

УДК 004.65:002

А.Б. Антопольский, Д.В. Ефременко

О создании современной цифровой инфраструктуры для хранения и анализа научно-технической информации*

Обсуждаются основные принципы, рамочные условия и конкретные меры, направленные на создание современной цифровой инфраструктуры научно-технической информации в Российской Федерации, а также организационные, политико-экономические и правовые аспекты формирования цифровой инфраструктуры научно-технической информации.

Ключевые слова: научно-техническая информация, цифровая инфраструктура, национальные проекты и программы, информационные ресурсы и сервисы

ВВЕДЕНИЕ

Современные цифровые технологии оказывают существенное влияние на формы организации научно-исследовательской деятельности, способы распространения и обеспечения сохранности ее результатов. Поистине драматические изменения происходят и в научной коммуникации. Появление социальных сетей, новые требования к экспертной оценке качества научных исследований и публикаций, обрекающие все большую поддержку в научно-образовательном сообществе движения открытого доступа и открытой науки – вот лишь малая толика тех новаций, которые характеризуют сегодня как науку в целом, так, в особенности, и сферу научно-технической информации. Все эти изменения способствуют тому, что и на уровне политического руководства страны возникает потребность в постановке задач, относящихся к развитию инфраструктуры научно-технической информации.

Поводом для появления настоящей статьи стал перечень из 12 поручений по итогам заседания Совета при Президенте по науке и образованию, состоявшегося 27.11.2018 г.¹

Правительство Российской Федерации совместно с федеральным государственным бюджетным учреждением «Российская академия наук» должны представить предложения «О создании современной цифровой инфраструктуры для хранения и анализа научно-технической информации, а также для обмена такой информацией».

Президентские поручения следует рассматривать как указание на то, что модернизацию инфраструктуры научно-технической информации (НТИ) следует увязать с приоритетными направлениями национального развития, сформулированными в майском Указе Президента России² и детализированными в национальных проектах и программах (в том числе в Национальном проекте «Наука», «Национальной технологической инициативе», Национальной программе «Цифровая экономика»). Для НТИ такая увязка означает, с одной стороны, частичное сохранение преемственности модели функционирования, но, с другой стороны, разработку существенно нового содержания целей, задач, форм и технологических решений.

* Статья отражает результаты работы по исследовательскому проекту № 18-00-002\18 «Интеграция научно-информационных ресурсов учреждений РАН (на примере языкознания) как части единого цифрового пространства РАН», реализуемому при поддержке РФФИ.

¹ Перечень поручений по итогам заседания Совета при Президенте по науке и образованию. – 29 декабря 2018 года. – URL: <http://kremlin.ru/acts/assignments/orders/59632> (дата обращения: 17.03.2019).

² Указ Президента Российской Федерации от 07.05.2018 г. № 204 «О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года». – URL: <http://kremlin.ru/acts/bank/43027> (дата обращения: 17.03.2019).

Один из парадоксов современной ситуации в сфере НТИ состоит в том, что ее нормативная база – формально действующее Постановление Правительства РФ от 24 июля 1997 г. № 950 «Об утверждении Положения о государственной системе научно-технической информации»³ – фактически описывает реалии, оставшиеся в прошлом. Созданная в советское время Государственная система научно-технической информации (ГСНТИ) в начале XXI в., к сожалению, утратила свою системную целостность и – в значительной степени – функциональность. В задачи настоящей статьи не входит обсуждение причин, объективных и субъективных факторов системной деградации ГСНТИ. Подробнее позиция одного из авторов по этому вопросу изложена в работе [1], а также двух авторов – в монографии [2]. Новые тенденции развития НТИ, не имеющие отражения в действующих нормативных документах, неоднократно обсуждались в последние годы такими исследователями как Р.С. Гиляревский, И.И. Родионов, В.А. Цветкова, О.В. Сюнтюренко, Е.В. Мельникова, М.Р. Биктимиров, Н. Е. Каленов, С.С. Терещенко [3–9]. Общим для всех этих исследователей является понимание того, что в российских условиях государство должно сохранить за собой роль решающего (но не единственного!) актора научно-технической индустрии.

Разделяя такой подход, мы хотели бы обратить внимание на изменяющийся экономический, социально-политический и технологический контекст. Следует осознавать, что важнейшими особенностями развития глобальной экономической системы становятся ослабление взаимосвязи труда и капитала, еще более усиливающее влияние этого фактора развитие цифровой экономики и как следствие – реструктуризация глобальных цепочек добавленной стоимости. Все это формирует тренд, во многом противоположный глобализации конца XX – начала XXI вв. и принципам либерального капитализма. Впрочем, и в преддверии «золотого века» глобализации, и даже в её кульминационной точке сохранялось важное исключение – ключевая роль государства в научно-технической политике. Выдающийся американский историк техники М. Кранцберг, рассуждая в конце 1980-х гг. о взаимосвязи научно-технической информации и процесса принятия политических решений, подчеркивал: «Признание того, что национальная безопасность и экономика страны зависят от ее научно-технической базы, означает, что правительство все больше заботится о направлении [научно-технического развития]; в то же время внутренний динамизм науки и техники требует сложных и дорогостоящих усилий, что делает их все более зависимыми от государственной поддержки и, следовательно, от общественного одобрения» [10]. Именно благодаря финансированию важнейших НИОКР аме-

риканским государством такие компании, как *Microsoft*, *Intel* и *Apple* сумели занять лидирующие позиции в глобальной IT-индустрии. Но достичь этого им удалось в первую очередь благодаря доступу к технологическим инновациям при низком уровне собственных затрат на НИОКР [11].

В ближайшем будущем автоматизация когнитивной деятельности, роботизация, интернет вещей и т.д. еще более усилят отрыв капитала от труда, поскольку «люди не смогут конкурировать по производительности со всё более совершенными машинами» [12]. При этом еще больше возрастет роль государства, которое, с одной стороны, продолжит выступать в качестве регулятора нарастающего дисбаланса между трудом и капиталом и, с другой стороны, будет вынуждено предпринимать экстраординарные усилия по защите своего собственного экономического и научно-технологического суверенитета в новых условиях. Последнее, по сути, и означает способность к перестройке в свою пользу глобальных цепочек добавленной стоимости. Такой способностью в современном мире обладают лишь немногие государства. Россия необходимым для этого потенциалом все еще располагает, хотя по некоторым очень важным направлениям ей, скорее всего, придется пойти на формирование технологического альянса с Китаем.

Политико-экономическую сверхзадачу майского указа 2018 г. и большинства связанных с ним национальных программ и проектов, можно видеть в достижении прорыва в реструктуризации цепочек добавленной стоимости в пользу нашей страны. При таком понимании сути национальных проектов именно государство является решающим актором соответствующих трансформаций. Существующая в России модель организации научно-образовательной деятельности также предполагает приоритет государства. Соответственно, в развитии инфраструктуры научно-технической информации государство сохраняет свои ключевые позиции, оставляя при этом достаточный простор для активности негосударственных акторов научной индустрии.

Вместе с тем на сегодняшний день остаются не до конца проясненными как общий вектор развития сферы НТИ, так и ее инфраструктурные составляющие, позволяющие интегрировать науку как социальную подсистему в формирующееся цифровое общество. Выполнение президентского поручения должно, на наш взгляд, способствовать выстраиванию системы приоритетов модернизации сферы НТИ и формированию эффективной организационно-технической модели научно-информационной деятельности.

СОСТОЯНИЕ ПРОБЛЕМЫ

Для разработки предложений о создании современной цифровой инфраструктуры хранения, анализа и обмена научно-технической информацией необходимо, прежде всего, проанализировать существующее положение дел в данной области, учесть основные действующие и планируемые проекты, а также

³ Постановление Правительства РФ от 24 июля 1997 г. № 950 «Об утверждении Положения о государственной системе научно-технической информации» – URL: <http://base.garant.ru/11901351/> (дата обращения: 17.03.2019).

предложить некоторую модель системы НТИ в рамках будущей цифровой экономики России.

Перечислим основные тенденции, которые характерны для пространства научно-технической информации (инфосферы НТИ) в последние десятилетия.

1. Научные и образовательные коммуникации неуклонно переходят в цифровую форму, хотя этот процесс оказался медленнее ожидаемого и, очевидно, в обозримом будущем сохранятся обе формы коммуникации.

2. Наука и образование переходят к открытым формам коммуникации (открытому доступу), что существенно меняет экономические модели функционирования научных журналов и других научно-информационных продуктов.

3. Библиотеки демонстрируют тенденцию к разделению на два типа – культурно-досуговый и научно-информационный, причем второй тип функционально сближается с органами НТИ (ГПНТБ России, БЕН РАН, ВПТБ, ЦНМБ, ЦНСХБ и др.).

4. Генераторами общедоступных ресурсов НТИ стало множество научных и образовательных структур разного уровня, вплоть до отдельных ученых и преподавателей, в результате количество таких ресурсов разного типа возросло многократно; созданы десятки тысяч сайтов, электронных библиотек (ЭБ), баз данных (БД) и других автоматизированных информационных систем (АИС).

5. Ресурсы НТИ стали играть важную роль в оценке эффективности научных и образовательных организаций (инфометрика, вебметрика, альтметрика).

В последние годы ведется целый ряд проектов, которые призваны в той или иной степени агрегировать (интегрировать) научную и образовательную информацию в электронной форме. Перечислим некоторые из них:

- Национальная электронная библиотека (<https://нэб.рф>) (основание – Федеральный закон «О библиотечном деле», ст. 181);
- Национальный энциклопедический портал (Основание – Распоряжение Правительства РФ от 25 августа 2016 г. № 1791-р);
- Национальный портал открытого образования (<http://npoed.ru/about>);
- Электронная библиотека диссертаций (<http://diss.rsl.ru>);
- Интегратор научной периодики Научная электронная библиотека и Российский индекс научного цитирования (<https://eLibrary.ru>);
- Интеграторы библиографической информации (Либнет, ЭКБСОН, АРБИКОН);
- Интеграторы материалов научных конференций (<https://konferencii.ru>);
- Интегратор ресурсов открытого доступа (Киберленинка);
- Репозитории научно-образовательной информации (НОРА, СОЦИОНЕТ, Федеральная резервная система банков знания);
- Государственный каталог Музейного фонда РФ (<https://goskatalog.ru/portal>);
- Центральный фондовый каталог Росархива (<http://cfc.rusarchives.ru/CFC-search>).

Одновременно создается множество других информационных ресурсов – это полнотекстовые электронные библиотеки; энциклопедические, словарные, аналитические, фактографические базы данных; тематические, проблемные, видовые АИС. Каталог «Наука в Рунете» (<https://elementy.ru/catalog>) содержит около 7 тыс. ссылок, Каталог Рамблера «Наука» – 2,3 тыс. ссылок (<https://top100.rambler.ru/navi/?theme=1113>), Навигатор информационных ресурсов по общественным наукам (<http://niron.inion.ru>) – свыше 3 тыс. ссылок.

Общее количество информационных ресурсов по науке и образованию, созданных в России в электронной форме, – не менее 50 тыс. Сюда входят информационные системы федерального уровня и значения, которые поддерживаются такими авторитетными организациями как ВИНТИ РАН, ИНИОН РАН, РГБ, ГПНТБ России, БЕН РАН, ФИПС, ЦНМБ, ЦНСХБ, Президентская библиотека им. Б.Н. Ельцина, или такими популярными как Википедия, НЭБ e-Library, Киберленинка.

Следует отметить, что в последнее время в рамках вновь принятых нормативных и концептуальных правительственных документов появилось несколько новых проектов, которые можно рассматривать как направленные на организацию национальной системы НТИ. Так, в Национальном проекте «Наука», в Федеральном проекте «Развитие передовой инфраструктуры для проведения исследований и разработок в Российской Федерации» п. 1.4. сформулирован как «Цифровая система управления сервисами научной инфраструктуры коллективного пользования, включая цифровые коллекции и банки данных». Можно упомянуть также Положение о Национальной электронной библиотеке⁴, которое предусматривает интеграцию различных научно-образовательных ресурсов в Единое российское электронное пространство знаний на основе Национальной электронной библиотеки. К этой сфере относятся также нормативные акты, регламентирующие создание Национального энциклопедического портала, Государственного каталога Музейного фонда и других проектов.

ОБЩИЕ ЗАМЕЧАНИЯ

Представим наше видение основных принципов создания национальной инфраструктуры НТИ в современных условиях.

Для нашего рассмотрения понятие инфраструктуры НТИ целесообразно приравнять к понятию инфосферы НТИ, под которой мы понимаем совокупность институций, ресурсов и сервисов, используемых в научной коммуникации.

Сделаем также пояснения о соотношении научной и образовательной инфраструктуры. Во-первых, в данном контексте речь идет об информационной инфраструктуре только высшего и специального обра-

⁴ Положение о федеральной государственной информационной системе "Национальная электронная библиотека". Утверждено Постановлением Правительства РФ от 20 февраля 2019 г. № 169. – URL: <http://government.ru/docs/35825/>

зования. Очевидно, задачи информационного обеспечения общего образования должны решаться отдельно, что подтверждает произошедшее недавно разделение федеральных органов управления на Министерство науки и высшего образования РФ и Министерство просвещения РФ. Во-вторых, действительно существуют информационные ресурсы, специфические для науки и специфические для образования, а также используемые в обеих сферах, причем последних большинство. Для нашего рассмотрения разделение этих типов ресурсов и сервисов несущественно.

Интеграция существующих ресурсов и сервисов в инфраструктуру НТИ потребует мотивировать основных участников информационной деятельности к участию в соответствующих процессах. Этого можно добиться экономическими, административными и правовыми методами, однако наиболее перспективный путь – обеспечить заинтересованность научных и образовательных учреждений, отдельных ученых в коллаборации как основной методологии формирования и поддержки ресурсов НТИ. Эффективность такого подхода демонстрирует, в частности, Википедия.

Предлагаемая в настоящей статье инфраструктура НТИ должна обеспечивать преемственность в отношении Государственной системы научной и технической информации (ГСНТИ), функционирующей в современных организационных, правовых, экономических и технологических условиях.

ЦЕЛИ ИНФРАСТРУКТУРЫ НТИ

Создание современной системы научно-технической информации, ориентированной на решение приоритетных задач социально-экономического и научно-технологического развития России, обеспечение свободного доступа российских пользователей к актуальной, полной, достоверной научной и образовательной информации, создаваемой в России и странах постсоветского пространства, преимущественно в электронной (цифровой) форме – это цель работ по формированию цифровой инфраструктуры научно-технической информации.

Основные задачи инфраструктуры НТИ:

- научно-информационное обеспечение достижения целей национального развития России;
- информационное обеспечение стратегического планирования и управления научно-технологическим и инновационным развитием России, разработки и реализации научно-технологической и инновационной политики, приоритетных программ и проектов социально-экономического и научно-технологического развития России;
- координация и оптимизация научно-информационной деятельности в России, в том числе по созданию научных и образовательных электронных ресурсов;
- международное сотрудничество в области информационного обеспечения науки и образования;
- популяризация и пропаганда науки, научных знаний и научно-технических достижений;
- наукометрические исследования научно-образовательных институций России и стран постсо-

ветского пространства, а также выявление новых и перспективных научных направлений.

Перечисленные основные задачи инфраструктуры НТИ можно назвать внешними.

Кроме того, при создании инфраструктуры НТИ необходимо обеспечить решение задач, направленных на оптимизацию самой научной инфосферы, которые можно назвать внутренними:

- мониторинг научно-образовательных электронных ресурсов (НОЭР), создаваемых в России и странах постсоветского пространства;
- экспертиза НОЭР и отбор наиболее качественных из них как составная часть инфраструктуры НТИ. К этому направлению можно отнести и активную деятельность по оптимизации системы научных журналов;
- согласованное развитие качественных и востребованных НОЭР;
- организация доступа к НОЭР, навигация и поиск в них из единого окна;
- создание новых информационных продуктов и сервисов, обеспечивающих предоставление актуальной, достоверной информации в современной и удобной форме;
- разработка логико-лингвистических и программных средств, обеспечивающих обработку и преобразование НОЭР, извлечение знаний, проверку научной информации на новизну и других задач, диктуемых развитием науки и стандартами научной этики;
- согласованная программа оцифровки традиционных НОЭР;
- архивирование и сохранность НОЭР;
- обмен с международными системами представления знаний;
- стандартизация информационных процессов и продуктов;
- разработка и реализация экономической модели функционирования научно-информационных сервисов;
- совершенствование правового обеспечения инфраструктуры НТИ.

ПРИНЦИПЫ И ПОСЛЕДОВАТЕЛЬНОСТЬ ФОРМИРОВАНИЯ ИНФРАСТРУКТУРЫ НТИ

Современное состояние научной инфосферы диктует стратегию ее развития на основе конвергенции существующего хаотического множества информационных систем и ресурсов в упорядоченную совокупность знаний и данных. Эта стратегия может быть реализована в несколько этапов.

Первый этап – мониторинг инфосферы

Формирование инфраструктуры НТИ включает проведение инвентаризации научно-образовательной инфосферы, организацию постоянного мониторинга и учета научно-образовательных ресурсов и сервисов, а также научных институций – генераторов и операторов этих ресурсов и сервисов. Результатом первого этапа должен стать общегосударственный

каталог научно-образовательных электронных ресурсов. Этот каталог может быть единым, но, вероятнее, будет создано несколько каталогов по основным научным направлениям:

- естественные и точные науки;
- социальные и гуманитарные науки,
- технические науки;
- медицинские науки;
- сельскохозяйственные науки.

Прототипом такого каталога может послужить созданный в ИНИОН РАН Навигатор информационных ресурсов по общественным наукам (НИРОН) (<http://niron.inion.ru>).

Конечно, мониторинг научных и образовательных электронных ресурсов и формирование каталогов может осуществляться и на более детальном уровне – по отдельным научным дисциплинам, субдисциплинарным направлениям или по отдельным категориям этих ресурсов. Например, в настоящее время разрабатывается проект создания специализированного Центра лингвистических ресурсов на базе Института русского языка им. В.В. Виноградова РАН. Очевидно, что первичной задачей этого центра должен быть мониторинг лингвистических ресурсов России, их учет и каталогизация.

Возможно организовать учет и мониторинг по типам НОЭР или по региональным отделениям научных учреждений. Например, учет и мониторинг университетских репозиториев выполняет НЭИКОН в рамках проекта «Национальный агрегатор открытых репозиториев» (НОРА). Важно, чтобы мониторинг научно-информационных продуктов происходил согласованно.

Мониторинг инфосферы и формирование каталога должны производиться регулярно, например, ежегодно.

Второй этап – оценка и отбор НОЭР

Очевидно, что ресурсов, которые их создатели относят к категории НОЭР, слишком много, их качество далеко не всегда соответствует современным стандартам и пользовательским требованиям. Поэтому необходим второй этап – оценка и отбор наиболее качественных НОЭР.

Критериями отбора НОЭР должны стать различные показатели – в том числе:

- пользовательские (открытость, видимость, цитируемость, скачиваемость);
- содержательные (уникальность данных, полнота, актуальность, достоверность);
- правовые (легитимность, правовой статус, режим доступа);
- информационно-технологические (например, качество поиска, надежность программного обеспечения).

Важными критериями для отбора ресурсов должны стать целесообразность включения НОЭР в процессы интеграции и агрегации, а также их использование для формирования ядра системы НТИ, иначе называемого ядром ЕРЭПЗ – Единого российского электронного пространства знаний (или ЕЦПНЗ – Единого цифрового пространства научных знаний).

Назовем отобранные НОЭР *основными* ресурсами инфраструктуры НТИ.

В состав научных и образовательных электронных ресурсов цифровой инфраструктуры НТИ могут быть включены ресурсы различных типов, содержащие как первичную, так и обработанную научную и образовательную информацию, в том числе:

- электронные библиотеки;
- архивы периодических изданий;
- репозитории, другие коллекции неопубликованных документов;
- оцифрованные архивные и музейные коллекции, фонды, выставки и проч.;
- фактографические базы данных, реестры, регистры, каталоги, перечни;
- энциклопедии, справочники, указатели;
- учебная и методическая литература, онлайн-курсы;
- первичные данные научных исследований (журналы наблюдений и др.);
- географические информационные системы;
- экспертные системы, базы знаний;
- банки изображений, видео- и аудиодокументов, фильмов;
- интерактивные продукты, 3D-модели, продукты виртуальной и дополненной реальности.

Разумеется, возможно создание и других научных информационных продуктов в электронной форме, что обусловлено динамичным развитием так называемых *E-science*. Мы исходим из расширительной интерпретации *E-science* как применение компьютерных технологий для проведения современных научных исследований, включая подготовку, постановку экспериментов, сбор данных, распространение результатов, долгосрочное хранение и обеспечение доступности всех материалов, полученных в ходе научного процесса. Они могут содержать моделирование и анализ данных, электронные/оцифрованные лабораторные записи, необработанные и сгруппированные данные, черновые версии, препринты, а также печатные и / или электронные публикации [13].

В настоящее время количество российских научных и образовательных электронных ресурсов, претендующих на научную, образовательную и просветительскую функции, по экспертной оценке, составляет 50-100 тыс., поэтому процедура их отбора для включения в цифровую инфраструктуру НТИ представляется необходимой.

Отбор НОЭР может предполагать их агрегацию. Например, многочисленные электронные библиографические указатели, каталоги и базы данных, создаваемые в научных и образовательных учреждениях, могут входить в состав основных ресурсов НТИ через библиографические агрегаторы, такие как ЛИБНЕТ, ЭКБСОН, АРБИКОН.

Без отбора в число основных НОЭР вероятно должны быть включены электронные ресурсы, имеющие официальный статус государственных информационных систем, такие как Национальная электронная библиотека, АИС Федерального института промышленной собственности (ФИПС) или Государственный каталог Музейного фонда. Представляется целесообразным и

создание специализированных научно-информационных центров.

Дискуссионным является включение в основные НОЭР российских ресурсов на иностранных языках (например, переводных журналов), ресурсов, создаваемых совместно с иностранными участниками, русскоязычных ресурсов, владельцами которых являются нерезиденты, и других спорных категорий ресурсов. В то же время в число основных НОЭР могут быть включены ресурсы, создаваемые, например, в рамках Союзного государства России и Белоруссии или в процессе функционирования наднациональных органов Евразийского экономического союза.

В любом случае методика отбора НОЭР и организация их экспертизы должны стать результатом широкого обсуждения научно-информационным сообществом и быть зафиксированы в нормативно-техническом документе.

Основные НОЭР, прошедшие экспертизу и отбор, получают официальный статус и соответствующую поддержку, включая государственное архивирование, и как было отмечено, должны стать источником ядра инфраструктуры НТИ (ЕРЭПЗ, ЕЦПНЗ). Вероятно, в рамках экономической модели инфраструктуры НТИ основные НОЭР должны получить определенные предпочтения.

Третий этап – формирование ядра инфраструктуры НТИ

Формирование ядра инфраструктуры НТИ (ядра ЕРЭПЗ или ЕЦПНЗ) – третий и наиболее сложный этап создания национальной системы научной информации. Это ядро включает две части – Онтологию научных знаний и Базу знаний.

Онтология научного знания должна унаследовать терминологию, номенклатуру, парадигматику и концептуальные схемы информационных языков и систем метаданных всех основных НОЭР инфраструктуры НТИ с тем, чтобы обеспечить поиск и навигацию в этих ресурсах из единого окна. Важное условие – формирование онтологии не должно приводить к переработке основных ресурсов. Поиск в основных НОЭР следует поддерживать с помощью поисковых средств, которые были заложены в этих ресурсах.

Естественно, кроме специфических видов поиска, необходим и обычный сквозной лексический поиск для тех НОЭР, которые формируют поисковый лексический индекс.

Онтология должна обеспечивать поиск и навигацию в Базе знаний ЕРЭПЗ. Под Базой знаний здесь понимается формализованное представление сведений энциклопедического и справочного характера в форме, пригодной для компьютерной обработки. База знаний ЕРЭПЗ может представлять собой совокупность энциклопедических и справочных сведений, как в виде, пригодном для компьютерной обработки, так и в человекочитаемом виде. В качестве прототипа Базы знаний ЕРЭПЗ можно указать на DBpedia (<https://ru.wikipedia.org/wiki/DBpedia>).

Процедуру формирования исходной версии Базы знаний можно представить следующим образом:

- формирование перечня наиболее авторитетных и популярных энциклопедий, терминологических словарей и справочников по различным наукам;
- загрузка этих энциклопедий, терминологических словарей и справочников в единую среду, их интеграция;
- выявление противоречий, ошибок, неоправданного дублирования информации;
- содержательное редактирование выявленных ошибок;
- индексирование информационных объектов Базы знаний с помощью онтологии;
- разработка технологии пополнения Базы знаний.

База знаний может быть как централизованной, так и разделенной на области науки. В последнем случае должна быть сформирована междисциплинарная (общенаучная) База знаний.

Формирование Базы знаний следует осуществлять на принципах открытости, актуальности, достоверности, авторизованности тематического университета НТИ.

При этом в Базе знаний могут различаться просветительский, базовый (образовательный) и профессиональный уровни. Навигация в Базе знаний на основе онтологии также должна различаться по этим уровням.

В Базу знаний необходимо заложить возможность отражения различных (в том числе противоречивых) научных теорий и концепций, а также механизм предотвращения включения в Базу знаний лженаучной и паранаучной информации. Очевидно, что решающая роль в этом должна принадлежать РАН как основной экспертной организации страны. Информация, содержащая государственную или коммерческую тайну, в Базу знаний не включается.

На основе Базы знаний и онтологии могут разрабатываться и реализовываться дополнительные (монетизированные) сервисы, развитие которых будет способствовать повышению экономической эффективности инфраструктуры НТИ.

Четвертый этап – выявление новых знаний

Разработка системы выявления нового знания на основе потока научной информации, генерируемого российскими научными организациями, – четвертый этап формирования инфраструктуры НТИ.

Эта система должна быть основана на существующих информационно-аналитических и реферативных службах. Специалистам-аналитикам необходимо иметь в своем распоряжении сервис, позволяющий определять новизну в документе или ином научно-информационном продукте по сравнению с Базой знаний. Если такая новизна обнаружена, то новые информационные объекты (энциклопедические статьи и справочные данные) вводятся в Базу знаний. При появлении новых понятий, терминов, имен или связей между ними они также вводятся в онтологию.

Этот сервис может опираться на алгоритмы существующих систем выявления плагиата и использовать более высокий уровень семантических технологий.

Процедуры выявления нового знания и включения их в ядро инфраструктуры НТИ должны непосредственно отражаться в системе оценки эффективности научной деятельности. Эта система постепенно заменит существующие способы оценки на основе публикационной активности, которые демонстрируют устойчивое снижение своей адекватности и полезности.

Очевидна сложность и противоречивость предлагаемого подхода. Однако необходимость формирования новой системы научно-технической информации, отражающей реальность современного этапа научно-технического развития, не вызывает сомнений.

Четвертый этап формирования инфраструктуры НТИ реально может быть реализован только после появления приемлемого прототипа Базы знаний. До тех пор существующая система аналитических и реферативных служб должна сосуществовать с процедурами формирования Базы знаний.

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ИНФРАСТРУКТУРЫ НТИ

Участники

В состав инфраструктуры НТИ должны входить следующие участники.

1. Головная организация. Её основные функции:
 - организация проектирования НТИ;
 - программно-технологическая поддержка ядра НТИ;
 - разработка и поддержка стандартов инфраструктуры НТИ;
 - организация работ по созданию онтологии и базы знаний (ядра инфраструктуры НТИ) и их дальнейшая поддержка;
 - обеспечение архивации и сохранности основных НОЭР.

Функции головной организации инфраструктуры НТИ, очевидно, следует возложить на ту организацию, которая обладает наилучшими возможностями для координации научно-информационной деятельности в стране на основе современных информационных технологий. Несомненно, что этот вопрос потребует отдельного обсуждения. По нашему мнению, в настоящее время наибольшим потенциалом для этого среди организаций, подведомственных Министерству науки и высшего образования РФ, обладает Федеральный исследовательский центр «Информатика и управление».

2. Учреждения, обеспечивающие мониторинг научных и образовательных электронных ресурсов, их экспертизу, оценку и отбор основных ресурсов, координацию деятельности по их созданию, организацию информационно-аналитической деятельности, а также наукометрические исследования. Эти функции целесообразно возложить на информационно-библиотечные организации Министерства, в том числе:

- ВИНТИ РАН в сотрудничестве с БЕН РАН – естественные науки;

- ГПНТБ России – технические науки;
- ЦНСХБ РАН – сельскохозяйственные науки;
- ИНИОН РАН – общественные и гуманитарные науки.

Для медицинских наук эти функции могут быть возложены на ЦНМБ Минздрава РФ.

Как было отмечено, применительно к отдельным наукам и категориям НОЭР эти функции могут быть возложены и на другие научные учреждения. Например, для языкознания в таком качестве может выступить проектируемый Центр лингвистических ресурсов.

Точно также к этой категории могут быть отнесены организации, осуществляющие соответствующие функции в регионах России. Например, это БАН, ГПНТБ СО РАН, ЦНБ УРО РАН, а также вновь создаваемые научно-образовательные центры.

3. Операторы основных научно-образовательных электронных ресурсов, отобранных на втором этапе, например:

- Российская государственная библиотека (РГБ);
- Центр информационных технологий и систем (ЦИТиС);
- Федеральный институт промышленной собственности (ФИПС);
- МГУ им. М.В. Ломоносова;
- Президентская библиотека им. Б.Н. Ельцина;
- НЭБ e-Library;
- Национальный электронно-информационный консорциум (НЭИКОН);
- Киберленинка.

Очевидно, что количество научных, образовательных и информационно-библиотечных учреждений, осуществляющих создание и поддержку основных научных и образовательных электронных ресурсов, будет велико. Окончательный перечень операторов основных ресурсов определится в ходе проектирования цифровой инфраструктуры НТИ. Операторами могут быть организации различных организационно-правовых форм.

4. Учреждения, научные коллективы (общества) и отдельные специалисты, осуществляющие в распределенном и удаленном режиме анализ российской научно-технической информации с целью определения новизны и пополнения новыми данными НОЭР, онтологии и Базы знаний инфраструктуры НТИ.

Технологическая платформа инфраструктуры НТИ, должна поддерживать коллаборации по созданию, использованию и сохранности НОЭР, созданию и использованию онтологии и Базы знаний, совместимость с ведущими международными проектами в сфере НТИ. В качестве такой платформы может быть предложен *Semantic Web*.

Организационные принципы функционирования

Инфраструктура НТИ должна иметь межведомственный характер. Общее управление инфраструктурой НТИ, вероятно, должно быть возложено на Министерство науки и высшего образования РФ. Научное и методическое руководство может быть возложено на

Российскую академию наук или на созданный с этой целью Межведомственный координационный совет.

Инфраструктура НТИ создается и функционирует в рамках программы «Цифровая экономика» либо в тесной координации с проектами, реализуемыми в рамках этой программы.

С целью формирования и развития инфраструктуры НТИ может быть создан консорциум, в состав которого должны войти ведущие научно-информационные центры различных организационно-правовых форм, в том числе государственные, коммерческие и некоммерческие, взаимодействующие на принципах государственно-частного партнерства. Возглавлять цифровую инфраструктуру НТИ, обеспечивать организационную и финансовую поддержку её создания и функционирования должен специальный орган управления (наблюдательный совет).

В состав органов управления инфраструктуры НТИ должны входить координационные и совещательные органы, действующие в настоящее время в России в сфере научно-технической информации, в том числе:

- Межведомственный совет по развитию Национальной электронной библиотеки;
- Межведомственный координационный совет по организации предоставления лицензионного доступа к информационным наукометрическим базам данных и полнотекстовым научным ресурсам;
- Научно-издательский совет Российской академии наук;
- Межведомственная рабочая группа по вопросам, связанным с созданием на основе электронной версии Большой российской энциклопедии и других российских научных энциклопедий общенационального научно-образовательного интерактивного энциклопедического портала.

Экономическая модель

Финансирование инфраструктуры НТИ должно осуществляться из различных источников, включая субсидии из бюджетов разных уровней, грантовую и спонсорскую поддержку, а также поступления от монетизированных сервисов.

Инфраструктура НТИ должна функционировать на принципах частичной окупаемости.

Следует рассмотреть возможность создания специального фонда поддержки инфраструктуры НТИ (эндаумента).

Для основных технологических процессов создания и функционирования инфраструктуры НТИ необходимо разработать экономические нормативы и показатели. При разработке экономической модели инфраструктуры НТИ следует учесть, что ее создание позволит сократить многие существующие затраты на информационные продукты, системы и процессы.

Правовые основы

Общие принципы создания и функционирования инфраструктуры НТИ должны быть закреплены федеральным нормативным актом. В нем должен быть

определен статус (правовой режим) инфраструктуры НТИ в целом и ее составных частей.

Общий принцип функционирования инфраструктуры НТИ – это открытость и общедоступность инфраструктуры НТИ и ее составных частей. Возможные исключения могут быть сформулированы в регламенте функционирования инфраструктуры НТИ.

На правовом уровне следует обеспечить поддержку и сохранность ядра инфраструктуры НТИ и основных НОЭР.

Необходимо разработать правовой режим научных и образовательных электронных ресурсов, создаваемых в коллаборации государственных научных организаций с негосударственными и зарубежными партнерами.

В ходе проектирования инфраструктуры НТИ должны быть сформулированы необходимые изменения в информационном законодательстве, а также в 4-й части Гражданского кодекса и представлены законодателям.

МЕЖДУНАРОДНОЕ СОТРУДНИЧЕСТВО ПРИ СОЗДАНИИ И РАЗВИТИИ ИНФРАСТРУКТУРЫ НТИ

В процессе создания и дальнейшего становления инфраструктуры НТИ чрезвычайно важно анализировать опыт развития систем научной информации стран, лидирующих в науке и образовании, и внедрять положительные результаты этого опыта. Необходимо учитывать, что международное сотрудничество в области НТИ достаточно чувствительно к общему состоянию международных отношений, а совместные инициативы отдельных научных организаций, групп и объединений научного сообщества в данной сфере являются неотъемлемой частью научной дипломатии. Кроме того, отчасти продолжает функционировать созданный еще в советское время задел в кооперации с центрами научно-технической информации ряда стран, прежде всего, государств, входивших в состав СЭВ.

Нет никаких сомнений, что современная российская инфраструктура НТИ должна обеспечивать максимально полный и быстрый доступ к зарубежным источникам и массивам научно-технической информации. Очень важным представляется взаимодействие цифровой инфраструктуры НТИ с ведущими мировыми проектами по представлению знаний. Разумеется, такое взаимодействие следует осуществлять на равноправной (недискриминационной) основе. Ведущим институциям и акторам цифровой инфраструктуры НТИ необходимо своевременно реагировать на значимые инициативы зарубежных партнеров в сфере научно-технической информации, и своевременно доводить свои предложения по тематике этих инициатив до структур государственного управления, а также инициировать их экспертное обсуждение. К такого рода важным инициативам относится, например, «План S» [14], разработанный в 2018 г. ассоциацией *Science Europe* – объединением ведущих исследовательских организаций и финанси-

рующих агентств из 12 стран Европейского Союза. «План S» включает перечень принципов и основных шагов, позволяющих обеспечить открытый доступ ко всем научным публикациям, отражающим результаты исследований, выполненных за счет целевого государственного финансирования или финансирования из фондов ЕС.

В перспективе инфраструктура НТИ должна интегрировать всю научную информацию на русском языке, включая зарубежные русскоязычные ресурсы. Мы полагаем, что вполне оправданно стремиться к включению в этот информационный массив и научной информации на языках ряда других государств постсоветского пространства, включая – несмотря на удручающее состояние современных российско-украинских отношений – украинский язык. Соответственно, приоритетными для международного сотрудничества в сфере НТИ должны стать страны ЕАЭС, СНГ, ШОС и БРИКС.

Основными координаторами международного сотрудничества в сфере НТИ могут быть Международный центр научно-технической информации (МЦНТИ) и Евразийский информационно-аналитический консорциум.

ЗАКЛЮЧЕНИЕ

Таким образом, означает ли формирование цифровой инфраструктуры НТИ воссоздание в новых условиях и на другой технологической базе государственной системы научно-информационного обеспечения? Мы исходим из того, что структурные составляющие ГСНТИ, которые продолжают функционировать, производят пользующиеся спросом научного сообщества информационные продукты и услуги, а также проводят оригинальные исследования научной инфосферы, нуждаются в поддержке и развитии, что позволит им успешно интегрироваться в новую цифровую реальность. В то же время очевидно, что далеко не весь прежний функционал ГСНТИ востребован российским научным сообществом и не все задачи, поставленные перед ГСНТИ в советскую эпоху, сохраняют свою значимость для государственных институций современной России. Поэтому следует иметь в виду не сохранение элементов ГСНТИ ради самого их сохранения, а формирование эффективно действующей цифровой инфраструктуры НТИ, отвечающей актуальным политико-экономическим задачам российского государства.

Сегодня решающую роль государства в научной инфосфере нельзя сводить к простому статистическому доминированию организаций НТИ с государственной формой собственности. Создание цифровой инфраструктуры НТИ, по нашему убеждению, предполагает организационную деятельность государства, направленную на повышение эффективности научно-информационного обеспечения исследований и разработок, включая минимизацию инфраструктурных издержек. В конечном счете речь идет об обеспечении значимого вклада инфраструктуры научно-технической информации в решение политико-

экономических задач, связанных с достижением целей национального развития России в условиях ожесточенной глобальной конкуренции и перехода к цифровому обществу.

СПИСОК ЛИТЕРАТУРЫ

1. Антопольский А.Б. О возможности и целесообразности воссоздания национальной системы научно-технической информации в России // Информация и инновации: Оценки, тенденции, перспективы. – 2017. – № 1–2. – С. 15–21.
2. Антопольский А.Б., Ефременко Д.В. Инфосфера общественных наук России / под ред. В.А. Цветковой. – Москва-Берлин: Директ-Медиа, 2017. – 676 с.
3. Гиляревский Р.С., Родионов И.И., Цветкова В.А. Развитие национальной информационной инфраструктуры в научно-технической сфере // Информационные ресурсы России. – 2011. – № 5. – С. 16–18.
4. Гиляревский Р.С., Родионов И.И., Цветкова В.А. Система научной и технической информации для современной России: строим заново или учитываем имеющийся опыт? // Информационные ресурсы России. – 2016. – № 2. – С. 2–8.
5. Мельникова Е.В. Расширение функций современной системы НТИ России в контексте ее совершенствования и развития инновационной направленности // Научно-техническая информация. Сер. 1. – 2014. – № 2. – С. 9–13.
6. Каленов Н.Е. Об информационном сопровождении фундаментальных научных исследований // Научно-техническая информация. Сер. 2. – 2015. – № 4. – С. 1–5.
7. Сюттюренко О.В. Цифровая среда: тренды и риски развития // Научно-техническая информация. Сер. 1. – 2015. – № 2. – С. 1–7.
8. Терещенко С.С. Роль информационной и аналитической инфраструктуры в разработке и реализации стратегий развития науки // Научно-техническая информация. Сер. 2. – 2016. – № 11. – С. 12–19.
9. Биктимиров М.Р., Сюттюренко О.В. Цифровые информационные ресурсы современной инновационной инфраструктуры // Научные и технические библиотеки. – 2017. – № 1. – С. 8–18.
10. Kranzberg M. Interdependence of Scientific and Technological Information and Its Relation to Public Decision Making // The Annals of the American Academy of Political and Social Science. Vol. 495. Telescience: Scientific Communication in the Information Age. – Philadelphia, 1988.
11. Perelman M. Steal This Idea: Intellectual Property and the Corporate Confiscation of Creativity. – New York: Palgrave MacMillan, 2003.
12. Дисэн Г. Геоэкономика Китая и «новая холодная война». Международный дискуссионный клуб «Валдай» // Валдайские записки, № 96. – М., 2018.

13. Bohle S. What is E-science and How Should it Be Managed? // Nature.com, Spektrum der Wissenschaft (Scientific American). – URL: http://www.scilogs.com/scientific_and_medical_libraries/what-is-e-science-and-how-should-it-be-managed/ (дата обращения: 17.03.2019).
14. Plan S. Accelerating the transition to full and immediate open access to scientific publications // Science Europe. 4 September 2018. – URL: https://www.scienceurope.org/wp-content/uploads/2018/09/Plan_S.pdf (дата обращения: 17.03.2019).

Материал поступил в редакцию 18.03.19

Сведения об авторах

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

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