

УДК 002.63:004.77

А.Б. Антопольский, А.В. Босов, Г.И. Савин, А.Н. Сотников, В.А. Цветкова, Н.Е. Каленов, В.А. Серебряков, Д.В. Ефременко

Принципы построения и структура единого цифрового пространства научных знаний (ЕЦПНЗ)*

Рассматриваются принципиальные особенности и подходы к формированию Единого цифрового пространства научных знаний (ЕЦПНЗ). Отмечено, что работы по формированию единой среды информационных ресурсов, а далее – знаний ведутся почти четверть века. Приведены нормативные документы, направленные на разработку данной проблемы; отмечено, что существует неоднозначность в подходах к определению единого пространства. Актуализировано определение, сформулированы основные принципиальные особенности пространства знаний, в числе которых отмечены: необходимость включения достоверной фундаментальной и научно-популярной информации; наличие ряда подпространств, связанных между собой, относящихся к отдельным научным направлениям. Проанализированы ресурсы, составляющие основную базу для формирования ЕЦПНЗ. Показаны основные начальные направления работ. Отмечено, что данное научное направление должно стать ключевым в решении задач, поставленных перед научным сообществом в рамках национального проекта «Наука» и национальной стратегии «О развитии искусственного интеллекта в Российской Федерации» (2019-2030 годы).

Ключевые слова: единое пространство знаний, онтология, принципы построения, структура, информационные ресурсы, электронные библиотеки, метаданные, виртуальные выставки

DOI: 10.36535/0548-0019-2020-04-2

ВВЕДЕНИЕ

Проблема объединения электронных (цифровых) информационных ресурсов и создания единой среды, обеспечивающей их формирование, хранение и предоставление пользователям, привлекла внимание руководящих органов России четверть века назад.

В 1995 г., в начальный период бурного развития «эры Интернет», была разработана и одобрена Президентом РФ «Концепция формирования и развития единого информационного пространства России и соответствующих государственных информационных ресурсов»¹.

Согласно этой Концепции «Единое информационное пространство представляет собой совокупность баз и банков данных, технологий их ведения и использования, информационно-телекоммуникационных систем и сетей, функционирующих и взаимодействующих на основе единых принципов и по общим правилам...». Отдельным вопросам, связанным с понятием «единое информационное пространство», посвящен ряд публикаций в профессиональной печати 2013 и 2014 годов [1, 2].

В 2014 г. в документе «Об утверждении Основ государственной культурной политики», утвержденном

* Работа выполнена в рамках проектов РФФИ КОМФИ № 18-00-00294, 18-00-00372, 18-00-00297, 18-00-00298
The work was performed within the framework of RFBR COMFI projects № 18-00-00294, 18-00-00372, 18-00-00297, 18-00-00298

¹ Концепция формирования и развития единого информационного пространства России и

соответствующих государственных информационных ресурсов (одобрена решением Президента РФ от 23.11.1995 N Пр-1694). – URL: <https://rulaws.ru/acts/Kontseptsiya-formirovaniya-i-razvitiya-edinogo-informatsionnogo-prostranstva-Rossii-i-sootvetstvuyuschih-gosu/> (дата обращения 21.01.2020)

Указом Президента РФ², указывается на необходимость создания единого электронного пространства знаний как части информационного пространства страны.

В 2016 г. публикуется статья, в которой рассматриваются правовые вопросы создания Единого российского электронного пространства знаний [3], трактуемого как собрание оцифрованных объектов, в том числе, защищенных авторскими правами.

Документ «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы», утвержденный Указом Президента России в 2017 году³, включает несколько статей, касающихся информационного пространства, в том числе информационного пространства, «основанного на знаниях», и дает определение информационного пространства, несколько отличное от приведенного выше, сформулированного за 12 лет до этого. В стратегии информационное пространство трактуется как «совокупность информационных ресурсов, созданных субъектами информационной сферы, средств взаимодействия таких субъектов, их информационных систем и необходимой информационной инфраструктуры» (стр. 1, п. 4д); здесь же (стр. 24) формулируются цели формирования пространства знаний – «Целями формирования информационного пространства, основанного на знаниях (далее – информационное пространство знаний), являются обеспечение прав граждан на объективную, достоверную, безопасную информацию...». В ст. 25 указывается: «Формирование информационного пространства знаний осуществляется путем развития науки, реализации образовательных и просветительских проектов, создания для граждан общедоступной системы взаимоувязанных знаний и представлений, обеспечения безопасной информационной среды для детей, продвижения русского языка в мире, поддержки традиционных (отличных от доступных с использованием сети "Интернет") форм распространения знаний».

Заметим, что в этом документе отсутствует упоминание «электронного пространства знаний», а речь идет о существенно более широком понятии «информационное пространство знаний» в его философской трактовке.

В начале 2019 г. Правительством РФ было утверждено «Положение о Национальной электронной библиотеке»⁴, в котором пространство знаний трак-

туется как эквивалент единого российского электронного пространства знаний (ЕРЭПЗ): «Пространство знаний – единое российское электронное пространство знаний, представляющее собой совокупность взаимно интегрированных на основе Национальной электронной библиотеки информационных систем и иных информационных ресурсов, сформированных на базе научного, исторического и культурного достояния народов Российской Федерации, образцов зарубежных научных, культурных и исторических ценностей и функционирующих на основе единых информационных технологий и принципов, которые обеспечивают семантическую взаимосвязь их содержимого, а также инструменты поиска и извлечения информации по запросу пользователей».

В приведенных государственных документах один и тот же термин трактуется по-разному. Если в документах предыдущих лет информационное пространство знаний рассматривалось как некая философская категория, обращение к которой позволяло решать многочисленные проблемы, связанные с различными сторонами жизни общества, то в Положении о Национальной электронной библиотеке (НЭБ) прямым текстом сказано, что ЕРЭПЗ представляет собой объединение существующих информационных ресурсов на основе НЭБ. Фактически, это означает, что речь идет лишь о создании надстройки, обеспечивающей поиск данных в различных (прежде всего, документальных) информационных системах.

Таким образом, речь идет лишь о создании надстройки, обеспечивающей поиск данных в различных (прежде всего, документальных) информационных системах.

Необходимо отметить, что подобную задачу в области научно-технической информации пытались решать на государственном уровне в Советском Союзе, начиная с конца 1970-х гг. (естественно, на существовавшей в то время организационной, технологической и технической базе), в рамках программы по созданию «Государственной автоматизированной системы научно-технической информации (ГАСНТИ)» [4]. В связи с коренными изменениями политической и экономической структуры страны, с 1992 г. ГАСНТИ фактически перестала существовать.

В современных условиях для успешного развития образования и науки требуется не только обеспечить формирование и поиск электронных ресурсов, отражающих историческое и культурное достояние, но и «необходимо создать принципиально новую цифровую среду, отражающую историю и динамику развития науки и оперирующую не столько документальной информацией, сколько научными фактами, извлеченными из достоверных источников».

Исходя из необходимости создания подобной среды, в 2019 г. авторами была сформулирована задача формирования единого цифрового пространства научных знаний (ЕЦПНЗ). На исследования, связанные с формированием ЕЦПНЗ четырех организациям – МСЦ РАН (филиал ФГУ НИЦ НИИСИ РАН), ФИЦ ИУ, ИНИОН РАН и БЕН РАН –

² «Об утверждении Основ государственной культурной политики». Указ Президента РФ № 808 от 24.12.2014 г. – URL: <http://kremlin.ru/acts/bank/39208>.

(дата обращения 21.01.2020).

³ Указ Президента Российской Федерации от 09.05.2017 г. № 203. О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы. – URL: <http://www.kremlin.ru/acts/bank/41919>

(дата обращения 02.02.2020)

⁴ Постановление Правительства РФ от 20.02.2019 г. № 169 « Об утверждении Положения о федеральной государственной информационной системе "Национальная электронная библиотека" и методики отбора объектов Национальной электронной библиотеки». – URL: <http://static.government.ru/media/files/yYAeqiBDCwITsAXr8LtXsdsdefAbLzi9.pdf> (дата обращения 02.02.2020)

выделены гранты РФФИ по направлению КОМФИ (проекты № 18-00-00294, 18-00-00372, 18-00-00297, 18000-00298).

Настоящая статья отражает результаты, полученные в ходе выполнения указанных проектов в 2019 г.

ЕЦПНЗ: ПРИНЦИПЫ ПОСТРОЕНИЯ

Согласно определению, приведенному в [5], под Единым цифровым пространством научных знаний понимается «компьютерная среда, обращаясь к которой пользователь (начиная от учёного и кончая учеником средней школы) должен получить ответы на вопросы, касающиеся тех или иных областей науки»; принципиальные особенности ЕЦПНЗ заключаются в следующем.

1. Среда должна содержать достоверную фундаментальную и научно-популярную информацию и состоять из ряда подпространств, связанных между собой, относящихся к отдельным научным направлениям.

2. Каждое подпространство, в свою очередь, включает две составляющие, одна из которых (базис) отражает зафиксированные научные знания, другая (надстройка) – новую научную информацию. Часть надстройки по истечении некоторого времени, после прохождения экспертного фильтра, вливается в базис, а информация, не прошедшая экспертного фильтра, из ЕЦПНЗ исключается.

Согласно предлагаемой модели каждое подпространство ЕЦПНЗ включает ряд связанных множеств, первое из которых (хранилище объектов) содержит цифровые объекты с их идентификаторами, указывающими, в частности, на принадлежность к базису или к надстройке. Второе множество представляет собой терминологическое описание области знания – предметную онтологию данного подпространства. Она включает рубрики основных классификационных систем (ГРНТИ, УДК, МКИ и др.), относящихся к данному научному направлению, и ключевые термины с установленными между ними связями, организованные в виде тезаурусов. Остальные множества отражают свойства объектов, а также условия их предоставления пользователям.

Множества свойств объектов формируются на основании профилей метаданных, специфичных для каждой предметной области, но формируемых по единому принципу, предусматривающему представление метаданных в виде совокупности связанных между собой фасетов «что /кто», «где», «когда».

Связь между подпространствами осуществляется по обобщенной онтологии ЕЦПНЗ, которая, в свою очередь, строится на основе предметных онтологий, входящих в отдельные подпространства.

3. ЕЦПНЗ должно обладать свойствами семантической интероперабельности [6], обеспечивающими возможность обмена данными с отечественными и зарубежными информационными системами. Основой для обмена данными в ЕЦПНЗ служат технологии *Semantic Web*, соответственно, существенным является применение в ЕЦПНЗ существующих предложений по стандартизации наборов элементов метаданных для *Semantic Web* и принципов формирова-

ния данных, декларируемых в рамках *Simple Knowledge Organization System (SKOS)* [7–10].

Предложения по наборам элементов метаданных ЕЦПНЗ являются развитием идей *Dublin Core Metadata Initiative (DCMI)* [11] в направлении дальнейшей детализации наборов элементов метаданных различных предметных областей науки. Они предусматривают обеспечение поддержки обмена как метаданными общего характера в рамках всего научного сообщества, так и более узкоспециализированными метаданными в рамках заинтересованных «предметных» сообществ (биологического, астрономического, математического и пр.).

В отличие от предложений *Dublin Core Metadata Terms* [12], которые содержат достаточно небольшой набор общих элементов метаданных, нацеленных в основе своей на описание веб-документов, структура предложений ЕЦПНЗ охватывает существенно более широкий набор понятий (например, «теорема» для математики, «реакция» для химии, «корпус языка» для языкознания и т.п.). Соответственно, при проектировании ЕЦПНЗ встает задача структурной организации элементов метаданных, выделения отдельных профилей, специализаций, подсхем, уровней детализации описания.

С технической точки зрения каждое подпространство ЕЦПНЗ является распределенной структурой, включающей ядро и периферию. Ядро содержит предметную онтологию данного подпространства, метаданные объектов подпространства и центральное хранилище объектов. Все ядра подпространств связаны с центром ЕЦПНЗ на уровне обобщенной онтологии, поддерживаемой в центре.

Необходимо отметить, что сформулированные выше принципы, положенные в основу создания ЕЦПНЗ, основаны на исследованиях, выполненных при участии авторов настоящей статьи по программам РАН и грантам РФФИ в прошлые годы. Результаты этих исследований нашли свое отражение в ряде успешно реализованных проектов, таких как ИСИР РАН [13, 14], ЕНИП РАН [15–17] и успешно функционирующей в настоящее время ЭБ ННР [18–20].

СТРУКТУРА И РЕСУРСЫ ЕЦПНЗ

Основное хранилище каждого подпространства Единого цифрового пространства научных знаний включает объекты следующих видов:

- текстовые файлы (распознанные оцифрованные печатные или рукописные документы) или документы, изначально сформированные в электронном виде;
- статические изображения (нераспознанные оцифрованные документы, оцифрованные или изначально сформированные в цифровом виде фотографии);
- цифровые или оцифрованные аудиозаписи;
- цифровые или оцифрованные видео/киноматериалы;
- 3D-модели различного рода;
- мультимедийные инсталляции (цифровые модели природных процессов и технических устройств, учебные материалы, виртуальные экскурсии и т.п.).

Основные хранилища подпространств ЕЦПНЗ могут быть распределенными, их объекты могут физически входить в электронные библиотеки (в том числе в НЭБ), архивные и музейные коллекции, фактографические базы данных и другие собрания, гарантированно поддерживаемые организациями различного типа, соблюдающими требования ЕЦПНЗ к профилям метаданных. Метаданные объектов, хранящихся децентрализованно, размещаются, преимущественно, в ядрах соответствующих подпространств. В ряде случаев возможно иерархическое представление метаданных – формируется профиль метаданных полной коллекции в виде объекта ЕЦПНЗ, хранящегося в ядре подпространства, а уточняющие метаданные отдельных объектов коллекции, хранятся у ее владельца.

Источниками перечисленных ресурсов ЕЦПНЗ должны служить, в первую очередь, уже существующие цифровые коллекции публикаций, объектов музейного и архивного хранения (в том числе 3D-модели и мультимедийные инсталляции, оцифрованные научные и научно-популярные кинофильмы), специализированные базы данных, поддерживаемые научными учреждениями и вузами.

Важной составляющей основных хранилищ всех подпространств ЕЦПНЗ должны быть статьи из тематических научных энциклопедических словарей и справочников со ссылками на опубликованные источники информации, которые представлены в цифровом виде в базе каждого подпространства.

Основными поставщиками таких публикаций для всех подпространств ЕЦПНЗ являются политематические электронные библиотеки (ЭБ), такие как Научная электронная библиотека (<https://elibrary.ru/>), включающая огромный массив статей из отечественных журналов; вышеупомянутая НЭБ (<https://rusneb.ru/>), в которой должны быть отражены наиболее важные научные монографии; ЭБ «Научное наследие России» – (далее ЭБ ННР - <http://e-heritage.ru/>), отражающая информацию о российских ученых, начиная с XVIII в., их публикациях, музейных и архивных документах, связанных с ними.

Поставщиком информации для ряда подпространств в области общественных наук должна стать Президентская библиотека им. Б.Н.Ельцина (<https://www.prlib.ru/collections>) в части научной составляющей ее ЭБ по российской государственности. Научные знания по тематике региональных исследований содержатся в специализированных ЭБ, в частности, в ЭБ «Научное наследие Урала» (<http://i.uran.ru/nasledie/>).

Значительную часть тематических подпространств ЕЦПНЗ должны составлять элементы существующих отечественных научных информационных систем, таких как *Math-Net* (<http://www.mathnet.ru/>), Соционет, ЭБ ННР, построенная по онтологическим принципам Информационная система по геологии и горному делу (<http://higeo.ginras.ru/>) [21, 22].

Из существующих цифровых коллекций научных музеев, подлежащих загрузке в тематические подпространства ЕЦПНЗ, можно назвать ряд разработок, выполненных в соответствии с онтологией ЭБ Научное наследие России: коллекцию минералов Минералогического музея им. В.И. Вернадского, интегриро-

ванную в эту библиотеку; коллекцию «Портреты по скелетам» (<http://acadlib.ru/>) [23], посвященную научному наследию антрополога М.М. Герасимова, созданную на основе фондов Государственного биологического музея им. К.А. Тимирязева и Государственного Дарвиновского музея; мультимедийный ресурс «Сад жизни» (<http://vim.benran.ru/>) [24], посвященный научной деятельности И. В. Мичурина, созданный на основе фондов Государственного биологического музея им. К. А. Тимирязева и Российского государственного архива кинофотодокументов.

Важнейшей составляющей контента ЕЦПНЗ являются онтологии предметных областей, обеспечивающие многоаспектный поиск ресурсов в отдельных подпространствах ЕЦПНЗ, формируемая на их основе общая предметная онтология, обеспечивающая целостность ЕЦПНЗ.

Онтологии предметных областей ЕЦПНЗ должны включать основные классификационные системы, используемые поставщиками цифровых объектов. Для отечественных публикаций это, в первую очередь, политематические системы – международная УДК, отечественные ББК и ГРНТИ; для патентов – наряду с ГРНТИ – международная классификация изобретений (МКИ). Для формирования онтологий отдельных тематических направлений необходимо использовать, в дополнение к указанным политематическим, специальные классификационные системы, принятые для описания ресурсов внутри этих тематических направлений (например, классификатор *MeSH* для медико-биологических ресурсов [25], классификатор *MSC* для математических ресурсов [26] и др.).

Наряду с рубриками, которые в силу своей природы не обладают необходимой динамичностью, в состав онтологий должны входить наборы ключевых терминов, связанных между собой тезаурусными отношениями.

ОСНОВНЫЕ НАПРАВЛЕНИЯ ПРЕДВАРИТЕЛЬНЫХ РАБОТ ПО СОЗДАНИЮ ЕЦПНЗ

В числе основных направлений работ по созданию ЕЦПНЗ необходимо указать следующие.

1. В рамках общесистемных исследований, связанных с вопросами проектирования и функционирования ЕЦПНЗ, следует определить организационную структуру пространства в целом и отдельных его подпространств; решить принципиальные вопросы формирования контента отдельных подпространств; проработать правовые вопросы включения ресурсов в ЕЦПНЗ, их хранения и предоставления пользователям; разработать требования к хранению и динамическому перераспределению ресурсов между базисом и надстройкой для различных подпространств; определить общие требования и порядок включения данных в ЕЦПНЗ; разработать требования к сервисной части ЕЦПНЗ; зафиксировать внутренние стандарты по использованию форматов представления данных; определить правила обмена данными между подпространствами внутри ЕЦПНЗ и с внешней средой.

2. Для решения вопроса о первоначальном наполнении отдельных подпространств ЕЦПНЗ необходимо провести мониторинг реально существующих в России цифровых ресурсов; выработать методологию их оценки и отбора; решить процедурные вопросы включения отобранных ресурсов в отдельные подпространства ЕЦПНЗ. Большая работа по выявлению отечественных информационных ресурсов социогуманитарного профиля проведена специалистами ИНИОН РАН. В 2017-2018 гг. был проведен мониторинг информационных ресурсов по общественным и гуманитарным наукам [27] учреждений РАН. В результате был создан Навигатор информационных ресурсов по общественным наукам (НИРОН), размещенный по адресу <http://niron.inion.ru>. Он включает свыше 3 тыс. ресурсов различных типов – как традиционных, так и цифровых, имеющихся в 214 учреждениях. В это число не входит свыше 13 тыс. архивных фондов и описей, отраженных в фондах Института системного анализа РАН (ИСА РАН) (<http://www.isaran.ru/?q=welcome>). Описание НИРОН содержится в [28]. Статистика типов ресурсов, отраженных в НИРОН, представлена в табл. 1.

Следует учитывать, что количество архивов, архивных фондов и описей в табл. 1 указано по РАН в целом, поскольку выделение архивов социогуманитарного профиля в отдельный блок не представляется возможным.

Данные табл. 1 отражают ресурсы по следующим наукам (тематика выражена в рубриках 1-го уровня ГРНТИ):

Общественные науки в целом
Философия
История. Исторические науки
Социология
Демография
Экономика и экономические науки

Государство и право. Юридические науки
Политика и политические науки
Науковедение
Языкознание
Литература. Литературоведение. Устное народное творчество
Информатика
Религия. Атеизм
Комплексное изучение стран и регионов

По другим общественным и гуманитарным наукам в НИРОН ресурсы почти не отражены, поскольку в учреждениях РАН они не создаются.

Точный учет доли цифровых ресурсов также весьма затруднен, поскольку статистика оцифровки традиционных фондов недоступна.

Подобную работу необходимо провести по всем научным направлениям, подлежащим включению в ЕЦПНЗ. Однако очевидно, что одномоментно этого сделать не удастся. Поэтому в процессе общесистемных исследований необходимо определить очередность формирования отдельных подпространств ЕЦПНЗ и проводить мониторинг ресурсов в соответствии с этой очередностью.

Мониторинг информационных ресурсов имеет свою специфику по различным наукам. В 2019 г. при поддержке РФФИ (проект № 18-00-00298) был разработан специализированный Навигатор информационных ресурсов по языкознанию (НИРЯЗ), который представляет собой существенно доработанный, расширенный и актуализированный фрагмент НИРОН по цифровым лингвистическим ресурсам. Тестовая версия НИРЯЗ размещена по адресу <http://niryaz2.alexo.beget.tech/>. Состав лингвистических ресурсов представлен в табл. 2.

Таблица 1

Информационные ресурсы Российской академии наук социогуманитарного профиля

№ п/п	Типы ресурсов	Количество
1	Библиотеки – учреждения и подразделения универсального и социогуманитарного профилей	86
2	Архивы – учреждения и подразделения	74
3	Музеи – учреждения и подразделения	112
4	Архивные фонды (данные ИСА РАН)	5498
5	Каталоги, библиографии, библиографические БД	550
6	Архивные описи (данные ИСА РАН)	7703
7	Электронные библиотеки и полнотекстовые коллекции изданий	247
8	Информационные системы	90
9	Энциклопедии, справочники	117
10	Лингвистические ресурсы	184
11	Персональные ресурсы	185
12	Периодические издания, их архивы	402
13	Материалы мероприятий (имеющиеся на сайтах владельцев)	184
14	Коллекции неопубликованных документов	420
15	Медиа ресурсы	109
16	Сайты владельцев ресурсов = число владельцев ИР	214
17	Прочие интернет-ресурсы	209
	Всего	16384

Типология и статистика лингвистических информационных ресурсов РАН

№	Типы лингвистических ресурсов	Количество
1	Библиографии, библиотечные каталоги, описи, каталоги ссылок	93
2	Электронные коллекции и библиотеки полных текстов	138
3	Периодические, продолжающиеся издания и архивы периодики	77
4	Корпуса	15
5	Лексикографические БД	83
6	Этнолингвистические и социоллингвистические БД	10
7	Лингвистические географические системы, атласы	5
8	Электронное представление памятников письменности	5
9	Активные ресурсы (алгоритмы, процессоры, программы)	28
10	Грамматики	11
11	Описания языков и комплексные лингвистические сайты	14
12	Информационные языки	10
13	Энциклопедии, справочники, реестры языков	75
14	Сведения об отдельных персонах (сайты и страницы ученых, личные фонды, био-библиографии)	473
15	Списки, перечни, указатели персон	31
16	Медиаресурсы	8
17	Сайты учреждений – владельцев лингвистических ресурсов	116
18	Сайты-сателлиты и ресурсы во внешних АИС	46
	Всего	1238

3. Разработка методологии формирования онтологий отдельных подпространств и пространства в целом должна включать несколько направлений: исследования в области установления связей между различными классификационными системами, используемыми для индексирования научных ресурсов; разработка методов и алгоритмов формирования базовых терминологических понятий в заданной предметной области; разработка методов и алгоритмов формирования тезаурусов (предметных онтологий) заданных областей научного знания. Основой для таких работ могут служить исследования, проводимые по этим направлениям (в том числе с участием авторов настоящей статьи) в прошлые годы [29–31], а также исследования, выполняемые в настоящее время по грантам РФФИ (проекты № 18-00-00294, 18-00-00372, 18-00-00297, 18-00 – 00298), в рамках которых реализована модель системы терминологических словарей [32]; предложена и реализована на практике методология извлечения ключевых терминов из научных текстов в области микробиологии [33], разрабатывается база данных поисковых терминов разнородных лингвистических ресурсов.

4. Разработка системы стандартов на метаданные информации, циркулирующей в ЕЦПНЗ, эти стандарты должны отвечать следующим требованиям:

- систематизировать основные типы информации, требующейся для решения всего комплекса задач ЕЦПНЗ;
- быть открытыми, т.е. обеспечивать доступ к соответствующей информации по метаданным;

- быть расширяемыми, т.е. обеспечивать возможность детализации описаний метаданных и импорта / экспорта информации;

- обеспечивать возможности уникальной идентификации информации;

- быть ориентированными на современные и перспективные технологии описания и использования информации (в современном понимании – ориентироваться на *Semantic Web*);

- создавать условия для интероперабельности с внешней средой.

5. Разработка инструмента конструирования ЕЦПНЗ, позволяющего выполнять:

- формирование многоуровневой модели предметной области;

- предоставление пользовательских и машинночитаемых интерфейсов, обеспечивающих пакетную загрузку данных, составляющих контент подпространств ЕЦПНЗ;

- создание/просмотр/редактирование: информационных ресурсов; онтологий предметных областей; коллекций информационных объектов;

- навигацию по контенту ЕЦПНЗ;

- импорт / экспорт данных;

- реализацию атрибутивного/семантического/полнотекстового поиска;

- выделение дополнительных семантических связей между элементами ЕЦПНЗ.

Для перехода от исследований и моделирования отдельных подпространств ЕЦПНЗ, выполняемых перечисленными организациями в рамках грантов

РФФИ, к промышленному проектированию ЕЦПНЗ, необходимо принятие серьезных организационных решений на государственном уровне с выделением ведущей организации, основных разработчиков и необходимого финансирования.

ЗАКЛЮЧЕНИЕ

Построение и развитие Единого цифрового пространства знаний, вне всякого сомнения, является общегосударственной задачей, решение которой должно стать неотъемлемой частью национальных проектов и программ, направленных на развитие сферы науки, культуры, образования, здравоохранения, демографии, цифровой экономики, охраны окружающей среды, в которых фактор знания играет ключевую роль.

Сформулированные президентом РФ В.В. Путиным в послании Федеральному собранию 15 января 2020 г. установки на создание собственных технологий и стандартов «по тем направлениям, которые определяют будущее» – искусственный интеллект, генетика, новые материалы, источники энергии, цифровые технологии⁵, имеют стратегическое значение. Они, в частности, отражают качественно новую ситуацию, когда Китай принял решение о создании собственной, автономной от Запада технологической платформы [34], а России предстоит выбор между присоединением к одной из платформ или формированием собственных экосистем развития ключевых технологий.

Сегодня реализуется ряд проектов, направленных на развитие информационной инфраструктуры Российской Федерации. Отметим два из них, для которых вопросы единого цифрового пространства научных знаний важны и актуальны. В Национальном проекте «НАУКА»⁶ в числе приоритетных задач отмечен «переход к передовым цифровым технологиям, ..., создание систем обработки больших объемов данных машинного обучения и искусственного интеллекта». В национальной стратегии «О развитии искусственного интеллекта в Российской Федерации» (2019-2030 гг.), утвержденной Указом Президента РФ от 10 октября 2019 г. № 490, отмечено, что для развития искусственного интеллекта «требуется репрезентативные, релевантные и корректные наборы данных (Раздел 2, п.8), а также поставлена задача повышения доступности и качества данных, необходимых для технологии информационного общества (Раздел 5, п. 24 (в)).

Безусловно, решение этих непростых задач должно базироваться на цифровой структуре, объединяющей многоаспектную и разнородную научную

информацию. В связи с этим очевидно, что комплексный подход к построению ЕЦПНЗ в полной мере отвечает масштабу тех вызовов, ответ на которые предстоит в ближайшие годы найти российскому государству и российскому обществу.

СПИСОК ЛИТЕРАТУРЫ

1. Васильева М.М. Формирование единого информационного пространства России в условиях глобализации // Вестник московского государственного лингвистического университета. – 2013. – № 684. – С. 92-104.
2. Саенко И.Б., Куваев В.О., Алышев С.В. Подход к построению системы показателей качества единого информационного пространства // Естественные и математические науки в современном мире. – 2014. – № 14. – С. 51–56.
3. Вершинин А.П. Единое российское электронное пространство знаний: вопросы права // Информационно-аналитический журнал «Университетская книга», декабрь 2016.г. – URL: <http://www.unkniga.ru/biblioteki/bibdelo/6630-edinoe-rossiyskoe-elektronnoe-prostranstvo-znaniy-voprosy-prava.html> (дата обращения 02.02.2020)
4. Короткевич Л.С. Государственная система научной и технической информации в СССР: итоги и уроки. – М.: ВИНТИ, 1999. – 273 с.
5. Антопольский А.Б., Каленов Н.Е., Серебряков В.А., Сотников А.Н. О едином цифровом пространстве научных знаний // Вестник Российской академии наук. – 2019. – Т. 89, № 7. – С. 728-735.
6. Батоврин В.К., Гуляев Ю.В., Олейников А.Я. Обеспечение интероперабельности – основная тенденция в развитии открытых систем // Информационные технологии и вычислительные системы. – 2009. – № 5. – С.7-15.
7. Zhu LJ. (Zhu, Lijun), Shi C. (Shi, Chen), Guo JF. (Guo, Jianfeng). Mapping discovery modeling and its empirical research for the scientific and technological knowledge concept in unified concept space // Cluster Computing-the Journal of Networks Software Tools and Applications. – 2015. – Vol. 18, Iss. 1. – P. 103-112.
8. Freire N., Isaac A., Robson G., Brooks J., Manguinhas H. A survey of Web technology for metadata aggregation in cultural heritage // Information Services & Use. – 2017. – № 37(4). – P. 425-436. DOI: 10.3233/ISU-170859.
9. Marcia Lei Zeng, Philipp Mayr. Knowledge Organization Systems (KOS) in the Semantic Web: a multi-dimensional review // International Journal on Digital Libraries. – 2019. – Vol. 20, Issue 3. – P. 209–230.
10. Highlighting the central roles in scientific communities // Информация и инновации: Специальный выпуск. – 2017. – С. 145-152.
11. Dublin Core Metadata Initiative. – URL: <http://dublincore.org/> (дата обращения 25.01.2020)

⁵ Путин В.В. Послание Президента Федеральному Собранию. 15.01.2020. // Официальный сайт Президента России. – URL: <http://kremlin.ru/events/president/news/62582> (дата обращения: 28.01.2020).

⁶ Национальный проект «Наука» (2019-2024 годы). Утвержден Президиумом Совета при Президенте РФ по стратегическому развитию и национальным проектам – Протокол от 24 декабря 2018 года № 16). – URL: www.consultant.ru/document/cons_doc_LAW_31304/ (дата обращения 02.02.2020).

12. Dublin Core Metadata Terms. – URL: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> (дата обращения 25.01.2020).
13. Бездушный А.Н., Меденников А.М., Серебряков А.М., Филиппова А.А., Лопатенко А.С. Метаданные ИСИР: определение и использование // *Электронные библиотеки*. – 2001. – Т. 4, № 4. – С. 3.
14. Бездушный А.А., Бездушный А.Н., Жижченко А.Б., Кулагин М.В., Серебряков В.А. RDF-схема метаданных ИСИР // *Современные технологии в информационном обеспечении науки: сб. науч. тр.* – М., 2003. – С. 141-159.
15. Бездушный А.А., Бездушный А.Н., Нестеренко А.К., Серебряков В.А., Сысоев Т.М. Предложения по наборам метаданных для научных информационных ресурсов ЕНИП РАН // *Электронные библиотеки*. – 2004. – Т. 7, № 5. – С. 1-23.
16. Сысоев Т.М., Дьяконов И.А. Интеграция научной информации в ЕНИП // *Новые технологии в информационном обеспечении науки: сб. науч. тр.* – М., 2007. – С. 259-267.
17. Филиппов В.И., Захаров А.А. Поддержка цифровых библиотек и музейных объектов в среде ЕНИП // *Информационное обеспечение науки: новые технологии: сб. науч. тр.* – М., 2009. – С. 113-124.
18. Каленов Н.Е., Савин Г.И., Серебряков В.А., Сотников А.Н. Принципы построения и формирования электронной библиотеки "Научное наследие России" // *Программные продукты, системы и алгоритмы. Электронный журнал*. – 2012. – Т. 4, № 100. – С. 30-40. – URL: <http://www.swsys-web.ru>,
19. Погорелко К.П. Динамика использования электронной библиотеки "Научное наследие России" // *Информационное обеспечение науки: новые технологии: сб. науч. тр.* – М.: БЕН РАН, 2017. – С. 192-200.
20. Каленов Н.Е., Кириллов С.А., Соболевская И.Н., Сотников А.Н. Современное состояние электронной библиотеки "Научное наследие России" // *Труды НИИСИ РАН. Математическое и компьютерное моделирование сложных систем: теоретические и прикладные аспекты*. – 2018. – Т. 8, № 6. – С. 166-169.
21. Каленов Н.Е., Малахова И.Г. Интегрированный общедоступный информационный ресурс "История геологии и горного дела" // *Информационные ресурсы России*. – 2017. – №1. – С. 19-23.
22. Второв И.П., Малахова И.Г. Опыт эксплуатации информационной системы «История геологии и горного дела» // *Информационное обеспечение науки: новые технологии: сб. науч. тр.* – М., 2017. – С. 208-215.
23. Кириллов С.А., Соболевская И.Н., Сотников А.Н. Формирование цифрового выставочного пространства на примере виртуальной выставки «Портреты по скелетам: научное творчество М.М.Герасимова» // *Программные продукты, системы и алгоритмы. Электронный журнал*. – 2018. – № 4. – URL: <http://swsys-web.ru/en/formation-of-digital-exhibition-space.html> (дата обращения 25.01.2020)
24. Каленов Н.Е., Кириллов С.А., Соболевская И.Н., Сотников А.Н., Шубина Ю.В. Виртуальная прогулка по «Саду жизни»: опыт совместного выставочного проекта // *Труды XVIII Всероссийской научной конференции «Научный сервис в сети Интернет»*. – М.: ИПМ им. М.В. Келдыша РАН, 2016. – С. 144-147.
25. Medical Subject Headings. – URL: <https://www.nlm.nih.gov/mesh/introduction.html> (дата обращения 25.01.2020).
26. Mathematics Subject Classification. – URL: <http://msc2010.org/mediawiki/index.php?title=MSC2010> (дата обращения 25.01.2020).
27. Антопольский А.Б. Информационные ресурсы общественных наук. Опыт организации мониторинга // *Библиосфера*. – 2017. – № 3. – С. 78-84
28. Антопольский А.Б., Антонов А.М. О навигаторе информационных ресурсов по общественным наукам // *Информация и инновации*. – 2018. – Т.13, №2. – С. 18-26.
29. Антопольский А.Б., Белоозеров В.Н., Каленов Н.Е., Маркарова Т.С. О развитии терминологической базы данных в виде комплекса отраслевых информационно-поисковых тезаурусов // *Информационные ресурсы России*. – 2018. – № 5(165). – С. 22-30.
30. Антопольский А.Б., Белоозеров В.Н., Калёнов Н.Е., Шабурова Н.Н., Якшин М.М. Разработка семантической сети ключевых слов на основе дефинитивных связей // *Научно-техническая информация*. – 2017. – № 11. – С. 19-23; Antopol'skii A.B., Beloozerov V.N., Kalenov N., Shaburova N.N., Yakshin M.M. The Development of a Semantic Network of Keywords Based on Definitive Relationships // *Scientific and Technical Information Processing*. – 2017. – Т. 44, № 4. – С. 261-265. DOI: 10.3103/S0147688217040062.
31. Белоозеров В.Н., Шапкин А.В., Щуко Ю.Н. Сеть классификационных систем ВИНТИ РАН // *Программные продукты, системы и алгоритмы. Электронный журнал*. – 2018. – № 4. – URL: <http://swsys-web.ru/en/classification-systems-network-viniti-ras.html> (дата обращения 27.01.2020)
32. Kalenov N., Senko A. Interactive system of terminological dictionaries as one of the elements in the ontology of scientific knowledge // *Software Journal: Theory and Applications*. – 2019. – № 4 (electronic journal). – URL: <http://swsys-web.ru/ru/interactive-system-of-terminological-dictionaries.html> (дата обращения 02.02.2020)
33. Цветкова В.А., Мохначева Ю.В., Харыбина Т.Н., Бескаравайная Е.В., Митрошин И.А. Пространство знаний: подходы к извлечению знаний из научных текстов // *Информационные ресурсы России*. – 2019. – №2. – С. 31-34.

34. Bremmer I. The End of the American Order: Ian Bremmer speech at 2019 GZERO Summit. 18.11.2019 // Eurasia Group. – URL: <https://www.eurasiagroup.net/live-post/end-of-american-order-ian-bremmer-2019-gzero-summit-speech> (дата обращения: 23.01.2010).

Материал поступил в редакцию 06.02.20.

Сведения об авторах

АНТОПОЛЬСКИЙ Александр Борисович – доктор технических наук, главный научный сотрудник, Институт научной информации по общественным наукам РАН (ИНИОН РАН), Москва
e-mail: ale5695@yandex.ru

БОСОВ Алексей Вячеславович – доктор технических наук, главный научный сотрудник, Федеральный исследовательский центр "Информатика и управление" (ФИЦ ИУ РАН), Москва
e-mail: ABosov@ipiran.ru

САВИН Геннадий Иванович – академик РАН, научный руководитель, Межведомственный суперкомпьютерный центр РАН – филиал ФГУ НИЦ НИИСИ РАН, Москва
e-mail: gsavin@jscc.ru

СОТНИКОВ Александр Николаевич – доктор физико-математических наук, зам. директора, Межведомственный суперкомпьютерный центр РАН – филиал ФГУ НИЦ НИИСИ РАН
e-mail: asotnikov@jscc.ru

ЦВЕТКОВА Валентина Алексеевна – доктор технических наук, главный научный сотрудник, Библиотека по естественным наукам РАН (БЕН РАН), Москва
e-mail: vats08@mail.ru

КАЛЕНОВ Николай Евгеньевич – доктор технических наук, главный научный сотрудник, Межведомственный суперкомпьютерный центр РАН – филиал ФГУ НИЦ НИИСИ РАН
e-mail: nkalenov@jscc.ru

СЕРЕБРЯКОВ Владимир Алексеевич – доктор физико-математических наук, главный научный сотрудник, Федеральный исследовательский центр "Информатика и управление" (ФИЦ ИУ РАН)
e-mail: serebr@ultimeta.ru

ЕФРЕМЕНКО Дмитрий Валерьевич – доктор политических наук, заместитель директора ИНИОН РАН
e-mail: efdv2015@mail.ru