

ПЕРЕДОВЫЕ ИССЛЕДОВАНИЯ В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ

УДК 004.8

ЦИФРОВОЙ КОВЧЕГ ЗНАНИЙ

© 2022 г. В. В. Горячко¹, А. С. Бубнов¹, Е. В. Раевский¹, Академик РАН А. Л. Семенов^{1,2,*}

Поступило 30.10.2022 г.

После доработки 06.11.2022 г.

Принято к публикации 08.11.2022 г.

В работе рассматривается общая проблема создания и использования энциклопедического знания в цифровой цивилизации. Создание и использование цифровой энциклопедии являются примером интеллектуальной деятельности коллективной расширенной личности. Представлены основные положения, относящиеся к созданию и функционированию “Цифрового ковчега знаний” МГУ на примере математических наук. Рассмотрены возникающие проблемы и пути их решения.

Ключевые слова: энциклопедия, БРЭ, расширенная личность, расширенное сознание, википедия, коллективное знание, открытые общедоступные информационные ресурсы, профессиональное сообщество, МГУ им. М.В. Ломоносова, математика

DOI: 10.31857/S2686954322070098

1. ВВЕДЕНИЕ

Как уже отмечалось [1], развитие человечества можно охарактеризовать информационными революциями. Эти революции определялись возникновением:

- Сознания,
- Речи,
- Письма,
- Автоматизации интеллектуальной деятельности — искусственного интеллекта.

Каждая из революций сопровождалась изменением того, как человек мыслит, общается, действует. Можно сказать, что каждая из них приводила к расширению сознания, личности человека [2, 3]. Одновременно каждая из этих революций приводила к сокращению, иногда — отмиранию тех или иных форм деятельности и связанных с ними способностей человека, как это отмечает Л.С. Выготский [4–6]. По отношению к возникновению письменности об этом писал Платон в диалоге “Федр”, ссылаясь на (бесписьменного) Сократа [7].

Концепция расширенной личности является продуктивной метафорой для:

- описания современного человека как участника производственных, социальных и экономических отношений,

- проектирования целей и содержания образования и оценивания их достижения.

Последняя из перечисленных революций началась около 100 лет назад внутри математики — когда математики, начав с построения формальных математических систем описания математических рассуждений, математической коммуникации построили универсальную формализацию интеллектуальной деятельности вообще [8]. Эта формализация, благодаря успехам естественных наук и технологии, материализовалась в первой половине XX века в создании компьютера. Сегодня человек передает компьютеру все большие фрагменты интеллектуальной деятельности, начиная с ее рациональных элементов, таких, как решение вычислительных задач, включая численное моделирование и компьютерную алгебру. Взрывной рост цифровых технологий привел в XXI веке к автоматизации интуитивной деятельности человека — машинному обучению, с которым часто отождествляется и весь искусственный интеллект. Мы не будем вдаваться здесь в терминологическую дискуссию.

Внутри самой революции искусственного интеллекта также произошли революционные события. Важнейшее из них — это накопление информации в общем пространстве, ее повсеместная доступность и организация простого доступа к ней. Эта “суб-революция” знания (подобная

¹ Московский государственный университет имени М.В. Ломоносова, Москва, Россия

² Московский физико-технический институт (национальный исследовательский университет), г. Долгопрудный, Московская область, Россия

*E-mail: alsemno@ya.ru

суб-революции Гутенберга внутри революции письменности) основывалась на:

- интернете,
- мобильных устройствах (смартфонах),
- поисковых системах.

Возможно, следующая революция, которая нам предстоит, будет базироваться на прямом интерфейсе “мозг — компьютер”. Сейчас этот интерфейс частично реализован в виде помощника для восприятия устной речи при некоторых видах глухоты (кохлеарный имплант), протезировании зрения и примитивной (пока) передачи нервных импульсов в электронные устройства, в частности, для людей с ограниченной подвижностью.

К важнейшим, действительно революционным событиям XXI века относится и запуск принципиально новых форм коллективной интеллектуальной деятельности — формирование коллективной расширенной личности. Наиболее ясной и убедительной формой такой деятельности стала Википедия.

2. СОЗДАНИЕ ВИКИПЕДИИ КОЛЛЕКТИВНОЙ РАСШИРЕННОЙ ЛИЧНОСТЬЮ

Собрания знаний о мире играли принципиальную роль в развитии науки, культуры и образования. Достаточно упомянуть “Orbis pictus” Коменского [9] — начальный текстово-картиночный полилингвальный инструмент ориентации ребенка в мире, и “Энциклопедию” Дидро и Д’Аламбера — озаменовавшую Век Просвещения [10].

Википедия — симбиоз Человечества и Технологии. Это — дело расширенной глобальной человеческой личности, которое стало возможной благодаря цифровым технологиям. Благодаря ей человеческое знание стало доступно большинству населения Земли: теперь этим знанием фактически пользуется намного больше людей, чем всеми справочными изданиями бумажного века. При этом количество изданий Википедии в единицу времени выросло на порядки по сравнению с бумажными энциклопедиями.

Основой для социальной технологии Википедии стали ее правила — Столпы, разумные ограничения: энциклопедичность, нейтральность, общедоступность, взаимное уважение участников, отсутствие жестких правил.

Таким образом, Википедия является важнейшим массовым ответом ИИ на невозможность для человека и человечества справиться с экспоненциальным информационным взрывом.

Серьезные проблемы и недостатки Википедии, критически существенные для отдельных вопросов и сообществ (некоторые из них обсуж-

даются ниже), незначительно влияют на ее важность в целом.

3. ИСПОЛЬЗОВАНИЕ ВИКИПЕДИИ РАСШИРЕННОЙ ЛИЧНОСТЬЮ

Важнейшим эффектом проекта Википедии стало активное использование Википедии населением Земли. Без этого проект, конечно, потерял бы смысл. Сегодня, когда ученый, литератор, деятель культуры читает лекцию в широкой аудитории и ссылается на какую-то неожиданную идею, не слишком известную историческую личность, использует специфическое понятие, концепцию, в аудитории бывает заметно движение: несколько человек достают свои мобильные телефоны. Это значит, что аудитория состоит из расширенных личностей, частью индивидуального знания которых являются энциклопедические статьи Википедии. Такое было невозможно в доцифровую эпоху: так ли часто любознательный молодой человек середины XX века обращался к “Большой советской энциклопедии”? Сегодня же даже средний школьник, получив задание, выходящее за рамки учебника, немедленно обращается к Википедии. Не менее важным является то, что учащиеся уже в начальной школе могут создавать и использовать собственные микро-википедии: коллективное знание о ближайшем окружении (природном, техногенном, социальном) учеников одного класса (а потом — шире, школы); этот подход мы реализуем в проекте курса “Будущий мир” для начальной школы [11].

Таким образом, упомянутая “суб-революция” знания сегодня охватила существенную долю населения Земли — куда большую, чем охваченную письменной грамотностью в XIX веке (вы помните, что тогда в России она не превосходила 15%). По существу, как отмечают многие авторы, изменилось представление о том, что значит, что человек что-то знает. Это представление мигрирует в направлении расширенной личности. Пример отражения такой миграции в научном дискурсе сегодня можно видеть уже в бакалаврских работах [12].

Как это уже бывало неоднократно в последнее столетие, технологическое достижение, возможность которого доказана на Западе, было освоено и интерпретировано Китаем. Созданная там на авторитарной основе альтернатива Википедии — Байдупедия (см. [13]) — быстро обогнала всю мировую Википедию по объемам.

Еще одним явлением, существенным для мирового научного сообщества, являются общедоступные (Open Access) публикации и культура электронных препринтов. Показателен, хотя и уникален, пример Г. Перельмана, получившего высшие математические награды за такие архив-

ные, в частности, не рецензируемые (и не проверяемые на плагиат)) публикации.

Наконец, есть огромное море размещенной в интернете профессиональной литературы, часто сомнительной с точки зрения качества оцифровки и правообладания.

4. НЕДОСТАТКИ И ПРОБЛЕМЫ ВИКИПЕДИИ

Вот наугад взятый из интернета пример критики Википедии:

“Однако, главная беда Википедии заключается в том, что модераторы и администраторы этого ресурса, являясь безусловными эрудитами и специалистами по вики-разметке, зачастую не являются профильными специалистами в тех областях, о которых пишут. То есть, если сильно упростить, то они лишь компилируют разрозненную по источникам информацию в единый текст. Правила Википедии не запрещают слесарю из штата Техас писать статьи о достижениях онкологии в России, например” [14]. Автор в своей публикации указывает не только на недостаток “демократичности” Википедии, но и на возможность эксплуатации этого недостатка для неблагоприятных, в том числе, личных, или политических целей, противоречащих исходной идее и идеологии Википедии.

Замечательно при этом, что во многом именно исходная демократическая идея обеспечила “доказательство возможности — proof of concept” важнейшего технолого-социального проекта.

Ключевые проблемы Википедии сегодня:

- различие в уровне мотивации и квалификации отдельных авторов (те, кто может написать профессионально, не пишут),
- качество статей,
- полнота ссылок и материала, куда ведут ссылки.

Несмотря на эти проблемы, ясно, что сегодня Википедия стала явлением, значительно дополняющим энциклопедические ресурсы, созданные по традиционным издательским и социальным технологиям и переведенные “в цифру”.

5. КОВЧЕГ ЗНАНИЙ. ОСОБЕННОСТИ ПРОЕКТА МГУ

В доцифровую эпоху нашей стране удалось достичь значительных энциклопедических достижений, таких, как энциклопедия Брокгауза и Эфрона, Большая советская энциклопедия (БСЭ) в трех изданиях. Математическая энциклопедия [15], создание которой совпало с завершением советского периода развития отечественной математики, стала уникальным мировым явлением. Доказательством этого является создание на ее основе The En-

cyclopedia of Mathematics Европейского математического общества [16].

Большая российская энциклопедия (БРЭ) продолжила традицию БСЭ в бумажном формате, а последнее десятилетие и — в цифровом [17]. Сотрудничество с БРЭ обеспечивает для профессионального сообщества возможность выхода на все общество и поддержку государства. Далее мы описываем идущие сейчас процессы взаимодействия профессионального научного сообщества с БРЭ на примере, прежде всего, математики и информационных технологий.

Заметим, что параллельно с БРЭ развивается российская википедия, сегодня входящая в первую пятерку википедий мира. Также расширяется система открытого доступа к публикациям.

Начав свое сотрудничество с БРЭ, мы столкнулись с (общемировой) проблемой низкой мотивации профессионалов к написанию энциклопедических статей. Прямое использование механизмов википедии приводило бы к снижению качества результата.

В целях решения проблемы были выработаны следующие подходы:

- возможность отделения функции создания статей от других функций (заказ, рецензирование, передача в БРЭ);
- возложение задач обеспечения качества статей на профессиональное сообщество, так, как это делается в журналах, советах по защите диссертаций и т.п.;
- создание **Ковчега**, где готовятся и предварительно размещаются рабочие материалы, которые после их одобрения передаются в БРЭ;
- использование как фундамента уже существующих источников, прежде всего “Математической энциклопедии”;
- технологическая поддержка комфортной платформы коллективной работы с документами — вики-редактирования, размещенной на серверах в РФ;
- учет публикуемых в БРЭ статей как научных публикаций в ведущих журналах.

Структура процессов, из которых формируется общий процесс создания энциклопедии и участники этих процессов, складывается следующим образом:

- **профессионал**, которому профессиональное сообщество уже вручило ответственность за какую-то область знания, и он эту ответственность принял: заведующий кафедрой (передача знания новым поколениям профессионалов), заведующий отделом исследовательского института (новые исследования и результаты), председатель ученого совета (оценка вклада других), главный редактор журнала (представление новых резуль-

татов сообществу), член академии, профессор РАН (избран профессиональным сообществом):

- Получает из научно-авторитетного в своей области источника (например, отделения РАН) приглашение принять участие в формировании содержания Энциклопедии в его области; такое участие не предполагает личного написания статей: речь идет о формировании перечня статей, выборе авторов, реакции на написанные статьи.

- Непринятие такого приглашения, или его полное игнорирование ставит этого профессионала в данном конкретном отношении вне своей референтной группы. Аналогия: не такая уж большая доля специалистов отказывается от вхождения в редколлегию научного журнала.

- Принятие предложения означает, что профессионал вместе со своими коллегами разделяет ответственность за полноту и качество представления их области в Энциклопедии.

- Принявшие приглашение профессионалы каждой области образуют **редакционную группу (РГ)** этой области (например, теории чисел, или генетики), куда они могут кооптировать коллег, как правило, имеющих академический статус (научную степень).

- Каждый из членов РГ, если не получает возражений от членов РГ, может написать статью сам или предложить любому **автору** написать статью на тему из области, относящейся к РГ, или разместить на странице группы открытое предложение написать статью на предлагаемую им тему.

- Каждый желающий может написать статью на тему, предлагаемую какой-то РГ, или предложить свою тему и направить статью для размещения в Ковчеге, при этом он может разместить ее, например, и в русской Википедии.

- Статья, направленная для размещения, попадает в Ковчег после того, как она получила одобрение хотя бы одного члена РГ.

- Создаваемые в Ковчеге статьи могут использоваться в качестве основы (со ссылкой) статьи из Математической энциклопедии, БСЭ, Математического энциклопедического словаря, открытых свободно распространяемых источников. РГ может предложить БРЭ разместить какую-то существующую или переработанную статью.

- **Научно-редакционная коллегия БРЭ** по данной области знания, например, математике, направляет в РГ для размещения в Ковчеге и переработки существующие статьи из энциклопедических источников. При отсутствии предложений от всех РГ БРЭ размещает статью у себя, в Ковчеге появляется соответствующая ссылка.

- Любой член редакционной группы может предложить имеющуюся в Ковчеге статью для передачи в БРЭ. Статья, не получившая возражений от членов РГ, передается в БРЭ, с автором заклю-

чается договор, авторство может быть указано в БРЭ, автор получает от БРЭ разрешение опубликовать статью со ссылкой на первоначальный источник частично или полностью, в оригинале или переводе на иностранные языки, после опубликования статьи в БРЭ, со ссылкой на БРЭ.

Если члены РГ предлагают написать статью авторитетному ученому, специалисту в теме статьи, то в качестве возможной формы организации работы по написанию статьи таким ученым предлагается следующая технологическая схема:

- РГ может предложить рекомендации по написанию, включая, например, объем того варианта статьи, который будет размещен в БРЭ, а также список вопросов для освещения в статье, информацию о ближайших в возможной сетевой структуре статьях: более общие статьи, откуда идет ссылка на данную, более специальные, частные статьи, соседние статьи, с которыми возможно пересечение. Автор в ходе работы над статьей может существенно отклониться от этих рекомендаций, предложив альтернативы.

- Автор привлекает для написания статьи одного или нескольких своих сотрудников и учеников и делает устный доклад-интервью, используя предложенный ему РГ список вопросов, слушатели доклада также задают нужные им вопросы. Доклад записывается (при необходимости с фиксацией письменных заметок, формул и т.п.)

- Запись доклада расшифровывается, при необходимости, с использованием ресурсов Ковчеге.

- Расшифровка и запись обрабатываются слушателями — участниками доклада, уточняются формулировки, вставляются формулы, ссылки и т.п.

- Результат обработки представляется автору, который дорабатывает текст; он может включить кого-то из слушателей доклада в соавторы и т.д.

Таким образом:

1. Традиционная структура, на которой базируется Ковчег — **профессиональное сообщество** с его связями, иерархией, мотивами и т.п., непрерывно выстраивающимися столетиями. У его членов может не хватать ресурсов и мотивов для создания материалов Ковчеге своими силами.

2. Эта структура усиливается **вики-сообществом**, соседствующим с профессиональным, расширяющим и дополняющим его. Это сообщество динамически формируется всеми людьми, обладающими мотивацией к вхождению в проект. У этого сообщества и отдельных его членов может не хватать компетентности и влияния для создания авторитетного и качественного источника знания.

Наше научно-образовательное сообщество может дать пример сбалансированного сочетания традиционных и цифровых механизмов.

Качество материалов на Ковчеге в БРЭ по той или иной области знания становится отражением качества научного знания профессионалов и их заинтересованности в поддержании этого качества, количество и объем материалов будут отражать также интерес широкого сообщества к данной проблематике.

6. МЕТОДОЛОГИЯ И ТЕХНОЛОГИЯ КОВЧЕГА

Ковчег представляет собой структуру, предназначенную для непрерывного сетевого создания сетевого продукта для сетевого потребителя с использованием профессиональных иерархий для обеспечения качества.

Она может рассматриваться как реализация метафоры машинного обучения в техно-социальной среде коллективной расширенной личности. В соответствующей “нейро-сети” присутствует слой потребителей разрабатываемого содержания, слой разработчиков — авторов-редакторов, слой экспертов — РГ, слой заказчика — БРЭ. Мы планируем процесс обучения, совершенствования правил взаимодействия, — как часть функционирования сети.

Ядро платформы построено на базе вики-системы, разработанной на факультете Вычислительной математики и кибернетики МГУ. Выбор этого вида программного обеспечения обусловлен целым рядом преимуществ:

- Обеспечение возможности одновременной работы большого числа удаленных пользователей.
- Сохранение подробной истории всех редакций формируемого пользователями информационного потока.
- Возможность для пользователей работать с системой “тонкий клиент”, поскольку программное обеспечение и информация хранятся в центральном ядре.
- Возможность использования системы унифицированного простого языка вики-разметки для формирования текстов, включая формулы.
- Возможность простого создания гипертекстовых документов и шаблонов, организации ссылок.

7. ДАЛЬНЕЙШИЕ НАПРАВЛЕНИЯ РАЗВИТИЯ КОВЧЕГА

Дальнейшие направления развития Ковчег могут охватывать:

- Размещение архивов препринтов, аналогично зарубежным архивам.
- Размещение материалов открытого доступа, в том числе, например, учебных мульти-медиа

курсов, видео-хостинг, в том числе — интервью, воспоминаний, относящихся к истории математики и истории науки.

- Размещение материалов с иными правилами разработки и доступа.
- Размещение учебных сред (LMS) или стыковка с ними.
- Система видео-конференций с автоматизированным аннотированием, расшифровкой.

РГ могут стать активными участниками издания электронных научных журналов, различных форм экспертизы.

8. ЗАКЛЮЧЕНИЕ

Научно-образовательный проект МГУ “Ковчег знаний” базируется на цифровой платформе, разработанной специалистами МГУ им. М.В. Ломоносова. Платформа обеспечивает возможность одновременной работы над базами данных практически неограниченного числа специалистов формирующих, развивающих и поддерживающих единую сетевую структуру.

Интерфейсом проекта с российским и русскоязычным сообществом является портал “Большой российской энциклопедии”.

БЛАГОДАРНОСТИ

Авторы выражают благодарность академику В.А. Садовничему за постановку задачи, С.Л. Кравцу за полезное обсуждение и поддержку.

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа А.С. Бубнова и Е.Н. Раевского была поддержана Школой математических методов анализа сложных систем МГУ, А.Л. Семенова — грантом РНФ 22-11-00177.

СПИСОК ЛИТЕРАТУРЫ

1. *Семенов А.Л., Зискин К.Е.* Расширенная личность как основной субъект и предмет философского анализа. Следствия для образования // Человек и системы искусственного интеллекта, ред. Лекторский В.А. СПб.: ООО “Издательство “Юридический центр””, 2022. С. 172–200. ISBN 978-5-94201-835-1.
2. *Clark A.* Being there: Putting brain, body, and world together again // MIT press, 1998. URL: <http://www2.econ.iastate.edu/tesfatsi/BeingThere.AClark1998.EntireBook.pdf> (дата обращения 08.11.2022).
3. *Serpp M.* Девочка с пальчик // М.: Ад Маргинем Пресс, 2016. Оригинал: Serres M. Petite Poucette. Éditions Le Pommier, Paris, 2012.
4. *Выготский Л.С.* Инструментальный метод в психологии // Собр. соч. В 6 т. Т. 1. 1982. URL:

- http://elib.gnpbu.ru/text/vygotsky_ss-v-6tt_t1_1982/go,108;fs,1 (дата обращения 08.11.2022).
5. *Vygotsky L.S.* The instrumental method in psychology // 1981, URL: <https://www.marxists.org/archive/vygotsky/works/1930/instrumental.htm> (дата обращения 08.11.2022).
 6. *Vygotsky L.S.* Mind in Society: The Development of Higher Psychological Processes // Harvard University Press, 1980.
 7. Платон. Федр // <https://www.plato.spbu.ru/TEXTS/PLATO/Academia/005-02.pdf>. См. также обсуждение: URL: <https://cyberleninka.ru/article/n/mif-ob-izobretenii-pismennosti-v-dialoge-platona-fedr-sovremennoe-osmyslenie-idei-sokrata-i-platona-o-preimushchestve-ustnogo-slova-2> (дата обращения 08.11.2022).
 8. *Turing A.M.* On Computable Numbers, with an Application to the Entscheidungsproblem // Proc. Lond. Math. Soc., series 2, vol. 42, 1937. Pp. 230–265. Correction: ibid. vol. 43, 1937. Pp. 544–546.
 9. *Komensky J.A.* Orbis Pictus. Wikipedia page // https://en.wikipedia.org/wiki/Orbis_Pictus
 10. Энциклопедия Дидро и Д’Аламбера. Виртуальная выставка // Электронные выставки Иностранки, URL: <https://press-libfl.tilda.ws/enciklopediya-didro-i-dalambere-triumf-prosveshcheniya> (дата обращения 08.11.2022).
 11. *Семенов А.Л., Булин-Соколова Е.И., Муранов А.А. и др.* Цифровые технологии в начальной школе. Вход в будущий мир // Информатизация образования и методика электронного обучения: цифровые технологии в образовании. Материалы VI Международной науч. конф., г. Красноярск, 20–23 сентября 2022 г. В 3 ч. Ч. 2 / под общ. ред. М.В. Носкова. – Красноярск: КГПУ им. В.П. Астафьева, 2022. С. 325–329. ISBN 978-5-907558-24-3.
 12. *Choi G.* The Internet-Extended Mind: The Psychological Ramifications and Philosophical Implications of Cognitive Offloading // Scripps Senior Theses, 1675. The Claremont Colleges Library, USA, 2021. URL: https://scholarship.claremont.edu/scripps_theses/1675 (дата обращения 08.11.2022).
 13. Энциклопедия Байду. Страница Википедии // URL: https://ru.wikipedia.org/wiki/Энциклопедия_Байду (дата обращения 08.11.2022).
 14. *Кузнецов С.* А редактор кто? Что не так со статьями в Википедии // Портал RuPosters, 26 марта 2020 г. URL: <https://ruposters.ru/news/26-03-2020/oruzhie-informatsionnoi-voini> (дата обращения 08.11.2022).
 15. Математическая энциклопедия // В 5 томах. Гл. ред. И.М. Виноградов. М.: Сов. энциклопедия, 1977.
 16. Encyclopedia of Mathematics // The European Mathematical Society, Springer Verlag. URL: <http://encyclopediaofmath.303.si> (дата обращения 08.11.2022).
 17. Большая российская энциклопедия — электронная версия // URL: <https://bigenc.ru> (дата обращения 08.11.2022).