

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ
ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 11

Москва 2020

ОБЩИЙ РАЗДЕЛ

УДК 004.6:004.774

Е.Ю. Дмитриева, О.В. Скютюренко

Актуальные задачи диверсификации технологий, информационных продуктов и услуг

Рассматриваются объективные факторы-предпосылки корректировки концептуальных положений и задач информационного обеспечения современной научно-промышленной сферы. Показан спектр актуальных и перспективных направлений создания и внедрения новых технологий, информационных продуктов и услуг в структуре Государственной системы НТИ. Представлен вариант методологического подхода к управлению информационным обеспечением на основе оценки уровня информационной поддержки целевых комплексных программ и проектов. Структурирован потенциальный экономический макроэффект, определяемый результативностью решения комплекса задач по повышению уровня информационной поддержки исследований, разработок, трансфера технологий.

Ключевые слова: информационные продукты, конвергенция технологий, интернет-ресурсы, информационная инфраструктура, аналитическая постобработка, навигация и поиск, реферативная информация, базы данных, трансфер технологий

DOI: 10.36535/0548-0027-2020-11-1

ВВЕДЕНИЕ

Задачи повышения эффективности информационной поддержки исследований, разработок, трансфера технологий чрезвычайно актуализируются в современных сложных условиях российской экономики и стимулируют поиск новых подходов и инновационных решений. Масштаб и размах происходящих изменений обуславливают рост деструктивных инноваций (инноваций, которые изменяют соотношение ценностей на рынке) практически во всех сегментах научно-промышленной сферы. Темпы развития и распространения инноваций оказываются беспрецедентно быстрыми. Таким образом, динамичность информационной среды, развитие и внедрение новейших технологий остро ставят вопрос о формировании новых информационных продуктов, услуг и использовании новых технологий информационного обеспечения. Особую значимость и актуальность приобретает производство информационно-аналитических продуктов и услуг для поддержки развития высокотехнологичных секторов экономики, технико-экономического анализа объектов и процессов в различных разрезах, многоаспектного моделирования и прогнозирования. Формирование цифровых информационных ресурсов и становление информационных технологий (как инструментов моделирования и продуцирования новых технологий и услуг) открывают новые приложения и возможности эффективного информационного обеспечения, моделирования, прогнозирования и планирования.

Существенную корректировку концептуальных положений и задач информационного обеспечения научно-промышленной сферы определяют следующие объективные факторы:

- доминирующий тренд новой информационной среды, что проявляется в быстром росте объема цифровых данных, интернет-ресурсов и перманентном расширении глобальной сети телекоммуникаций. Развитие глобальной сети Интернет влечет смену парадигмы функционирования системы информационного обеспечения исследований и разработок – от иерархической к сетевой. Создание распределенных сетевых информационных ресурсов (ИР) является наиболее бурно развивающимся направлением информатизации научно-промышленной сферы;
- конвергенция информационных, традиционных библиотечных, компьютерных и телекоммуникационных технологий. Самоорганизация (в смысле адаптивности структуры и функциональных ролей участников) глобальной сетевой институциональной среды;
- виртуальная и реальная интеграция разнородных информационных ресурсов в гетерогенной цифровой среде. Преимущественное использование интероперабельных программных средств, унифицированных и отчуждаемых программных и технологических решений, современных сервисно-ориентированных архитектур, специализированных систем навигации к проблемно-ориентированным информационным ресурсам;
- растущая востребованность развития современных методов и средств извлечения и объединения знаний из различных типов информационных источ-

ников с последующим наложением на них различных связей между исследуемыми объектами. Здесь значительные перспективы имеет мультипликативная аналитическая постобработка научно-технической и технико-экономической информации с использованием методов наукометрии, эконометрии и многомерного анализа данных;

- актуализация задач управления знаниями и информационной поддержки принятия решений. В прагматическом аспекте – это пакет технологий, включающий в себя комплекс формализованных методов, охватывающих: поиск и извлечение знаний; структурирование и систематизацию знаний; анализ знаний (выявление зависимостей и аналогий); обновление (актуализацию) знаний; распространение знаний; генерацию новых знаний.

Эти факторы и актуальные задачи развития российской экономики требуют создания качественно новых технологий информационной поддержки научных исследований и наукоемкого производства как ключевого фактора ускоренного технологического развития. Далее мы рассмотрим спектр актуальных и перспективных направлений создания и внедрения технологий, информационных продуктов и услуг в структуре Государственной системы НТИ и, прежде всего, в ВИНТИ РАН как ее ведущей организации.

НОВЫЕ ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ СОЗДАНИЯ ИНФОРМАЦИОННЫХ ПРОДУКТОВ И УСЛУГ

В прагматическом аспекте в современной научно-промышленной сфере необходимость создания новых методов обработки, а также продуктов и услуг предопределяет устойчивые тенденции временного сжатия цикла «исследование – разработка – производство» и роста удельного веса исследований и разработок междисциплинарного характера. Интеграция информационных технологий с индустриальными коренным образом перестраивает процессы производства и актуализирует создание новых инструментов и методов информационной поддержки.

С системных позиций на содержательном уровне кратко рассмотрим континуум {A,B,C,D,E,F,G,H,I,K,L} перспективных методов, продуктов и услуг, реализация которых позволит на качественно более высоком уровне решать задачи информационного обеспечения исследований и разработок.

А. Разработка системы навигации и поиска знаний в гетерогенной сетевой среде на основе универсального интеллектуального конвертора метаданных. В современных условиях, характеризуемых лавинообразным ростом объемов научной информации, разнообразием ее видов и форм представления, задача поиска информации критически усложняется. Сегодня теория научно-технической информации не располагает методами индустриальной интеграции знаний, представленных в разнородных источниках. В мировом информационном пространстве основным видом поиска научной информации является поиск по свободной лексике (лексический поиск), на котором основаны распространенные поисковые машины (Яндекс, Google). Однако такой поиск дает низкие

характеристики полноты и точности, в частности, потому, что при нем не учитываются семантические связи понятий. В ВИНТИ РАН создана методология и ведется разработка системы, обеспечивающей эффективный поиск информации в пространстве разнородных ресурсов, содержащих данные, проиндексированные по различным системам классификации, ключевым словам, средствам полнотекстового поиска. Этот проект, поддержанный грантами РФФИ (№ 17-07-00153 и № 20-07-00103), включает разработку и реализацию алгоритмов автоматического конвертирования поисковых запросов, поступающих на естественном языке, в форму, обеспечивающую поиск информации с использованием различных классификационных языков. Онтология пространства научных знаний может быть представлена как сеть семантических связей понятий, отображаемых ключевыми словами и классификационными рубриками. Специализированная база данных, поддерживающая разработанную онтологию, будет служить основой для смысловой навигации по источникам, структурированным различными системами индексирования. Это позволит обеспечить эффективный поиск научной информации в сетевых условиях разнородности информационных ресурсов, что является исключительно важной задачей в современных условиях [1-3].

В. Реализация тематико- и/или проблемно-ориентированного избирательного распространения информации (ИРИ) с использованием интернет-СМИ и научных журналов открытого доступа. Потенциал и актуальность этого направления определяются двумя тренд-факторами.

Первое. Рост интернет-ресурсов в последние десять лет приобретает лавинообразный характер. По данным International Data Corporation (IDC), мировой объем информации удваивается каждые два года. По некоторым оценкам суммарный объем всех научных журналов в мире за один год составляет более 1 Тб информации, которая из-за огромных объемов практически необозрима в открытом доступе. Поиск нужной информации в Интернете становится все более важной и сложной задачей. Можно с уверенностью утверждать, что потенциально любой специалист мог бы найти в СМИ достаточно много интересной, новой и актуальной информации научно-технического, правового, финансового и экономического характера.

Второе. В последнее десятилетие развитые страны обращают все большее внимание на проблему открытости и доступности результатов научных исследований. В перспективе это позволит повысить прозрачность науки, сократить нерациональные затраты, существенно снизить издержки финансирования дублирующих исследований. По плану Европейской комиссии в рамках 7-летней программы Horizon 2020 (бюджет \$80 млрд), более 80% всех публикаций европейских учёных, проводящих свои исследования за государственный счёт, будут размещаться в журналах открытого доступа. Некоторые страны (Австралия, Великобритания, США, и др.) уже сейчас на самом высоком уровне занимаются решением этой проблемы [4]. В России реализуется пилотный проект КиберЛенинка, где по лицензионному дого-

вору размещаются научные журналы открытого доступа. В списке имеющихся в eLibrary журналов открытого доступа свыше 3-х тыс. наименований.

Развитие интернет-ИРИ как новой системы информационного обслуживания должно базироваться на использовании механизма кластеризации потоковой информации из открытых источников с использованием методов построения адаптивных гипермедиа на основе технологии кластеризации неструктурированных данных и обеспечения способа донесения актуальной, лингвистически обработанной информации до различных целевых групп ее потребления (и отдельных пользователей) в соответствии с их персональными потребностями и ожиданиями. С некоторой долей условности можно говорить о создании ИРИ (избирательного распространения информации) нового поколения на основе конвергенции телекоммуникационных, компьютерных и информационных технологий. Качественно новый уровень конвергированного ИРИ характеризуется: практически неограниченным кругом источников (и пользователей), предельной минимизацией временного лага, высокой целевой избирательностью. При реализации информационного комплекса должны быть использованы методы вычислительной математики и компьютерной лингвистики, предназначенные для обработки текста на естественном языке, такие как вероятностный морфологический анализ, синтаксический анализ и ранжирование, синтаксический анализ и эксплицирование отношений, установление референтных связей и др. В полном объеме реализацию ИРИ нового поколения ВИНТИ РАН мог бы осуществить в достаточно сжатые сроки [4].

С. Реализация информационного обслуживания на основе политематического полнотекстового банка данных и федерального индекса научного цитирования. Политематический полнотекстовый банк данных ВИНТИ РАН помимо зарубежных изданий должен содержать все отечественные научно-технические журналы. Для реализации этого направления следует выполнить значительный объем подготовительных работ (создание аппаратной платформы, выбор и инсталляция программного обеспечения, разработка договорной базы и правовых вопросов, проведение стоимостной оптимизации и др.). Необходимое условие для этого – первоочередное решение задач комплектования входного потока научной литературы.

В ВИНТИ РАН функционирует оригинальная система автоматизированной обработки входного потока литературы, обеспечивающая взаимодействие десятков операторов на всех стадиях технологического процесса. Система охватывает все виды изданий: журналы, книги, депонированные рукописи, описания изобретений (патентную литературу), стандарты, материалы конференций и пр. Сформированная производственная база позволяет сканировать первоисточники, поступающие как по подписке, так и из других источников (библиотеки, личные экземпляры и др.). В настоящее время эта система в основном используется для обеспечения научных отделов Института копиями статей, необходимых для создания Реферативного журнала (РЖ) и полнотекстовой Техно-

логической БД первоисточников, поступающих в ВИНТИ. На имеющемся оборудовании обрабатывается свыше 0,5 млн статей в год (свыше 2 млн страниц).

Производственный процесс включает этапы постатейной регистрации первоисточников, сканирования и распечатки результатов сканирования, а также полного системного программного сопровождения этих этапов. Система решает задачи учета поступающих экземпляров, регистрации выпусков изданий, аналитической обработки выпусков (постатейно) и тематической разметки выделенных документов. Критической позицией является разработка и/или адаптация надежно функционирующей биллинговой¹ системы и организация системы взаимных расчетов с издательствами и другими поставщиками данных.

Один из важных ресурсов информатизации науки индекс научного цитирования – это принятая в научном мире мера "значимости" научной работы. Единственным источником сведений в индексах цитирования является только сама публикация, а политика взаимодействия индексов с государством и пользователями должна быть гибкой. Российский индекс научного цитирования (РИНЦ) не является государственным в полном смысле, а принципы его формирования, не предусматривающие критериев отбора изданий и предоставляющие организациям и авторам самим корректировать свои данные, создают сложности при его использовании для решения государственных задач учета и оценки результатов научной деятельности. Необходимо разработать индекс научного цитирования, отвечающий требованиям отбора и представления необходимой наукометрической информации на более высоком качественном уровне и, главное, являющийся государственным ресурсом. Наличие в ВИНТИ РАН реферативного информационного ресурса большого объема в электронном виде и фонда первоисточников – русскоязычных журналов и книг, позволяет достаточно оперативно и на хорошем уровне начать создание информационного ресурса нового типа – федерального (государственного) индекса научного цитирования [5]. Банк данных ВИНТИ (более 36 млн док.) значительно сокращает объем работ по описанию включаемых в индекс публикаций, однако не заменяет библиографическую обработку сносок и пристатейных списков литературы – самой затратной части создания ресурса. Кроме того, банк данных включает далеко не все необходимые публикации из выбранных изданий, поэтому дополнение его ретроспективной информацией также придется предусмотреть, как и решение многих других вопросов. Было бы целесообразно рассмотреть и возможность использования разработок *Thomson-Reuter* и *Elsevier* по созданию индексов цитирования *Web of Science* и *Scopus*. Это в дальнейшем позволит интегрировать и оперативно дополнить эти мировые базы данных сведениями о российских изданиях.

¹ Автоматизированная система расчетов, ответственных за сбор информации об использовании телекоммуникационных услуг, их тарификацию, выставление счетов абонентам, обработку платежей.

Д. Реализация практического применения новых методов и систем автоматического перевода научно-технических текстов. Задача минимизации негативных факторов, обусловленных языковым «барьером» и детерминирующим использование электронных информационных ресурсов, довольно успешно решается путем более широкого внедрения в практику поиска и обработки информации систем автоматического и автоматизированного перевода. Действующие системы машинного перевода ориентированы на конкретные пары языков (например, английский и русский). Качество машинного перевода зависит от объема словаря, объема информации, приписываемой лексическим единицам, от тщательности составления и проверки алгоритмов анализа и синтеза, от эффективности программного обеспечения.

В последнее десятилетие стал доминировать статистический подход к машинному переводу. Параметры статистических моделей, на основе которых генерируется перевод, являются производными от анализа двуязычных корпусов текста (*text corpora*). Компьютеры оценивают статистические закономерности в больших массивах ранее накопленного цифрового контента. Самообучение компьютера осуществляется посредством анализа достаточно большого (сотни тысяч) количества параллельных текстов – содержащих одинаковую информацию на разных языках. Например, Евросоюз и ООН выпускают множество текстов документов на всех основных языках стран-участниц [6]. Основным преимуществом статистических систем является их свойство не отставать от развития и подвижности языка: если в языке происходят какие-либо изменения, то система сразу это распознает и самостоятельно обучается, при этом качественно перевод отличается гладкостью [7, 8]. Периодические издания ВИНТИ на русском и английском языках, огромный объем реферативного БД (более 36 млн документов, с глубиной ретроспективы по некоторым предметным областям до 15 лет), значительный входной поток информации – всё это создаст реальные предпосылки для применения новых статистических методов машинного перевода с целью повышения эффективности информационного обслуживания российских и зарубежных ученых и специалистов.

Е. Производство информационно-аналитических продуктов и услуг с использованием методов наукометрии и многомерного анализа данных. Мультипликативная аналитическая постобработка научно-технической и технико-экономической информации с использованием методов наукометрии и многомерного анализа данных позволяет выявлять статистические закономерности, выражающие зависимости между распределениями различных параметров исследуемых систем и процессов, и характер изменения распределений во времени [9-12]. Использование методов аналитической постобработки реферативной и библиографической информации представляется весьма перспективным для решения ряда задач, в числе которых:

- анализ структуры отечественной и мировой науки;
- определение тенденций и процессов, происходящих в мировой и региональной науке;

○ выявление наиболее актуальных или, напротив, теряющих свою актуальность научных направлений;

○ отслеживание генезиса конкретных научных идей и истории их развития;

○ определение продуктивности работы исследователей в конкретной научной области и эффективности материальных затрат в этой области;

○ анализ структуры научного сообщества и изучение науки как социального организма.

Исходной ресурсной базой, помимо реферативного БНД ВИНТИ, могут быть и ресурсы БНД Российского фонда фундаментальных исследований (www.rfbr.ru), Федеральных научно-технических программ (www.fcntp.ru), Росстата (www.gks.ru) и др. Совместная постобработка информации БНД ВИНТИ и данных Росстата (ВВП, произведенной энергии, среднего годового дохода на душу населения, произведенного продукта с использованием высоких технологий и ряда других) – это перспективное множество представляющих практический интерес статистических показателей и распределений, позволяющих анализировать:

- сравнительный рост ВВП, расходов на образование, исследования и разработки, объема публикаций российских авторов;

- изменения структуры ВВП и структуры публикаций российских авторов;

- зависимость роста объемов инвестиций в народное хозяйство и роста объемов публикаций (по отраслям народного хозяйства);

- зависимость роста выпуска специалистов государственных и муниципальных вузов и роста объемов публикаций (по отраслям народного хозяйства);

- прогнозировать динамику изменений показателей многомерных, например технико-экономических, объектов и процессов во времени (например, корреляции роста индекса промышленного производства по отношению к предыдущему периоду и прироста инвестиций за тот же период).

Развитие и внедрение методов и средств (продуктов и услуг) постобработки цифровых информационных ресурсов стало бы значительным вкладом как в развитие информатики, так и в становление инновационной экономики в нашей стране, а также в перспективе этот процесс мог бы трансформироваться в новое научное направление – *сетевую мультипликативную аналитическую постобработку информации*.

Ф. Создание системы информационной поддержки инновационной деятельности и трансфера технологий. Для развития инновационных процессов (в отраслях промышленности) исключительно важна информационная поддержка взаимодействия ключевых аудиторий на этапах трансфера технологий (инновационного цикла «исследование – разработка – производство»). Насущно необходима разработка проблемно-ориентированного интернет-ресурса, обеспечивающего интерактивное взаимодействие и информационную поддержку участников инновационных процессов путем создания единой интегрированной информационной среды отбора, ведения и реализации инновационных проектов. В ВИНТИ следует разработать интерактивную подсистему, в ко-

торую будут включены следующие элементы: *индикативная БД инноваций, БД потенциальных инвесторов (частных и государственных), БД предприятий и организаций*, заинтересованных в поиске и внедрении тех или иных научно-технических разработок.

Концептуальным прототипом такого интернет-ресурса является информационная служба Евросоюза CORDIS – интерактивная информационная платформа в области европейских инноваций, исследований и разработок, которая предоставляет пользователям результаты исследований и разработок по всему инновационному циклу с помощью ряда подсистем, средств и 10 поисковых БД. К настоящему времени в ней зарегистрировано свыше 300 тыс. пользователей.

Сегодня реально функционирует лишь Федеральный портал по научной и инновационной деятельности (www.sci-innov.ru). Его отличительная особенность – ориентация на весьма ограниченную тематику, определяемую перечнем приоритетных направлений развития науки, технологий и техники и перечнем критических технологий РФ. Необходимо отметить, что существенным условием его устойчивого функционирования и развития является организация участия пользователей в информационном наполнении её баз данных. Важно также обеспечение информационного взаимодействия системы информационной поддержки инновационной деятельности и трансфера технологии с государственными информационными системами (ИС РФФИ, ИС РНФ, ИС Федеральной целевой научно-технической программы), хранящими полнотекстовую информацию о результатах исследований (в том числе фундаментальных) и разработок, которые могут иметь дальнейшую промышленную коммерческую реализацию.

Г. Разработка САПР информационной поддержки работ инновационного цикла. В современных условиях для разработки и производства новой продукции актуально и необходимо использование системы автоматизированного проектирования (САПР) информационного обеспечения работ по всему инновационному циклу (так же, как и использование конструкторских САПР, или САПР технологической подготовки производства). Такая система позволит осуществлять проектирование и эффективное управление комплексным информационным обеспечением во взаимосвязи с актуализирующимися задачами и действующими производственными планами по всему распределенному во времени инновационному циклу [13].

Концептуальная основа новой технологии информационного обеспечения работ по всем этапам инновационного цикла должна базироваться на следующих основных положениях:

- автоматизированное проектирование и управление комплексным информационным обеспечением цикла исследование – разработка – производство;

- предоставление комплексной информации, включающей научно-техническую, технико-экономическую, нормативную (в том числе стандарты на изделия и процессы), прогнозно-аналитическую информацию высокой степени обработки и др.;

- функционирование системы в корпоративной интегрированной информационной среде Интранет,

и использование информационных ресурсов сети Интернет;

- количественная оценка уровня организации комплексного информационного обеспечения отдельных этапов и всего инновационного цикла в целом.

Это должно сократить временной лаг и качественно повысить интегральный уровень информационной поддержки процессов создания новой наукоемкой продукции. Опосредованно это будет содействовать повышению качества и конкурентоспособности отечественной высокотехнологичной продукции. Средой функционирования САПР информационной поддержки должна быть специализированная платформа или, что прагматически более предпочтительно, универсальная полнофункциональная автоматизированная информационно-библиотечная система (АИБС) предприятия. На базе АИБС возможно хранение и ведение как внешней технико-экономической и научной информации, так и всех типов информационных ресурсов предприятия, а также удовлетворение различных информационных запросов специалистов, осуществление депозитарной функции и поддержка однородной информационной среды.

Н. Реализация технологии информационного обслуживания на базе электронного реферативного журнала (с индикативным рефератом) в сети Интернет. Задача реферирования мирового потока научной литературы не утратила своего значения, но существенно изменилась. Следует отметить, что в развитых странах по-прежнему издается около 3 тыс. реферативных журналов (РЖ), они выходят в электронном виде и выполняют информационно-поисковые и науковедческие функции, а их рефераты становятся в основном индикативными. Это нейтрализует действие брэдфордского закона рассеяния публикаций определенной тематики по всему массиву журналов, способствует развитию национальной науки, выработке собственной терминологии и собственной информационной политики.

Основные критические пункты реструктуризации электронного РЖ:

- переориентация на индикативный реферат;
- широкое использование аннотаций (резюме) статей в журнале; возможен минимальный вариант – по каждой статье дается реферат на языке оригинала и русский текст названия и аннотации после машинного перевода (для английского, немецкого, французского языков) с постредактированием;
- автоматическое индексирование статей;
- минимизация временного лага до 1–1,5 месяцев;
- реализация режимов: электронного ИРИ, предоставления данных по произвольным выборкам и срезам, информационного мониторинга (по работам, проектам и/или программам);
- детальная подготовка и проведение, параллельно с традиционной технологией, пилотного цикла с добавлением рисунков, формул, графики в текст реферата.

И. Создание доступной через сети общего пользования базы данных по производимой и потребляемой промышленной продукции (ПППП) и стандартам России (стран СНГ, стран БРИКС, ШОС). Источ-

никами комплектования БД ПППП будут служить промышленные каталоги и буклеты, материалы выставок, ресурсы Интернета. Эта БД может существенно дополнить информационную поддержку инновационной деятельности. Её прототип – Федеральный фонд промышленных каталогов. Предполагается установить взаимодействие с Министерством промышленности и торговли России, которое работает над созданием Государственной информационной системы промышленности, предусмотренной Федеральным законом от 31.12.2014 № 488-ФЗ (ред. от 13.07.2015) "О промышленной политике в Российской Федерации". «Эта система создается в целях автоматизации процессов сбора, обработки информации, необходимой для обеспечения реализации промышленной политики и осуществления полномочий федеральных органов исполнительной власти по стимулированию деятельности в сфере промышленности, информирования о предоставляемой поддержке субъектам деятельности в сфере промышленности, а также для повышения эффективности обмена информацией о состоянии промышленности и прогнозе ее развития» (Ст.1). Функционирование БД ПППП совместно с Системой поддержки трансфера технологий (п. F) и с БД по кабинету фирм (отечественных и зарубежных) существенно повысило бы уровень информационного обеспечения промышленности РФ.

К. Создание государственной вебометрической системы цифрового пространства научных библиотек. Современный кризис информационно-библиотечной системы, который приобрел перманентный характер, обусловлен стремительным развитием телекоммуникаций и информационных технологий. В значительной степени Интернет стал важным альтернативным источником общедоступной и специальной информации для ученых, специалистов, населения в целом. Постоянно снижаются показатели основной функциональной деятельности библиотек (как публичных, так и научных): числа читателей, посещаемости, объемов книговыдачи. При этом можно констатировать, что количество обращений к веб-серверам библиотек неуклонно растет в противовес обычным посещениям, частота которых неуклонно снижается. Под воздействием внешних факторов постепенно меняется менталитет потребителя (простого читателя, специалиста, ученого) – ему не нужна книга как таковая, ему нужны знания. Для сохранения библиотечной системы как институционального компонента информационной инфраструктуры ГСНТИ вебометрическая система должна стать современным эффективным инструментом развития библиотек в цифровой среде [14].

В качестве основных задач, решаемых в процессе создания вебометрической системы библиотек, выделим:

- повышение роли и значимости публичных и научных библиотек в обществе;
- сохранение и развитие функциональной деятельности библиотек (в зависимости от их типа и вида), поддержание позитивного имиджа в мировом веб-пространстве;

- совершенствование (опосредовано) состава и структуры фондов, оптимизацию комплектования библиотек;
- интенсификацию процессов цифровизации фондов библиотек;
- стимулирование процессов диверсификации библиотечных услуг и продуктов в цифровой среде;
- мониторинг и поддержку принятия управленческих решений;
- социологический мониторинг культурного и образовательного предпочтения россиян;
- формирование интегральной оценки уровня и рейтингового распределения библиотек.

Сбор данных должен осуществляться в автоматическом режиме с использованием API² продвинутых поисковых систем (Google, Yahoo, Яндекс).

Web-система должна поддерживать режим постоянного мониторинга web-сайтов библиотек и автоматическую аналитическую постобработку результатов вебметрических исследований с использованием современных методов наукометрии и многомерного анализа данных (для анализа структур научных сайтов следует использовать методы теории графов и метод главных компонентов).

Система должна обеспечивать неэкспертное автоматическое (автоматизированное) формирование сопоставительных, рейтинговых и комплексных оценок, выявление эмпирических закономерностей, получение интегральных характеристик web-сайтов библиотек в режиме квазиреального времени.

Л. Реализация технологии формирования информационных продуктов прогнозно-аналитического и обзорного характера. Важнейшим приоритетом является воссоздание на базе новых информационных технологий традиционного для ВИНТИ направления переработки информации с выходными продуктами прогнозно-аналитического и обзорного характера. Например, подготовка ежемесячных выпусков предметно-тематических и/или проблемно-ориентированных экспресс-информационных материалов следующей структуры: краткий обзор (10 тыс. знаков); библиографический указатель (75–150 тыс. знаков). Ключевыми задачами являются: определение актуальных тематик и создание условий для привлечения к сотрудничеству квалифицированных специалистов, номинация информационных продуктов и услуг, оценка издержек и расчёт ценообразования.

В целом реализация этих задач помимо приносимых статусных и экономических выгод влияет на расширение возможностей использования результатов прогнозно-аналитической и наукометрической деятельности в научно-промышленной сфере и управлении народным хозяйством, а также создает реальную основу для:

- сопоставительного анализа структуры и уровня отечественной и мировой науки;

- определения тенденций и процессов в научно-технической сфере;
- прогнозирования развития технологий, экономики, общества;
- выявления точек роста, наиболее актуальных и/или стагнирующих научных направлений;
- мониторинга структуры (программ) отечественного научно-промышленного комплекса.

Перспективной и важной сферой деятельности ВИНТИ является выполнение целевых заказных работ обзорно-аналитического характера по крупным проектам и комплексным программам.

ОЦЕНКА УРОВНЯ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ЦЕЛЕВЫХ ПРОГРАММ И КРУПНЫХ НАУЧНО-ТЕХНИЧЕСКИХ ПРОЕКТОВ

Актуализация задач развития высокотехнологичных секторов экономики, мировые тенденции сокращения временных характеристик цикла «исследование – разработка – производство», необходимость эффективной реализации национальных проектов РФ на период 2019-2024 гг. – все это выдвигает новые, более высокие требования к организации и качеству информационного обеспечения (ИО). На примере оценки уровня информационного обеспечения целевых комплексных программ (ЦКП), в рамках которых обычно осуществляется создание качественно новых систем и образцов техники, рассмотрим основы методологического подхода к управлению информационным обеспечением. Для этого введем комплексную меру U_{Σ} .

Показатель (коэффициент) охвата комплекса работ научной и технико-экономической информацией по проблемам организаций – участников разработки и реализации одной ЦКП определяется выражением:

$$Q = \sum_{i=1}^S (g_i \cdot W_{Ni} \cdot W_{Ti}^{-1}), \quad (1)$$

где: S – количество тематических направлений, по которым необходимо осуществлять информационное обслуживание; W_{Ni} – доля исполнителей комплексной программы, которые получают информацию по i -й тематике; W_{Ti} – число исполнителей ЦКП, которым требуется информация по i -й тематике; g – приведенный вес i -й тематики (по важности, актуальности), $g_i = g'_i \cdot \sum_i g'_i$, где g'_i – вес i -й тематики.

Показатель качества информационного обеспечения R определяется как:

$$R = M^{-1} \cdot \sum_{j=1}^M (K_{0j} \cdot K_j^{-1}); \quad M \leq S \quad (2)$$

где: M – количество тематических направлений, по которым ведется информационное обслуживание; K_{0j} – реальный уровень качества информационного обеспечения разработчиков по j -му тематическому

² API (Application Programming Interface) – это специальный интерфейс или приложения, с помощью которого одна программа/приложение может взаимодействовать с другой; с помощью API различные программы и приложения могут использовать функции и ресурсы друг друга.

направлению; K_j – предельно достижимый (или оптимальный) уровень качества информационного обеспечения разработчиков по j -му тематическому направлению.

В общем случае уровень качества определяется с помощью экспертных методов оценок и представляет собой функцию вида:

$$K_j = \varphi(P, O, F, C, D), \quad (3)$$

где: P – полнота, достоверность и точность представляемой информации; O – оперативность представления вторичных и первичных документов, фактографической информации, данных логической обработки; F – форма представления и выдачи информации; C – уровень логической обработки данных; D – наличие ограничений и доступность информации.

На основании выражений (1) и (3) определяется комплексная мера уровня информационного обеспечения одной ЦКП как:

$$U_1 = Q \cdot R = M^{-1} \cdot \sum_{j=1}^M (K_{0j} \cdot K_j^{-1}) \cdot \sum_{i=1}^S (g_i \cdot W_{Ni} \cdot W_{Ti}^{-1}) \leq \alpha_0, \quad (4)$$

где: α_0 – заданный или оптимальный, или предельно достижимый уровень информационного обеспечения.

С помощью выражений 1-4 комплексная мера U_Σ уровня информационного обеспечения совокупности целевых программ (например одной отрасли) определяется как:

$$U_\Sigma = L^{-1} \cdot \sum_{i=1}^L (Q_i \cdot R_i), \quad (5)$$

Используя выражения (4) и (5), можно дать как абсолютную, так и относительную оценку уровня информационного обеспечения отдельных ЦКП (крупных проектов и/или инновационных этапов) и их совокупности, например, по отношению к тому же уровню в других ЦКП или отраслях.

Разработка методологии комплексной оценки уровня информационного обеспечения требует итерационного решения целого ряда задач связанных, прежде всего, с развитием экспертных методов оценок, формализации и определения численных значений и единиц измерения разнородных компонент (P, O, F, C, D) и их корректной нивелировки, а также накопления статистических данных для проведения сопоставительного анализа.

ЗАКЛЮЧЕНИЕ

В современных экономических условиях проблема преодоления тенденций инерционного развития национальной информационной системы требует для своего решения новых идей, новых концептуальных подходов, новых технологий, новых информационных продуктов и услуг. Особую значимость и актуальность приобретает производство информацион-

ных продуктов и услуг на основе аналитической постобработки информации, технико-экономического анализа объектов и процессов в различных разрезах, многоаспектного моделирования и прогнозирования. Широкое применение в структуре ГСНТИ цифровых информационных ресурсов, новых информационных технологий содействует более эффективному решению задач информационного обеспечения фундаментальных и прикладных исследований, инновационной деятельности [15], что обеспечивается рациональным, сбалансированным развитием информационной инфраструктуры, информационных ресурсов, информационных технологий. Комплекс задач по более глубокой переработке информации, извлечения новых знаний является частью общих проблем информационного обеспечения научных исследований и разработок и имеет определенное экономическое измерение [16]. Информационный компонент научно-технического комплекса России прямо или косвенно отражается в эффекте:

- мультипликации использования новых научно-технических результатов, знаний и информационных ресурсов;
- комплексного подхода к инвестициям и инновациям в научно-промышленной сфере;
- экономии общественно необходимого времени и материально-технических ресурсов за счет типовых проектных решений;
- трансфера технологий и использования частных технических решений (в разных отраслях).

Для реализации масштабных задач модернизации национальной информационной инфраструктуры необходимо привлечение дополнительных средств из госбюджета, государственных и целевых федеральных программ, научных и технологических фондов. В частности, такие перспективные направления как технологии Большие Данные (Big Data) и широкополосный Интернет в настоящее время существенно отстают от мирового уровня. Представляется целесообразным сформировать для ВИНТИ РАН, как головной организации ГСНТИ, целевое госзадание на общую координацию работ, разработку «дорожной карты», развитие общесистемной нормативно-методической базы, мониторинг формирования и использования цифровых информационных ресурсов, проведение научных исследований проблем сбора, аналитико-синтетической переработки, хранения, поиска, распространения и использования научно-технической информации. Следует отметить, что вне рамок настоящей работы (в силу статейных ограничений) остались вопросы, актуальные для современной российской экономики, связанные с задачами информационного обеспечения процессов импортозамещения в сфере высоких технологий, а также с необходимостью частичной конверсии оборонной промышленности для производства продукции гражданского назначения. В заключение следует отметить, что в России имеется необходимый научный, технический и экономический потенциал для формирования соответствующей требованиям времени информационной инфраструктуры, и этот потенциал необходимо реализовать в кратчайшие сроки.

СПИСОК ЛИТЕРАТУРЫ

1. Антошкова О.А., Белоозеров В.Н., Дмитриева Е.Ю., Шапкин А.В. Разработка онтологии НТИ на основе библиографических классификаций // Информационное обеспечение науки: новые технологии: Сб. научн. трудов / ред. Н.Е. Каленов, В.А. Цветкова. – Москва : БЕН РАН, 2017. – С. 292-300. – ISBN 978-5-201-13141-8.
2. Антошкова О.А., Белоозеров В.Н., Дмитриева Е.Ю. и др. Построение онтологии информационных ресурсов в виде сети библиографических классификаций // Материалы научно-практической конференции «Перспективные направления научных исследований и критические технологии в классификационных системах» (25-27 октября 2017 г.). – Москва: ВИНТИ РАН, 2017. – 96 с. – С. 20-25. – ISBN 978-5-94577-072-0.
3. Сюттюренко О.В., Белоозеров В.Н., Дмитриева Е.Ю. и др. Сеть классификаций по науке и технике как механизм смысловой навигации и поиска информации в пространстве знаний // Депонировано в ВИНТИ РАН 19.12.2019, № 120-B2019.
4. Сюттюренко О.В. Перспективы использования интернет-СМИ, журналов открытого доступа и социальных медиа // Научно-техническая информация. Сер. 1. – 2015. – № 6. – С. 30-36; Syuntyurenko O.V. Prospects for Using Online Media, Open-Access Journals, and Social Media Networks in the Field of Science and Technology // Scientific and Technical Information Processing. – 2015. – Vol. 42, № 2. – P. 112-118.
5. Биктимиров М.Р., Гиляревский Р.С., Сюттюренко О.В. Новая концептуальная основа развития информационной деятельности ВИНТИ РАН // Научно-техническая информация. Сер. 1. – 2016. – № 1. – С. 1-8; Biktimirov M.R., Gilyarevskii R.S., Syuntyurenko O.V. A New Conceptual Basis for the Development of the Information Activities of the All-Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences // Scientific and Technical Information Processing. – 2016. – Vol. 43, № 1. – P. 1-7.
6. Brynjolfsson E., McAfee A. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. – New York: Norton & Company, 2016. – 320 p.
7. Дроздова К.А. Машинный перевод: история, классификация, методы // Материалы III междунар. науч. конф. «Филологические науки в России и за рубежом» (Санкт-Петербург, июль 2015 г.). – СПб: Свое издательство, 2015. – С. 139-141. – URL <https://moluch.ru/conf/phil/archive/138/8497/> (дата обращения: 28.12.2018).
8. Колганов Д.С., Данилов Е.А. Обзор аналитической, статистической и нейронной технологии машинного перевода // Международный студенческий научный вестник. – 2018. – № 3-2. – URL: <http://eduherald.ru/ru/article/view?id=18262> (дата обращения: 28.12.2018).
9. Борисова Л.Ф., Сюттюренко О.В. Реферативный банк данных ВИНТИ РАН: перспективы постобработки информации с использованием методов анализа данных // Научно-техническая информация. Сер. 1. – 2007. – № 11. – С. 6-11.
10. Калачихин П.А. Принципы построения государственной наукометрической системы // Научно-техническая информация. Сер. 2. – 2016. – № 7. – С. 11-23; Kalachikhin P.A. The Principles of the Design of the State Scientometric System // Automatic Documentation and Mathematical Linguistics. – 2016. – Vol. 50, № 4. – P. 161-172.
11. Сюттюренко О.В., Гиляревский Р.С. Использование методов наукометрии и сопоставительного анализа данных для управления научными исследованиями по тематическим направлениям // Научно-техническая информация. Сер. 2. – 2016. – № 12. – С. 1-12.
12. Сюттюренко О.В. Цифровая среда: аналитическая постобработка информации с использованием методов наукометрии и анализа данных // Научно-техническая информация. Сер. 1. – 2019. – № 4. – С. 8-16; Syuntyurenko O.V. Digital Environment: Information Analytical Postprocessing Using the Scientometric and Data Analysis Methods // Scientific and Technical Information Processing. – 2016. – Vol. 46, № 2. – P. 59-66.
13. Сюттюренко О.В., Булычева О.С. Концептуальный облик перспективного технологического пакета информационной поддержки наукоемкого производства // Научно-техническая информация. Сер. 2. – 2016. – № 4. – С. 1-10; Syuntyurenko O.V., Bulycheva O.S. The Conceptual Form of an Advanced Technology Package for the Information Support of Knowledge-Intensive Production // Automatic Documentation and Mathematical Linguistics. – 2016. – Vol. 50, № 2. – P. 47-55.
14. Булычева О.С., Сюттюренко О.В. Концептуальные положения и предпосылки создания веб-метрической системы цифрового пространства библиотек // Сборник Президентской библиотеки. Сер. «Электронная библиотека». – 2018. – Вып. 8. – С. 19-31.
15. Сюттюренко О.В., Дмитриева Е.Ю. Государственная система научно-технической информации в структуре задач цифровой экономики // Научно-техническая информация. Сер. 1. – 2019. – № 9. – С. 1-11. Syuntyurenko O.V. Dmitrieva E.Yu. The State System for Scientific and Technical Information within the Objectives of the Digital Economy // Scientific and Technical Information Processing. – 2016. – Vol. 46, № 4. – P. 288-297.
16. Родионов И.И., Гиляревский Р.С., Цветкова В.А. Информационная деятельность как инфраструктура национальной экономики. – СПб: Издательство, 2016. – 200 с.

Материал поступил в редакцию 14.07.20.

Сведения об авторах

ДМИТРИЕВА Елена Юрьевна – кандидат технических наук, заведующая научно-методологическим отделением ВИНТИ РАН, e-mail: niipio@mail.ru

СЮНТЮРЕНКО Олег Васильевич – доктор технических наук, профессор, ведущий научный сотрудник ВИНТИ РАН, БЕН РАН, e-mail: olegasu@mail.ru