москва

E-mail: admin@ifz.ru

СУХОРОСЛОВ Олег Викторович – кандидат технических наук, старший научный сотрудник Института проблем передачи информации РАН, Москва

E-mail: oleg.sukhoroslov@gmail.com

ХОЛОДКОВ Кирилл Игоревич – инженер Института физики Земли им. О.Ю. Шмидта РАН, Москва

E-mail: keir@ifz.ru

ШОГИН Алексей Николаевич (1959-2013 гг.) – кандидат технических наук, зав. Центром развития информационных систем ВИНТИ РАН, Москва

ДОКУМЕНТАЛЬНЫЕ ИСТОЧНИКИ ИНФОРМАЦИИ

УДК 025.48 : [3+7/9] ИНИОН

Р.Р. Мдивани

Тезаурусы ИНИОН РАН по социальным и гуманитарным наукам

Описаны результаты научно-практической работы по созданию Серии: Информационно-поисковые тезаурусы по социальным и гуманитарным наукам Автоматизированной информационной системы Института научной информации по общественным наукам РАН: проблемы формирования лексического массива; особенности терминологии и ведение отраслевых тезаурусов; роль структуры тезаурусов как средства расширения возможностей навигации в концептуальном пространстве документов библиографического фонда.

Ключевые слова: общественные науки, информационно-поисковые тезаурусы, ведение тезаурусов, структура тезаурусов, тезаурусные расширения, навигация в информационных ресурсах

Основным лингвистическим средством навигации в массиве научных ресурсов Автоматизированной информационной системы по общественным наукам (АИСОН) Института научной информации по общественным наукам ИНИОН РАН является Серия: Информационно-поисковые тезаурусы по социальным и гуманитарным наукам.

В настоящее время работа над отраслевыми выпусками Серии завершена и включает 10 тезаурусов, изданных в виде книг и на компакт-дисках (в скобках указано число терминов): «Экономика. Демография» (5 тыс.); «Философия» (3,3 тыс.); «Правоведение» (8,1 тыс.); «Социология» (2,3 тыс.); «Политология» (6,9 тыс.); «Языкознание» (3 тыс.); «Религиоведение» (1,8 тыс.); «Литературоведение» (1,1 тыс.); «Науковедение» (6,6 тыс.); «Исторические науки. Т.1. История. Археология (10 тыс.); Т.2. Этнология. Антропология (5 тыс.)» [1].

Программа специализированной лингвистической базы данных, созданной в ИНИОН, обеспечивает проверку парадигматических связей и формирование Алфавитного лексико-семантического указателя, Систематического указателя по Рубрике ИНИОН [2], Классификационного указателя по Рабочим таблицам для массовых библиотек Библиотечно-библиографической классификации [3] и Пермутационного указателя ключевых слов (KWIC). Кроме того, в состав тезаурусов включаются вспомогательные фасеты (алфавитные списки лексических единиц одной категории: «Языки мира», «Литература отдельных стран и народов», «Археологические культуры», «Народы мира», «Политико-географический фасет» и др.).

Как отмечал Дж. Солтон в своей классической работе, «одним из главных недостатков, присущих те-

заурусу, является то, что его надо вести» [4, с. 494-495]. Дж. Солтон предлагает по меньшей мере четыре способа корректировки тезауруса:

1) использовать исходный тезаурус без изменений и на возросшем массиве документов (при этом ухудшение эффективности поиска составляет примерно 4%);

2) добавить новые термины из влитых в массив документов;

3) ввести новые классы для новых терминов;

4) полностью перестроить тезаурус, создав новую классификацию на базе пополненного словаря.

Практика разработки и ведения тезаурусов Серии показала целесообразность ориентации на второй способ. Это обусловлено следующими обстоятельствами. Во-первых, на начальном этапе разработки тезаурусов использовался репрезентативный массив накопленной за определённый период отраслевой нормализованной лексики. Во-вторых, для пополнения терминами исходного состава тезаурусов и формирования словарных статей проводился анализ индексирования документов с учётом информативности отбираемых терминов.

Работа с большим массивом лексики ПОДов, представленных в базе данных отечественных и иностранных документов, показала относительно невысокую долю появления новых дескрипторов. Это связано со спецификой социальных и гуманитарных наук, которая обуславливает определённый консерватизм понятийного состава тезаурусов по общественным наукам. В противоположность естественным наукам, обладающим синхронно движущимся исследовательским фронтом и кумулятивным подходом к исследуемым фактам, в социальных и гуманитарных

науках существует множество различных школ, достаточно независимых друг от друга. Эти школы, имеющие дело с одними и теми же социальными и гуманитарными проблемами, не приходят при этом к единому их толкованию. Темы, которые обсуждались веками, могут возникнуть вновь, и даже при рассмотрении новых проблем обществовед должен ссылаться на старые материалы и идеи. Эта специфика обуславливает существование больших ретроспективных баз данных, так как нет объективных критериев, которые позволяют сказать, что документ устарел и его можно исключить из базы знаний общественных наук [5, с. 265].

В социальных и гуманитарных науках развитие научной мысли идёт (если не учитывать конкретных данных) за счёт различных интерпретаций и уточнения смысла терминов при сохранении их плана выражения и конвенционального значения. Так, например, обстоит дело с использованием в современном политическом дискурсе таких слов-понятий, как «суверенитет», «демократия», «права человека» [6, с. 20-34]. Поэтому можно утверждать, что «в сознании не рефлексирующего по поводу своих языковых способностей носителя языка значение языкового выражения выступает как непосредственная данность» [7, с. 10].

Программное обеспечение лингвистической базы данных, конечно, облегчало работу над тезаурусами, но вся интеллектуальная работа по отбору терминов, формированию словарных статей и систематизации терминов по Рубрикатору ИНИОН и по Рабочим таблицам для массовых библиотек ББК проводилась «вручную». В определённой степени работу осложнял неизбежный «человеческий фактор», с которым связаны, как справедливо отмечает Н.В.Лукашевич, уровень квалификации и качество работы индексаторов, а также трудности определения сложных семантических отношений и снятия омонимии [8, с. 10-12]. Естественно, что при работе «в ручном режиме» снятие омонимии не является проблемой. В социальных науках (в частности, в науках экономических и исторических), база данных которых имеет довольно большую глубину ретроспекции, актуальна проблема синонимии. Авторы публикаций в определённой степени более свободны в употреблении того или иного синонимичного термина, чем специалисты в области естествознания и техники. Ретроспективный анализ базы данных показал, что превалирование того или иного синонимичного термина в документопотоке является характерным для определённых периодов развития науки. Например, термин «товары широкого потребления» использовался в период 1980-1994 гг. (103 документа); в последующие годы он был заменён термином «потребительские товары» (390 документов). В иных случаях авторы публикаций, переосмысляющие прошлое, выражают своё идеологизированное отношение к анализируемому предмету, пользуясь различными формулировками названия того или иного события. Примером может служить массив документов, индексированных термином «Великая Октябрьская социалистическая революция», и публикации, индексированные как «Октябрьская революция 1917 г». Чтобы сохранить возможность ретроспективного поиска по теме без

проблем переиндексирования документов в БД, такие термины связываются в тезауусе перекрёстной ссылкой «см. также». При этом соответствующие парадигматические связи включаются в словарную статью одного из них. Такие случаи встречаются нечасто и выявляются на основе анализа документов в БД в процессе разработки и ведения тезауусов.

Всё изложенное в определённой степени влияет на темпы пополнения лексики тезауусов много лет функционирующей базы данных. Обязательным при этом является строгое соблюдение общепринятого в информатике принципа смыслового соответствия, поскольку в системе понятий «данные – информация – знания» список терминов (дескрипторов и аскрипторов) является носителем формы, тогда как соответствующее содержание – это информация и знания, которые выдаются пользователю системы благодаря структурированному в результате индексирования информационному ресурсу.

Следует отметить, что «ручное» создание информационно-поискового тезаууса и его дальнейшее обновление (ведение) с учётом развития научной области знания и изменения терминологии связано с существенными затратами интеллектуального труда. Однако тезауус конкретной предметной области необходим, так как универсальные тезауусы, энциклопедии и словари плохо представляют концептуальные связи богатой и специфичной терминологии научных дисциплин [9, с. 201-202].

Что касается пополнения тезауусов, то оно проводится периодически на основе анализа индексирования документов, поступивших в базу данных, с учётом информативности новых терминов и последующим определением их парадигматических связей. Включение новых терминов в представленные в электронном виде тезауусы позволяет избежать весьма дорогостоящего процесса бумажных переизданий.

С точки зрения эффективного поиска в массиве научных ресурсов, представленных в автоматизированной библиографической базе данных, тезауусы служат средством навигации в концептуальном поле документов фонда, поэтому структура каждого отраслевого тезаууса ориентирована на оптимизацию формулировки запроса. Это требует включения в тезауус его собственных навигационных средств, дающих возможность выявления дополнительных концептуальных отношений между входящими в его состав терминами: расширение запроса с помощью тезаууса обычно увеличивает эффективность поиска. Поэтому тезауусы АИСОИ строятся по принципу расширения за счёт включения указателей, синонимических с точки зрения понятийного состава.

1. Алфавитный лексико-семантический указатель содержит дескрипторы с соответствующими семантическими (парадигматическими) отношениями с аскрипторами и другими дескрипторами. Наличие семантических связей, представленных в словарной статье дескриптора, позволяет расширить формулировку запроса, что в большинстве случаев влияет на полноту выдачи, но не исключает и влияния на степень релевантности.

Кроме того, поскольку база библиографических данных АИСОИ построена как инверсный файл, пользователь может использовать в запросе булевы

операторы, добавляя при необходимости лексику из различных вспомогательных фасетов: названия стран, народов мира и т.п.

2. Систематический указатель включает дескрипторы, классифицированные в алфавитном порядке по рубрикам Рубрикатора АИСОН, который представляет собой иерархическую классификационную таблицу отраслей социальных и гуманитарных наук. Важно, что три верхних уровня Рубрикатора АИСОН совместимы с Государственным Рубрикатом НТИ и с рубрикатом других систем, базирующихся на соответствующих разделах Государственного Рубрикатора, что обеспечивает взаимодействие с ресурсами других библиотек и центров. Наличие тематических рубрик выполняет двойную функцию. Во-первых, рубрика, если она соответствует или близка интересам пользователя, облегчает просмотр терминов, отбираемых для запроса. При этом рекомендуется последующее обращение к ним в Алфавитный указатель, где представлены расширенные за счёт словарных статей концептуальные отношения каждого отобранного из Систематического указателя термина. Во-вторых, индексы рубрик могут использоваться непосредственно при формулировке запроса, так как ими индексированы конкретные документы в базе данных. Во всяком случае, просмотр документов по теме рубрики может быть полезным при желании обеспечить высокую степень полноты выдачи.

3. Классификационный указатель содержит дескрипторы, распределённые по рубрикам Библиотечно-библиографической классификации (ББК). Этот вспомогательный указатель предназначен для читателей и работников массовых библиотек, которые могут использовать хорошо знакомую им систему для выбора дескрипторов и формирования запросов при обращении к библиографической базе данных ИНИОН; следует отметить, что использование кодов ББК непосредственно для поиска документов в АИСОН невозможен. Поскольку число дескрипторов того или иного тезауруса значительно превосходит число рубрик соответствующего отраслевого отдела, использованы методические указания к Таблицам ББК, допускающие выделение частных, более дробных делений путём присоединения к индексам соответствующих разделов дополнительных индексов и типовых делений справа, что сделало терминологию тезаурусов обозримой. Например, в отделе **60.5 Социология** представлен раздел **60.54 Социология групп**. Методические указания к этому разделу и Алфавитно-предметный указатель позволили с помощью дополнительных индексов выделить 8 подразделов, по которым классифицировались соответствующие дескрипторы тезауруса. Например:

- 60.54 Социология групп. Социоморфология (19 терминов)
- 60.54-1 Теория больших групп (35 терминов)
- 60.54-2 Социально-демографические группы (35 терминов)
- 60.54-3 Этнические, этно-социальные общности. (75 терминов) и т.п.

Всего по таким подразделам было классифицировано 239 дескрипторов.

4. Пермутационный указатель ключевых слов в контексте (KWIC) выполняет вспомогательную

функцию по отношению к Алфавитному лексико-семантическому указателю и содержит дескрипторы и аскрипторы без словарных статей. В нём по однословным терминам и отдельным ключевым словам, входящим в контекст терминологических словосочетаний, формируются словарные гнезда. Словарное гнездо компактно отражает различные терминологические словосочетания, семантически связанные с «заглавным» ключевым словом, что существенно облегчает поиск дескрипторов, представленных в основном (Лексико-семантическом) указателе отраслевого тезауруса, упорядоченном по алфавиту.

Поскольку в общественных науках термины-словосочетания употребляются достаточно широко, был проведён статистический анализ ключевых слов, не стоящих в алфавитном ряду на первом месте. Например, понятие «рабочая сила» так представлено в Пермутационном указателе тезауруса «Правоведение» (самостоятельного дескриптора в нём нет):

внешняя миграция	рабочей силы
внутренняя миграция	рабочей силы
движение	рабочей силы
международная миграция	рабочей силы
миграция	рабочей силы.

Ясно, что вхождение ключевого слова в контекст словосочетания определяет его роль в структуре терминов Алфавитного лексико-семантического указателя. Более важно, что Пермутационный указатель выделяет ключевые слова, которые могут использоваться самостоятельно для поиска по текстам (в библиографии – по заголовкам и аннотациям) с применением стандартных операций приведения производных слов и словоформ к общей основной форме путём удаления аффиксов, т.е. стемминга и лемматизации. Это позволяет рассматривать Пермутационный указатель как некое расширение входов в общее концептуальное поле тезауруса.

Общее число дескрипторов и аскрипторов, представленных в алфавитных лексико-семантических указателях 10-ти тезаурусов, составляет около 50 тыс. терминов. Анализ соответствующих пермутационных указателей показал средний результат расширения понятийного объёма отраслевых тезаурусов на 20-25% за счёт ключевых слов, не стоящих на первом месте в составе терминов-словосочетаний и не являющихся однословными терминами.

Завершение работы над отраслевыми тезаурусами позволило использовать ещё один способ расширения возможности навигации в библиографической базе данных системы. В Центре информатизации ИНИОН на базе отраслевых тезаурусов создан сводный информационно-поисковый тезаурус по социальным и гуманитарным наукам, представленным в библиографической БД АИСОН. В этом тезаурусе термины даны в алфавитном порядке, причём совпадающие термины повторяются с указанием отраслевой принадлежности без словарных статей. Однако при обращении к тому или иному дескриптору на экран выводится его словарная статья, причём каждый из дескрипторов имеет парадигматические связи, специфические для отрасли, к которой он относится. Это не исключает наличия в связях одинаковых дескрипторов. Например, дескриптор ВИКТИМОЛОГИЯ представлен в трёх отраслях:

Науковедение (Н)	Правоведение (Пр)	Социология (Соц)
ВИКТИМОЛОГИЯ ВТ криминология РТ девиантность насилие социальная педагогика социальные проблемы	ВИКТИМОЛОГИЯ ВТ криминология РТ виктимизация населения жертва потерпевшие преступления	ВИКТИМОЛОГИЯ РТ криминология насилие преступность социология преступности убийство

Таким образом, пользователь может получить расширенную информацию о парадигматических связях термина «виктимология». Для наглядности мы свели термины в общую словарную статью (в настоящее время программа этого не делает). Буквами в скобках обозначены отрасли:

ВИКТИМОЛОГИЯ

- криминология (Н, Пр, Соц)
- виктимизация населения (Пр)
- девиантность (Н)
- жертва (Пр)
- насилие (Н, Соц)
- потерпевшие (Пр)
- преступления (Пр)
- преступность (Соц)
- социальная педагогика (Н)
- социальные проблемы (Н)
- социология преступности (Соц)
- убийство (Соц).

Такое расширение семантического поля дескриптора существенно улучшает возможность поиска по межотраслевым проблемам.

Многолетняя эксплуатация лингвистических средств АИСОН ИНИОН РАН, отражающих тематику научных публикаций за период с 80-х гг. прошлого века, подтвердила эффективность этих средств, поскольку «мощность языка запросов соответствует мощности языковых средств» [10, с. 220]. В результате, семантические метаданные, относящиеся к контенту документа (словарь понятий, класс рубрикатора), в сочетании с внешними структурированными данными (элементы библиографического описания: автор, тип документа, год издания и т.п.) обеспечивают эффективное функционирование базы данных Автоматизированной информационной системы по общественным наукам.

СПИСОК ЛИТЕРАТУРЫ

1. Тезаурусы ИНИОН по социальным и гуманитарным наукам: Серия / РАН. ИНИОН. Центр информатизации. Фундаментальная библиотека. – М.: ИНИОН РАН.
 - а) Информационно-поисковый тезаурус ИНИОН по экономике и демографии / Архангельская В.А. при участии Мдивани Р.Р. (Классификационный указатель ББК); ред. Мдивани Р.Р. – М.: ИНИОН РАН, 2001. – 492 с.
 - б) Информационно-поисковый тезаурус ИНИОН по философии / Гинзбург Б.П. при участии Мдивани Р.Р. (Классификационный указатель по ББК); ред. Мдивани Р.Р., Серебряная Е.И. – М.: ИНИОН РАН, 2002. – 324 с.
 - в) Информационно-поисковый тезаурус ИНИОН по правоведению / Базарнова С.В., Шевцова С.О. при участии Мдивани Р.Р. (Классификационный указатель по ББК); ред. Крысанова Н.В., Мдивани Р.Р., Слива А.И. – М.: ИНИОН РАН, 2002. – 808 с.
 - г) Информационно-поисковый тезаурус ИНИОН по социологии / Гинзбург Б.П. при участии Мдивани Р.Р. (Классификационный указатель по ББК); ред. Гирко Л.В., Ли О.С., Мдивани Р.Р. – М.: ИНИОН РАН, 2003. – 240 с.
 - д) Информационно-поисковый тезаурус ИНИОН по политологии / Базарнова С.В., Мдивани Р.Р., Слива А.И., Шевцова С.О.; ред. Верченнов Л.Н., Ильин М.В., Магай Е.В., Мдивани Р.Р., Мелешкина Е.Ю.; отв. ред. Пивоваров Ю.С. – М.: ИНИОН РАН, 2007. – 632 с.
 - е) Информационно-поисковый тезаурус ИНИОН по языкознанию / Смиренский В.Б.; ред. Антонюк Г.С., Магай Е.В.; отв. ред. Мдивани Р.Р., Стрельченко П.Я. – М.: ИНИОН РАН, 2007. – 200 с.
 - ж) Информационно-поисковый тезаурус ИНИОН по религиоведению / Денискин А.А. при участии Мдивани Р.Р. (Классификационный указатель по ББК); ред.: Кимелев Ю.А., Магай Е.В., Мдивани Р.Р. – М.: ИНИОН РАН, 2008. – 164 с.
 - з) Информационно-поисковый тезаурус ИНИОН по литературоведению и фольклористике / Курашова М.Я., Смиренский В.Б., Швырёва Г.П., Этова О.В. при участии Мдивани Р.Р. (Классификационный указатель по ББК); ред.: Магай Е.В.; отв. ред. Мдивани Р.Р. – М.: ИНИОН РАН, 2010. – 112 с.
 - и) Информационно-поисковый тезаурус ИНИОН по науковедению: наука и образование / Магай Е.В., Мдивани Р.Р., Хадияров Г.Г.; ред.: Макешин Н.И., Мдивани Р.Р. – М.: ИНИОН РАН, 2011. – 496 с.
 - к) Информационно-поисковый тезаурус ИНИОН по историческим наукам: Т.1. История. Археология / Базарнова С.В., Матвеева Е.Ю. при участии Громовой Д.Н., Окорочковой И.Л., Русаковой М.Ю.; ред.: Магай Е.В., Мдивани Р.Р.; отв. ред. Пивоваров Ю.С., Шевырин В.М. – М.: ИНИОН РАН, 2012. – 610 с.; Т.2. Этнология. Антропология / Магай Е.В., Мдивани Р.Р.,

- Уварова Т.Б.; ред.: Мдивани Р.Р., Уварова Т.Б. – М.: ИНИОН РАН, 2012. – 300 с.
2. Рубрикатор автоматизированной информационной системы по общественным наукам (АИСОН): Классификационная таблица / РАН. ИНИОН. Фундам. б-ка. Отд. каталогизации и электрон. каталогов, Отд. науч.-библиогр. информ.; сост. Антонюк Г.С. и др.; отв. ред. Глинский В.А., Слива А.И. – М., 2000. – 192 с.
 3. Библиотечно-библиографическая классификация: Рабоч. табл. для мас. б-к. – М.: Либерия, 1997. – 688 с.
 4. Солтон Дж. Динамические библиотечно-информационные системы. – М.: изд-во «Мир», 1979. – 560 с.
 5. Хобом Х.-К. Информация и документация по социальным наукам // Теория и практика общественно-научной информации. – М.: ИНИОН, 2002. – Вып. 17. – С. 262-280.
 6. Ильин М.В. Слова и смыслы. Опыт описания ключевых политических понятий. – М.: Российская политическая энциклопедия (РОССПЭН), 1997. – 431 с.
 7. Рубашкин В.Ш., Лахути Д.Г. О языке и средствах диалога с экспертом предметной области в системе ведения семантического словаря // Научно-техническая информация. Сер. 2. – 2002. - № 7. – С. 7-15.
 8. Лукашевич Н.В. Тезаурусы в системе информационного поиска. – М., 2011. – 512 с.
 9. Маннинг К.Д., Рагхаван П., Шютце Х. Введение в информационный поиск / пер. с англ. – М.: ООО «И.Д. Вильямс», 2011. – 528 с.
 10. Адамович И.М., Заикин М.Ю., Земсков Д.В., Пешков А.Н. Локальные информационно-упорядоченные подпространства в глобальном пространстве Интернет // Системы и средства информатики / Ин-т проблем информатики РАН. – М.: Наука, 2005. – Вып. 15. – С. 219-232.

Материал поступил в редакцию 29.04.13.

Сведения об авторе

МДИВАНИ Роберт Робертович - кандидат филологических наук, старший научный сотрудник Института научной информации по общественным наукам (ИНИОН) РАН, ведущий научный сотрудник, руководитель Группы разработки и ведения тезаурусов Центра информатизации ИНИОН, Москва
E-mail: mamdi13@mail.ru

Индекс цитирования, учитывающий скрытую диффузию научных знаний

Предложен новый наукометрический показатель, который оценивает диффузию знаний и имеет две составляющие: первая эквивалентна обычному индексу цитирования, т. е. описывает видимую диффузию научных знаний; вторая – отражает скрытую диффузию научных знаний и выражается через количество неявных цитирований. Практическая ценность предложенного показателя заключается в том, что он позволяет легко идентифицировать скрытых инициаторов научного мейнстрима. Отличительная особенность таких ученых – большое значение предложенного индекса цитирования и малое значение обычного индекса цитирования.

Ключевые слова: наукометрия, индекс цитирования, неявное цитирование, диффузия знаний, инициатор мейнстрима

ВВЕДЕНИЕ

Для оценивания полезности научной деятельности сегодня совместно с экспертными заключениями все чаще используют и наукометрические показатели. В первую очередь наукометрические показатели востребованы для оценки фундаментальных исследований, результаты которых непосредственно не связаны с экономическим эффектом. Фундаментальные разработки направлены на развитие науки, поэтому их востребованность оценивают через отзыв научного сообщества на публикации с результатами исследований. Формально этот отзыв выражают индексом цитирования – суммарным количеством публикаций, ссылающихся на рассматриваемые работы. Варианты индекса цитирования модифицируют его следующим образом:

- учитывают личный вклад ученого [1, 2], разделяя количество цитирований между соавторами, число которых в одной статье иногда превышает 3000;
- учитывают авторитетность цитирующего издания, используя импакт-фактор журнала или другой аналогичный коэффициент [3, 4];
- игнорируют самоцитирование или цитирование соавторами [5, 6], что существенно снижает рейтинг "ученого-затворника", публикации которого интересуют только его самого;
- игнорируют повторное цитирование одной работы одним и тем же ученым [7], что уменьшает влияние комплементарного цитирования.

Использование этих и других модификаций не устраняют следующие два недостатка индекса цитирования. Первый недостаток связан с забыванием имен классиков, когда авторы считают, что вклад предшественников настолько хорошо известен любому из соответствующей области науки, что нет смысла об этом упоминать [8]. Соответственно, работы таких ученых высшего ранга быстро перестают

цитировать. Как правило, этот процесс ускоряет включение научных результатов классиков в учебники и учебные пособия. Поэтому стремительно возрастает диффузия этих научных знаний, т.е. их адаптация и использование в широком спектре научных и инженерных исследований и разработок [9].

Второй недостаток индекса цитирования заключается в низком рейтинге концептуальных работ с принципиально новыми идеями. Это происходит из-за достаточно распространенного сокрытия первоисточников, т.е. невключения концептуальных работ предшественников в перечень цитируемой литературы. Создание нового научного результата – это процесс, включающий этапы [8]: 1) генерирования идеи и разработки концепции, 2) модификации идеи (концепции) для улучшения точности, быстродействия, упрощения и т.д., 3) продвижения и практического использования. В истории науки есть много примеров, когда работы 2-го и 3-го этапов значительно чаще цитируют, чем концептуальные статьи. Впечатляющим примером, из приведенных в [8] являются статьи А. Фолина [10] и О. Лоури [11]. О. Лоури модифицировал реактив А. Фолина для колориметрического определения белка путем добавления еще одного компонента, что расширило границы применения метода. Сегодня, по данным Google Scholar, статью О. Лоури процитировали 247 480 раз, тогда как работу А. Фолина – 1950. Следовательно, среди 247 480 работ, цитирующих статью [11] и, соответственно, использующих идеи О. Фолина, на которых она основана, почти нет таких, которые явно ссылаются на его статью [10].

Цель настоящей статьи состоит в модификации индекса цитирования за счет разработки механизмов компенсации указанных выше недостатков. Идея заключается в дополнении индекса цитирования составляющей, учитывающей диффузию знаний. В [12] индикатором диффузии знаний выбра-

но цитирование работы. Но это прямое, видимо, использование знаний. Существует и скрытая диффузия знаний, состоящая в использовании идей без прямой ссылки на ее автора, но с возможностью идентификации первоисточника через цепочку цитирований. В приведенном выше примере это цитирование только работы А. Лоури без прямой ссылки на статью О. Фолина. Такие случаи назовем скрытой диффузией знаний, а цепочку ссылок, приводящую к первоисточнику, назовем неявным цитированием. Соответственно, подсчитав каким-то образом число неявных цитирований и добавив его с некоторым весом к обычному индексу цитирования, получим рейтинг ученого с учетом скрытой диффузии знаний. Аналогичные принципы оценивания персонала используются в сетевом маркетинге, когда сотрудник получает баллы как за свои прямые продажи, так и за продажи, которые осуществила завербованная им команда. Здесь персональные продажи сотрудника соответствуют прямому цитированию, а продажи его команды – неявному цитированию. Эти же принципы применяются в спорте для оценки по системе "гол плюс пас" атакующих игроков команды. Соответственно, аналогом гола является прямое цитирование, а паса – неявное.

ФОРМАЛИЗАЦИЯ ПОСТАНОВКИ ЗАДАЧИ

Будем считать известным список из M публикаций в таком формате:

$$\mathbf{A}_j = \langle j, \mathbf{F}_j, \mathbf{T}_j \rangle, \quad j = \overline{1, M}, \quad (1)$$

где $\mathbf{F}_j = \{F_j^1, F_j^2, F_j^3, \dots\}$ – множество номеров публикаций, на которые ссылается работа \mathbf{A}_j ;

$\mathbf{T}_j = \{T_j^1, T_j^2, T_j^3, \dots\}$ – множество номеров публикаций, в которых цитируется работа \mathbf{A}_j .

С математической точки зрения, задача состоит в нахождении отображения $\mathbf{A}_j \rightarrow C_j$ ($j = \overline{1, M}$), где C_j – новый индекс цитирования, который учитывает скрытую диффузию знаний.

УЧЕТ СКРЫТОЙ ДИФфуЗИИ ЗНАНИЙ

Индекс цитирования, учитывающий скрытую диффузию знаний, определим следующим образом:

$$C_j = D_j + \alpha \cdot I_j, \quad (2)$$

где $D_j = |\mathbf{T}_j|$ – обычный индекс цитирования, учитывающий только прямые ссылки ($|\cdot|$ – обо-

значает мощность множества, т.е. количество его элементов);

I_j – индекс неявного цитирования;

$\alpha \in [0, 1]$ – весовой коэффициент важности неявного цитирования.

В случае $\alpha = 0$ новый показатель (2) эквивалентен обычному индексу цитирования. Когда $\alpha = 1$ вклады прямого и скрытого цитирования будут равнозначными.

Диффузию знаний учтем следующим образом. Каждая научная работа содержит как новые знания, так и знания из цитируемых работ. Предположим, что доля заимствованных знаний постоянна. Обозначим ее через $\beta \in (0, 1)$. Рассмотрим простую цепочку цитирований $C \rightarrow B \rightarrow A$, когда работа A ссылается только на работу B , а работа B – только на работу C . Соответственно, β -доля знаний перешла из B в C , а затем β -доля знаний из B перешла в A . Тогда, из C в A перешли знания, доля которых составляет $\alpha = \beta \cdot \beta$, хотя это и не подтверждается прямым цитированием. Такое неявное использование знаний или скрытая диффузия обусловлена тем, что автор включает в перечень литературы наиболее релевантные источники, уровень использования которых превышает некоторый порог [8].

Рассмотрим случай, когда список литературы в B состоит из N источников, среди которых и работа C . Тогда будем считать, что из C в B перешли знания, доля которых составляет $\frac{\beta}{N}$. Обобщая выкладки, получаем такую формулу индекса неявного цитирования:

$$I_j = \sum_{\substack{\forall i \in \mathbf{T}_j \\ j \notin \mathbf{F}_i}} \frac{|\mathbf{T}_i|}{N_i}, \quad (3)$$

где $N_i = |\mathbf{F}_i|$ – длина списка литературы в i -й публикации.

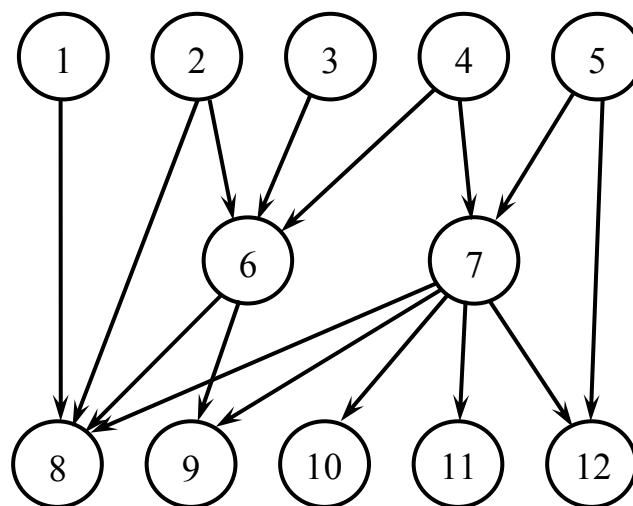
Условие $j \notin \mathbf{F}_i$ в (3) указывает, что в случае прямого цитирования неявное цитирование не учитывается.

Пример. Исходные данные приведены в табл. 1. Соответствующая сеть цитирования изображена на рисунке. Расчет нового индекса цитирования (2) по этим данным иллюстрирует табл. 2. Из нее видно, что публикации № 2, 4, 5 и 6 получили одинаковое количество прямых цитирований. Среди них самый высокий рейтинг имеет публикация №4 в связи с максимальным количеством неявных цитирований, две из которых посредством работы № 6 и пять – посредством работы № 7.

Таблица 1

Исходные данные по цитированию в формате (1)

j	F_j	T_j
1	неизвестно	{8}
2	неизвестно	{6, 8}
3	неизвестно	{6}
4	неизвестно	{6, 7}
5	неизвестно	{7, 12}
6	{2, 3, 4}	{8, 9}
7	{4, 5}	{8, 9, 10, 11, 12}
8	{1, 2, 6, 7}	–
9	{6, 7}	–
10	{7}	–
11	{7}	–
12	{5, 7}	–



Сеть цитирования

Таблица 2

К расчету индекса цитирования (2)

j	1	2	3	4	5	6	7
D_j	1	2	1	2	2	2	5
I_j	0	$\frac{1}{3} = 0,33$	$\frac{2}{3} = 0,67$	$\frac{2}{3} + \frac{5}{2} = 3,17$	$\frac{4}{2} = 2$	0	0
$C_j, \alpha = 0,1$	1	2,03	1,07	2,32	2,2	2	5
$C_j, \alpha = 0,5$	1	2,17	1,33	3,58	3	2	5

СРАВНЕНИЕ С АНАЛОГАМИ

Нами обнаружено, что для ранжирования научных публикаций неявное цитирование используется только в [13] для расчета индекса Хирша отдельной статьи. В этом случае, индекс Хирша статьи составляет h , если на нее ссылаются h работ, каждую из которых процитировано минимум h раз.

Ближайшим аналогом является подход [14], в котором неявные ссылки используются для расчета кумулятивного индекса цитирования патентов. Для этого применяется аналог формулы (2) для всей цепочки цитирований. При этом, чем длиннее цепочка неявного цитирования, тем меньше весовой коэффициент. Принципиальное отличие нашего подхода заключается в учете длины списка литературы при расчете неявного цитирования в (3), тогда как в [14] длина соответствующего списка патентов не учитывается. Игнорирование длины списка литературы может привести к тому, что составляющая от неявного цитирования будет больше, чем если бы имело место прямое цитирование. Например, для фрагмента сети цитирований с вершинами 4, 6, 7 и 8 (см. рис. 1), работа № 8 обеспечила бы 2 балла работе № 4 за неявное цитирование, тогда как в случае прямого цитирования $4 \rightarrow 8$ работа № 4 получила лишь бы 1 балл.

ВЫВОДЫ

Предложен новый наукометрический показатель, оценивающий диффузию знаний. Он имеет две составляющие, первая из которых эквивалентна обычному индексу цитирования, т. е. описывает видимую диффузию научных знаний. Вторая составляющая отражает скрытую диффузию научных знаний и выражается через количество неявных цитирований. Предложенный подход наследует механизмы оценки деятельности персонала в сетевом маркетинге и атакующих игроков в хоккее, футболе и других командных играх спорта.

Практическая ценность предложенного показателя заключается в том, что он позволяет легко идентифицировать креативных ученых, которые сгенерировали новые идеи сподвигнули научное сообщество на создание значительного количества востребованных (высокоцитированных) работ. При этом, сами генераторы идей остались в тени. За такими учеными часто «охотятся» лидерские исследовательские компании, но автоматически их обнаружить было достаточно сложно из-за низких наукометрических показателей. Теперь идентификацию скрытых инициаторов научного мэйнстрима можно формализовать, выявив ученых, которые одновременно имеют большое значение предложенного индекса цитирования и малое значение традиционного индекса цитирования. Практическое применение

предложенного показателя возможно лишь на основе автоматизированного учета научных публикаций, например, на базе систем Web of Knowledge, Scopus, Google Scholar или eLibrary.ru.

СПИСОК ЛИТЕРАТУРЫ

1. Egghe L. Mathematical theory of the h- and g-index in case of fractional counting of authorship // Journal of the American Society for Information Science and Technology. – 2008. – Vol. 59, № 10. – P. 1608–1616.
2. Schreiber M. A modification of the h-index: The h(m)-index accounts for multi-authored manuscripts // Journal of Informetrics. – 2008. – Vol. 2, № 3. – P. 211–216.
3. Buela-Casal G. Assessing the quality of articles and scientific journals: Proposal for weighted impact factor // Psychology in Spain. – 2004. – Vol. 8, № 1. – P. 60–76.
4. Van Noorden R. A profusion of measures // Nature. – 2010. – Vol. 465. – P. 864–866.
5. Garcia-Perez M.A. The Hirsch h index in a non-mainstream area: methodology of the behavioral sciences in Spain // The Spanish Journal of Psychology. – 2009. – Vol. 12, № 2. – P. 833–849.
6. Сойфер В.Н. Международная Соросовская программа образования. Часть 2. Результаты именных конкурсов // Соросовский образовательный журнал. – 1996. – № 1. – С. 4–16.
7. Franceschini F., Maisano D., Perotti A., Proto A. Analysis of the ch-index: an indicator to evaluate the diffusion of scientific research output by citers // Scientometrics. – 2010. – Vol. 85. – P. 203–217.
8. Кара-Мурза С.Г. Цитирование в науке и подходы к оценке научного вклада // Вестник АН СССР. – 1981. – № 5. – С. 68–75.
9. Chen C., Hicks D. Tracing knowledge diffusion // Scientometrics. – 2004. – Vol. 59, № 2. – P. 199–211.
10. Folin O., Ciocalteu V. On tyrosine and tryptophane determinations in proteins // The Journal of Biological Chemistry. – 1927. – Vol. 73, № 2. – P. 627–650.
11. Lowry O.H., Rosbrough N.J., Farr A.L., Randall R.J. Protein measurement with the Folin phenol reagent // The Journal of Biological Chemistry. – 1951. – Vol. 193, № 1. – P. 265–275.
12. Gao X., Guan J. Network model of knowledge diffusion // Scientometrics. – 2012. – Vol. 90, № 3. – P. 749–762.
13. Schubert A. Using the h-index for assessing single publications // Scientometrics. – 2009. – Vol. 78, № 3. – P. 559–565.
14. Atallah G., Rodriguez G. Indirect patent citations // Scientometrics. – 2006. – Vol. 67, № 3. – P. 437–465.

Материал поступил в редакцию 02.04.13.

Сведения об авторах

ШТОВБА Сергей Дмитриевич - доктор технических наук, профессор кафедры компьютерных систем управления Винницкого национального технического университета, Украина
E-mail: shtovba@gmail.com

ШТОВБА Елена Валерьевна - кандидат экономических наук, доцент кафедры менеджмента и моделирования в экономике Винницкого национального технического университета, Украина
E-mail: olena.shtovba@yahoo.com

УДК 004.738.5:02

М. А. Плющ

Об использовании Интернета для сбора информации по истории книги и библиотек: на примере сведений о реестре А.И. Богданова

Излагаются и интерпретируются опубликованные в электронных и печатных источниках сведения о составленном А. И. Богдановым реестре рукописей из библиотеки Д. М. Голицына, конфискованной московской конторой Канцелярии конфискации. Отмечается причина составления реестра – задание, полученное А. И. Богдановым от И. Д. Шумахера. Рассматриваются качества реестра, отмечается его связь с московской описью рукописей и книг светского содержания, находившихся в библиотеке Д. М. Голицына. Оценивается значение Интернета в сборе информации, включая библиографические описания печатных работ, электронные статьи, оцифрованные дореволюционные издания, указания на архивные документы.

Ключевые слова: библиография, реестр рукописей Д. М. Голицына, персоналии, А. И. Богданов

По версии Летописи Библиотеки Академии наук А. И. Богданов в 1737 г. составил опись рукописей и книг уникальной библиотеки князя Дмитрия Михайловича Голицына [1, с. 69; 2, с. 134-156]*. В электронной энциклопедии «Книга» содержится утверждение, что А. Богданов «описывал и регистрировал книги, поступавшие [в Библиотеку Академии наук] из собраний Феофана Прокоповича, А. И. Остермана, Д. М. Голицына и др.» [3].

По данным из электронного «Словаря русского языка» А. Богданов якобы «описывал собрание ру-

кописей Феофана Прокоповича, которое поступило в библиотеку [Академии наук] после его смерти», в 1739 г. «составил «реестр» книгам и рукописям, конфискованным у Д. М. Голицына», в 1741 г. «разбирал рукописи, поступившие в библиотеку из дома А. И. Остермана» [4]. Аналогичные утверждения можно найти в словарной статье на сайте Института русской литературы в «Словаре русских писателей XVIII века» [5]. Они повторяются на сайтах Интернета с приведением библиографических ссылок. Однако Летопись БАН верно сообщает, что только некоторые рукописи из огромной библиотеки Феофана Прокоповича поступили в Библиотеку Академии наук [1, с. 67]. Тем самым опровергается мнение, что Богданов составлял реестр рукописей этой библиотеки. Источник указанной ошибки установить не удалось.

В опубликованной в 1941 г. статье Н. Н. Аблов утверждает без уточнений, что по поручению Академии наук Богданов разобрал «книги, оставшиеся от ... князя Дмитрия Голицына» и составил «библиографический «Реестр книгам, взятым из дому бывшего князя Дмитрия Голицына, на российском языке историческим и гражданским и прочим светским печатным и рукописным...» [6, с. 136-137]. Н. Н. Аблов не дает ссылок на печатные публикации и ссылается только на опубликованный архивный документ («Реестр книгам») в четвертом томе «Материалов по истории Императорской Академии наук» (см. далее).

* В статье под названием «О доступных источниках информации о библиотеках А. П. Вольнского: сведения о библиотеках и их поступлении в Библиотеку Академии наук» (Научно-техническая информация. Сер. 1. – 2012. - № 9. – С. 25-32) нами были допущены две ошибки. В ней говорится, что по версии Летописи Библиотеки Академии наук после конфискации в 1737 г. и доставки из Архангельского в Санкт-Петербург библиотеки князя Дмитрия Михайловича Голицына А. И. Богданов составил опись ее рукописей и книг. В действительности в Летописи не упоминается доставка библиотеки из Архангельского и само сведение является неверным. Сведение о доставке библиотеки Д. М. Голицына из Архангельского приводит к выводу, что библиотека находилась в имении князя. Этот миф возник в дореволюционной литературе и был развенчан, в том числе в пятой главе монографии [2]. Кроме того, в этой статье было отмечено, что, по данным публикации без автора на сайте «Музей-усадьба «Архангельское» (<http://www.arhangel'skoye.ru/hist-bibl>), опись библиотеки Д. М. Голицына на месте не проводилась, но, по подсчетам чиновников, число книг составило примерно 6 тысяч. В реальности указанное мнение о размере библиотеки имеет совершенно иное происхождение.

В книге И. Н. Кобленца, посвященной жизни и деятельности А. И. Богданова, отмечается, что после передачи библиотеки Д. М. Голицына в казну «в составленном Богдановым реестре книг и рукописей» было «перечислено 355 названий русских и переводных сочинений». По мнению И. Н. Кобленца, при конфискации в библиотеке Д. М. Голицына «оказалось около 6000 книг на немецком, голландском, испанском, английском, шведском и итальянском языках и множество переводов на русский язык» [7, с. 123]. Это мнение о количестве книг было опровергнуто С. П. Лупповым в монографии «Книга в России в первой четверти XVIII века» [8, глава 5]. И. Н. Кобленц отмечает, что в реестре Богданова было зарегистрировано 355 книг и только 9 из них были печатными [7, с. 124]. При рассмотрении качества богдановского реестра будет показано, что эти цифры неточны. Можно предположить, что указание о 355 книгах является опечаткой, так как реестр заканчивается строкой за номером 335, при этом отсутствуют 52 номера внутри нумерации. При просмотре указателя имен и текста в книге И. Н. Кобленца «Источники и деятели русской библиографии XV-XVIII вв.» (1991 г.) обнаруживается, что А. И. Богданов в ней не упоминается.

В монографии [8] отмечается, что из библиотеки Д. М. Голицына Богданов «отобрал 283 книги, перечень которых опубликован в четвертом томе «Материалов для истории Императорской Академии наук» [8, с. 208 и сноска № 81]. В действительности в богдановском реестре зафиксировано много материалов, не имеющих ничего общего с летописями, рукописными переводами научных книг и печатными изданиями. В некоторых строках зафиксировано несколько материалов под одним номером.

В 1973 г. в третьем томе издания «Сборник статей и материалов Библиотеки АН СССР по книговедению» вышла статья Г. Н. Моисеевой, посвященная вопросу об участии А. И. Богданова в создании «Краткого российского летописца» М. В. Ломоносова. Г. Н. Моисеева со ссылкой на книгу [7] пишет без уточнений, что Богданов в 1737 г. (в действительности в 1739 г., см. далее) «составлял опись рукописей и книг опального верховника Д. М. Голицына» [9, с. 286]. Отсутствие уточнений создает впечатление, что Богданов якобы произвел опись всей библиотеки князя Д. М. Голицына.

М. С. Слуховский утверждает, что «часть некогда громадной, в 6 тыс. экземпляров, библиотеки князя Д. М. Голицына стала «ценной коллекцией, поступившей в Академическую библиотеку из конфискованного имущества» [10, с. 197]. В этом утверждении неверно как названное количество экземпляров, так и мнение о передаче части библиотеки князя в Библиотеку Академии наук.

Ссылаясь на четвертый том «Материалов по истории Императорской Академии наук» [11] (далее «Материалы...»), М. И. Слуховский пишет, что Богданов приложил к рапорту о разборке библиотеки Д. М. Голицына «реестр 335 голицынских книг, почти сплошь рукописных» [10, с. 197-198]. Очевидно, исследователь не учел 52 пропуска в нумерации и другие особенности реестра. В итоге М. И. Слуховский

сделал ошибочный вывод, что номер последней строки в нумерации (№ 335) аналогичен количеству «книг» в реестре. Мнение, что Богданов произвел именно разборку библиотеки князя Дмитрия Михайловича также ошибочно. В реальности Богданов проводил отбор материалов для Библиотеки Академии наук и фиксировал в «реестре» отобранные материалы, опираясь на составленную в Москве опись светских книг и рукописей. Необходимо особо подчеркнуть, что согласно данным В. Г. Парушевой [2, с. 154-156] разборка библиотеки проводилась в Москве. Мнение, что библиотека хранилась в селе Архангельское, является мифом.

В статье Б. А. Градовой, Б. М. Клосса и В. И. Корецкого о рукописях библиотеки Д. М. Голицына справедливо отмечено, что «в литературе бытует неверное утверждение, будто бы опись книг Голицына составил А. И. Богданов». Авторы не уточняют, считают ли они «опись книг Голицына» описью всей его библиотеки или описью некоторой ее части. Именно в этой статье, вероятно впервые, почти верно указано, что А. И. Богданов в 1739 г. отбирал «наиболее интересные голицынские рукописи для академической библиотеки» [12, с. 181 и сноска № 9]. К сожалению, авторы не отметили наличие в богдановском реестре материалов, не имеющих ничего общего с летописями, научными работами и их рукописными переводами.

Как уже было сказано, Летопись Библиотеки Академии наук со ссылкой на статью Г. Н. Моисеевой (1973 г.) сообщает о составлении А. И. Богдановым описи «рукописей и книг опального верховника князя Д. М. Голицына» [1, с. 69]. При этом, как и в выше упомянутых публикациях, не уточняется, является ли богдановский «реестр» описью всей библиотеки Д. М. Голицына или описью ее части.

В четвертом томе издания «Материалов...», доступном в оцифрованной форме в Интернете, опубликован документ № 265 – рапорт Богданова, согласно которому он 26 июля 1739 г. получил от Академии наук письмо с требованием посетить Канцелярию конфискации. Богданову было поручено привезенные из Москвы «объявленные книги бывшего князя Дмитрия Голицына пересмотреть и отобрать, кои надобны будут, в академию наук» [11, с. 178]. Богданов отмечает в рапорте, что отобранные им книги записаны в приложенный к рапорту «реестр», а другие книги гражданской печати являются дублетами, уже имеются в академической Библиотеке. За этим документом следует «реестр взятым из дому бывшего князя Дмитрия Голицына на российском языке историческим и гражданским и протчим светским печатным и рукописным [книгам] по литерою АВ», датированный, как и рапорт, 18 августа 1739 г.

В нумерации строк реестра, составляющих на первый взгляд по номерам 335 строк, отсутствуют 52 номера. Во многих случаях на одну строку приходится несколько текстов, в нескольких случаях на один номер приходится часть той или иной книги. Прямое указание, что текст является печатным, имеется в 9 строках, но на строку № 321 с таким указанием приходится три текста. Указание Богданова в конце реестра, что печатными являются три книги

неверно. В действительности три печатные книги зафиксированы в последнем разделе реестра. Этому разделу предшествует регистрация книг, в том числе нескольких печатных, более крупных форматов.

Качество записей в реестре крайне низко из-за того, что он составлялся по описи, сделанной малограмотными чиновниками московской конторы Канцелярии конфискации (см. далее). В трех строках Богданов зарегистрировал «собрания летописцев», «хронические собрания». В некоторых местах он ввел такие записи, как «летописец собран из разных историй» (строка № 38), «летописец эллинский и римский» (строка № 39), «описание города Амстердама; тут же часть летописца Димитрия, митрополита; тут же хронология китайского государства; тут же летописец, собранный из книг бытеевских» (строка № 61), «книга Аристотеля о разных вещах физических; тут же собрания разныя духовныя» (строка № 89). Не менее 27 единиц в реестре занимают «книга записная...», «выписка из посольских книг», «вирши поздравительные» и подобного рода тексты. Большое количество строк занимают отдельные летописцы и рукописные тексты. Помимо явно научных книг в виде их рукописных переводов Богданов зарегистрировал много текстов, не имеющих научного значения, а также материалы, не являющиеся книгами. Записи в реестре имеют вид: «Иоанна Фридерика о state владетельском» (строка № 7), «наука артиллерийская и инженерная» (строка № 27), «роспись всяким разным книгам, а чьей библиотеки – не написано» (строка № 163), «описание о граде Риме, еще повесть о турках, еще перевод с письма права конфидератского и проч.» (строка № 251), «поздравление иезуита ко князю Вешновецкому» (строка № 306). По свидетельству С. П. Луппова [9, с. 208] и судя по переписке Академии наук, Канцелярии конфискации и Кабинета Ее Императорского Величества [11], зафиксированные Богдановым в «реестре» рукописные переводы, летописцы и другие материалы так и не поступили в фонд Библиотеки.

Авторами статьи [12] было сделано важное наблюдение. По их словам, на большинстве рукописей из голицынской библиотеки, оказавшихся в итоге в Государственной публичной библиотеке им. М. Е. Салтыкова-Щедрина (ныне РНБ), «видны следы четырех нумераций», а в некоторых случаях имеется пятая нумерация. Пятая нумерация соответствует номерам в «описи конфискации 1737-1738 гг. и представлена в 1739 г. ... А. И. Богдановым» [12, с. 180-181]. В этом сведении имеется в виду опись библиотеки Д. М. Голицына, составленная в Москве. Факт совпадения номеров пятой нумерации на большинстве сохранившихся в РНБ рукописей с номерами московской описи подтверждает, что Богданов использовал московскую опись для отбора в петербургской Канцелярии конфискации экземпляров из собрания Д. М. Голицына. Впервые этот вывод был сделан С. П. Лупповым в изданной в 1973 г. монографии «Книга в России в первой четверти XVIII века». Ученый пришел к нему на основе сличения богдановского реестра с московской описью русскоязычных книг светского содержания [8, с. 208], точнее преиму-

щественно рукописных книг, переводов и летописцев. Эта опись хранится в Российском государственном архиве древних актов в фонде Канцелярии конфискации (фонд 340, опись 1, часть 5, дело № 13891, листы 65-73)**.

В 1979 г. в статье [13] вышеназванная опись была опубликована в приложении № 1 [13, с. 241-249]. В нее входят 336 строк со сплошной ненарушенной нумерацией. При их сличении выявляется практически полная, в том числе лексическая и грамматическая тождественность указанной описи и богдановского реестра. Таким образом, не совсем верно, что Богданов отбирал из остатков библиотеки Д. М. Голицына книги, интересные для Академии наук. Точным содержанием работы Богданова являлся отбор преимущественно летописцев и рукописных переводов научных книг по московской описи материалов светского содержания [13, с. 249].

Исходя из опыта сбора информации для написания настоящей и предыдущих статей, нельзя не отметить следующее. Интернет при внимательном использовании способен сыграть выдающуюся роль в накоплении разнообразных сведений, включая библиографические описания печатных статей и электронные научные статьи, которых нет в печатной литературе. Очень большое значение имеет обнаружение в Интернете оцифрованных дореволюционных изданий и сносок на печатные старинные издания. В свою очередь, в них могут содержаться как сведения по теме статьи, так и отсутствующие в иных публикациях сноски на другие источники информации. Ссылки в электронных публикациях иногда включают указания на архивные документы, которые в обратном случае были бы упущены. Использование Интернета даже в пределах свободного доступа к веб-сайтам и платного доступа к нескольким электронным библиотекам является важным направлением сбора исторической и библиографической информации по истории книги и библиотек в России.

СПИСОК ЛИТЕРАТУРЫ

1. Летопись Библиотеки Академии наук. – Том 1. – 1714-1900. – СПб.: БАН, 2004. – 413 с. – URL: http://www.rasl.ru/e_editions/Chronicle.pdf
2. Крючкова М. А., Парушева В. Г. Русский Версаль: усадьбы князей Голицыных Архангельское и Николо-Урюпино. – М.: Русский Мир, 2012. – 336 с.
3. Энциклопедия «Книга». – URL: <http://polygraphicbook.narod.ru/text/statiy/2/bogda.htm>
4. Словарь русского языка XVIII века / отв. ред А. М. Панченко. – М.: Институт русской литературы и языка, 1988-1999. – Вып. 1. – URL: http://russo.com.ua/slovar_russkogo_yazyika_veka/page/bogdanov_andrey_ivanovich.86

** В монографии [8] опись с архивным делом № 13891 указана на с. 207 в сноске № 77 как опись № 1, на с. 208 в сноске № 82 как опись № 2. Проверка по другим печатным публикациям показывает, что верным является указание на опись № 1.

5. Институт русской литературы (Пушкинский дом) // Электронные публикации ИРЛИ. Справочные издания. Словарь русских писателей XVIII века. – URL: <http://www.lib.pushkinskiydom.ru>
6. Аблов Н. Н. Сподвижник Ломоносова, первый русский книговед – Андрей Богданов // Советская библиография. Сборник. – Вып. 1 (19). – М.: Изд-во Всесоюзной книжной палаты, 1941. – 239 с.
7. Кобленц И. Н. Андрей Иванович Богданов. 1692-1766. Из прошлого русской исторической науки и книговедения. – М.: изд-во АН СССР, 1958. – 214 с.
8. Луппов С. П. Книга в России в первой четверти XVIII века. – Л.: Наука, Ленинградское отделение, 1973. – 374 с.
9. Моисеева Г. Н. Об Андрее Богданове, первом русском книговеде (к вопросу об его участии в работе Ломоносова над «Кратким российским летописцем») // Сборник статей и материалов Библиотеки АН СССР по книговедению / отв. ред. А. А. Моисеева. – Л.: Изд. отдел Библиотеки АН СССР, 1973. – Т. 3. – 495 с.
10. Слуховский М. И. К биографии первого русского книговеда А. И. Богданова // Книга. Исследования и материалы. Сборник XXVI. – М.: изд-во «Книга», 1973. – 239 с.
11. Материалы для истории Императорской Академии наук. – Том четвертый (1739-1741). – СПб.: Типография Имп. Академии наук, 1887. – 824 с. // Сайт «Ушинский. РФ. Открытый архив опубликованных и неопубликованных документов». – URL: <http://www.ushinskiy.ru>
12. Градова Б. А., Клосс Б. М., Корецкий В. И. О рукописях библиотеки Д. М. Голицына в Архангельском // Археографический ежегодник за 1980 г. – М.: Наука, – 1981. – 367 с.
13. Градова Б. А., Клосс Б. М., Корецкий В. И. К истории Архангельской библиотеки Д. М. Голицына // Археографический ежегодник за 1978 год. – М.: Наука, 1979. – 367 с.

Материал поступил в редакцию 23.04.13.

Сведения об авторе

ПЛЮЩ Максим Андреевич – кандидат технических наук, старший научный сотрудник ВИНТИ РАН, Москва
E-mail: m92ash68@rambler.ru

РОССИЙСКАЯ АКАДЕМИЯ НАУК

Федеральное государственное бюджетное учреждение науки ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ РОССИЙСКОЙ АКАДЕМИИ НАУК

предлагает научным работникам, аспирантам и другим специалистам в области естественных, точных и технических наук, желающим быстро и эффективно опубликовать результаты своей научной и научно-производственной деятельности, использовать способ публикации своих работ через *систему депонирования*.

«Депонирование (передача на хранение) – особый метод публикации научных работ (отдельных статей, обзоров, монографий, сборников научных трудов, материалов научных мероприятий – конференций, симпозиумов, съездов, семинаров) узкоспециального профиля, разрешенных в установленном порядке к открытому опубликованию, которые нецелесообразно издавать полиграфическим способом печати, а также работ широкого профиля, срочная информация о которых необходима для утверждения их приоритета. Депонирование предусматривает прием, учет, регистрацию, хранение научных работ и обязательное размещение информации о них в специальных информационных изданиях».

Подготовка и передача на депонирование научных работ происходит в соответствии с «Инструкцией о порядке депонирования научных работ по естественным, техническим, социальным и гуманитарным наукам» (М., 2003).

Результатом депонирования является публикация информации о депонированных научных работах в информационных изданиях ВИНТИ РАН – Реферативном журнале и аннотированном библиографическом указателе «Депонированные научные работы».

В соответствии с “Положением о порядке присуждения ученых степеней”, утвержденным Постановлением Правительства Российской Федерации от 30.01.2002 № 74 (в ред. Постановлений Правительства РФ от 20.04.2006 № 227, от 02.06.2008 № 424, от 20.06.2011 № 475), научные работы, депонированные в организациях государственной системы научно-технической информации, признаны публикациями, учитываемыми при защите кандидатских и докторских диссертаций.

Подать научную работу на депонирование можно обратившись в Отдел депонирования ВИНТИ РАН по адресу:

125190, Москва, ул. Усиевича, 20.

ВИНТИ РАН, Отдел депонирования научных работ.

Тел.: 8 (499) 155-43-28, Факс: 8 (499) 943-00-60.

e-mail: dep@viniti.ru

С инструкцией о порядке депонирования можно ознакомиться на сайте ВИНТИ РАН:
<http://www.viniti.ru>