

В. И. БУДЗКО, В. И. МЕДЕННИКОВ, И. В. СОЧЕНКОВ
Федеральный исследовательский центр «Информатика и управление» РАН, Москва

ЦИФРОВОЙ ИНСТРУМЕНТ УПРАВЛЕНИЯ НАУЧНО-ОБРАЗОВАТЕЛЬНЫМИ РЕСУРСАМИ В РОССИИ

Рассматривается динамика объемов и структуры информационных научно-образовательных ресурсов, отраженных на сайтах аграрных ВУЗов и НИИ, под влиянием пандемии COVID-19 и введенных санкций с целью выработки цифрового инструмента для адекватной реакции регулирующих органов на возникшие новые запросы общества, бизнеса. Для чего рассматривается формализованное представление эффективности выполнения указанными ресурсами триединой роли: коммуникационная среда ученого сообщества, образовательная, трансферная указанных ресурсов в экономику.

Введение. Научные знания на всем протяжении развития человечества выполняли триединую роль с различным акцентом на их приоритеты во временном разрезе: совершенствование и поиск новых направлений самих научных исследований путем развития различных средств коммуникации (конференции, печатные издания трудов, открытая переписка и пр.); повышение уровня образованности различных слоев жителей за счет трансформации научных знаний в образовательные, поскольку научные знания способны выполнить функции стимулирования научно-технологического прогресса лишь при достижении населением либо ее частью некоторого уровня интеллектуального потенциала; трансфер научных знаний в экономические инновации. В эпоху всеобщей цифровизации человечества для каждой из них начали формироваться свои цифровые инструменты, зачастую, онтологически, информационно и инструментально никак не связанные между собой. Хотя современные цифровые технологии способны осуществить интеграцию научных знаний в единую цифровую платформу. Масштабные исследования указанных ролей их показало значительное пересечение этих знаний, которые назовем информационными научно-образовательными ресурсами (ИНОР), а цифровую платформу сокращенно – ЦПИНОР [1, 2]. Предлагаемый доклад посвящен разработке цифрового инструмента эффективного управления выполнением ИНОР триединой роли при ограниченности выделенных средств, адекватно реагирующего на возникшие новые запросы общества, бизнеса, связанные с пандемией и последствиями тотальных санкций.

Математическая модель эффективности выполнения ИНОР триединой роли. Как видно из изложенного, интегральная эффективность использования ИНОР зависит от эффективности использования их в каждой конкретной роли, структура ИНОР в которых одна и та же: разработки, публикации, консультационная деятельность, нормативно-правовая информация (НПИ), дистанционное обучение (ДО), пакеты прикладных программ (ППП), базы данных (БД) [1, 2]. Тогда выбор наилучшего значения интегральной эффективности использования ИНОР будет следствием решения оптимизационной трехкритериальной задачи при финансовых ограничениях на инвестиции на развитие ролей ИНОР. Для чего сведем данную задачу к виду, удобному для рассуждений в терминах оптимизационной многокритериальной задачи, пригодной для оценки эффективности использования ИНОР с последующим стимулированием развития их ролей уже каждой конкретной организации. При этом для того, чтобы уйти от большого количества форм представления ИНОР на сайтах, путем онтологического моделирования содержимого сайтов НИИ и ВУЗов введем некоторые цифровые стандарты форм хранения их в следующем виде: неупорядоченный каталог (список), упорядоченный каталог, неупорядоченное полноформатное представление и упорядоченное полноформатное представление. После чего введем выражения.

i – код роли ИНОР, $i=1, 2, 3$;

m – номер ВУЗа, $m \in M$;

n – код вида ИНОР, $n = (1, 2, \dots, 7)$;

j – код формы хранения ИНОР, $j = (1, 2, 3, 4)$;

$f_i^m(x)$ – функция эффективности i -ой роли ИНОР m -го ВУЗа;
 $x = (x_{mij})$, где x_{mij} – объем n -го вида ИНОР m -го ВУЗа для i -ой роли в j -ой форме хранения ИНОР, $x_{mij} \in X_{ni} \subset X_n$, $X_n = \bigcup_j X_{ni}$;

$G(x)$ – выделенные инвестиции на развитие ролей ИНОР для всех ВУЗов;

Z_i^m – выделенные инвестиции на развитие i -ой роли ИНОР m -му ВУЗу при ограничениях

$$\sum_{mi} Z_i^m \leq G(x). \quad (1)$$

Будем считать, что при оптимизации по Парето, как наиболее развитому подходу, руководство каждого ВУЗа стремится к увеличению функций $f_i^m(x)$. Для поиска приемлемого решения обычно используют так называемую свертку критериев путем сведения многокритериальной задачи к скалярной. Опишем наиболее популярную из них, а именно свертку методом взвешивания. В этом случае общий по m -му ВУЗу критерий W^m выглядит так

$$W^m = \max \sum \alpha^i f_i^m(x), \text{ по } x \text{ при } \sum \alpha^i = 1, 0 \leq \alpha^i, \quad (2)$$

то есть функции $f_i^m(x)$ в общий критерий входят с определенными весами α^i . Тогда распределение инвестиционных средств может быть осуществлено пропорционально величинам $W^m / \sum W^m$. Поскольку содержание каждого вида ИНОР для всех трех ролей должно отличаться по смыслу, то для кластеризации множеств X_{ni} при формировании ЦПИНОР можно было бы воспользоваться системами интеллектуального анализа текстов (text-mining) [3, 4]. В качестве такой технологии может выступать Цифровая платформа агрегации и анализа научного контента, с применением элементов которой реализованы и функционируют ряд систем информационной поддержки научно-технической деятельности на базе НИЦ «Институт им. Н.Е. Жуковского» (Экспертная информационная система поиска, анализа и учета знаний в авиастроении) и ФИЦ ИУ РАН («ИАС Приоритеты»), Цифровая платформа информационно-аналитической поддержки научно-исследовательской деятельности в области аграрно-промышленного комплекса). Перечисленные решения обладают следующим набором функций, реализованных на интегрированных источниках данных:

1. Поиск текстовой информации по запросу пользователя с поддержкой фасетного поиска по метаданным (как текстовым, так и нетекстовым) в режимах семантического, вопросно-ответного, концептного, а также поиска по отдельным ключевым словам.
2. Быстрое ознакомление с темой документов с помощью автоматически генерируемых по тексту ключевых словосочетаний.
3. Эксплоративный поиск документов (тематически похожих на эталонный документ).
4. Поиск смысловых текстовых заимствований.
5. Формирование, сопоставление и тематический анализ пользовательских коллекций документов (тематическая кластеризация, построение тематических словосочетаний и т. п.).

Эти инструменты позволяют анализировать документные полнотекстовые коллекции с метаданными, включая научные статьи, препринты, патентную информацию, авторефераты, курсовые работы, научно-образовательные материалы (выпускные квалификационные и курсовые работы, учебно-методические пособия и т. п.) и осуществлять мониторинг и выявление тенденций в наукоёмких областях с опорой на репрезентативные данные (по состоянию на весну 2022 г. в системных индексных базах данных содержится индексы свыше 100 млн документов на русском и английском языках) [5]. Таким образом, имеющийся задел может быть положен в основу создания промышленного решения – ЦПИНОР.

Анализ мониторинга контента ИНОР. На основании приведенной модели можно дать более ясное и подробное пояснение влияния размеров инвестиций на ту или иную роль ИНОР. Казалось бы, при возникших новых запросов общества, бизнеса при ограничивающих их деятельность последствий пандемии, до конца еще неосознаваемых последствий тотальных санкций, в результате которых произошло изменение структуры спроса на образовательные

услуги, на потребительские товары из-за снижения покупательной способности населения, из-за нарушений в производственно-логистических цепочках, ужесточения санитарных норм, накладывающих определенный отпечаток на материальный, образовательный и трудовой рынок, а также на доходность предприятий, Минобрнауки должен был бы кардинально изменить структуру инвестиций на разработку цифровых инструментов ролей ИНОР путем изменения весов α^i , адекватно отражающих новые реалии. Например, вместо финансирования отправки публикаций научными сотрудниками в зарубежные издания направить средства на разработку упомянутой выше ЦПИНОР, предназначенной для наиболее эффективного исполнения триединой роли ИНОР. Итог же текущих приоритетов Минобрнауки виден в представленных ниже данных, отражающих динамику изменения качества и объемов ИНОР на сайтах аграрных ВУЗов и НИИ в течение трех этапов мониторинга ВУЗов и двух этапов НИИ, последний из которых был произведен в конце 2021 г. [1, 2].

Так, у НИИ число разработок снизилось с 18806 до 5410, публикаций – с 43718 до 8274, БД – с 238 до 124, число консультантов – с 231 до 14. У ВУЗов число разработок снизилось с 4660 до 3359, БД – с 675 до 0, число консультантов – с 259 до 76. В то же время на сайтах ВУЗов произошел всплеск числа публикаций с 19401 до 41001, что связано с переходом их на удаленный формат обучения. Однако реализация их вызывает много вопросов к технологиям размещения на сайтах. Прежде всего, несистемный подход к внедрению большого количества (свыше 10) электронных библиотечных систем (ЭБС), обнаруженных на сайтах ВУЗов, онтологически никак несвязанных друг с другом, дублирующих контент, ведет к значительным финансовым издержкам, как учебных организаций, так и государства. При этом в условиях эпидемии инвестиции в ДО в виде внедрения специальных закрытых сервисов сопровождаются ликвидацией данного ИНОР из открытого контента сайтов.

Заключение. Исследования показали, что существующий цифровой разрыв между современным состоянием и потенциалом технологий разработки и наполнения содержимым сайтов продолжает стремительно увеличиваться, что демонстрирует слабую методологическую роль Минобрнауки по формированию цифрового инструмента эффективного управления ИНОР в интересах всех потребителей этих ресурсов с адекватной реакцией на возникшие новые запросы общества, бизнеса в настоящее время. А пока в работе продемонстрирована отрицательная корреляция требований общества и рынка с результатами регулирующих усилий Минобрнауки, что проявилось в тенденции резкого снижения количества и качества ИНОР на сайтах всех аграрных ВУЗов и НИИ.

Работа поддержана грантом Министерства науки и высшего образования РФ, внутренний номер 00600/2020/51896, договор № 075-15-2022-319.

ЛИТЕРАТУРА

1. **Меденников В.И., Муратова Л.Г., Сальников С.Г.** Методика оценки эффективности использования информационных научно-образовательных ресурсов. М.: Аналитик, 2017. 250 с.
2. **Меденников В.И., Муратова Л.Г., Сальников С.Г., Луппов В.В., Личман А.А.** Эффективность использования информационных интернет-ресурсов научно-исследовательских учреждений аграрного направления. М.: Аналитик. 2018. 237 с.
3. **Ядринцев В.В., Клюбина К.В., Тихомиров И.А., Гершельман А.Ф.** Выбор серверного решения для цифровой платформы поиска и анализа текстов. *Системы и средства информатики*. 2018. Т. 28. №. 3. С. 26-38.
4. Volkov S., Devyatkin D., Tikhomirov I., Sochenkov I. Data Driven Detection of Technological Trajectories. International Conference on Data Analytics and Management in Data Intensive Domains. Springer, Cham, 2020. Pp. 204–215.
5. Belenkov V.G., Budzko V.I., Devyatkin D.A., Kan A.V., Mikhailin I.S., Sochenkov I.V., Tikhomirov I.A., Shapkin V.S. Methodology for identifying centers of excellence in aviation science based on publication and patent activity. Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS). 2020. V. 32. No. 4. Pp. 21-40.

V.I. Budzko, V.I. Medennikov, I.V. Sochenkov (Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russia)

Digital tool for scientific and educational resources management in Russia

The dynamics of the volume and structure of information scientific and educational resources reflected on the websites of agricultural universities and research institutes, under the influence of the COVID-19 pandemic and the imposed sanctions, is considered in order to develop a digital tool for an adequate response of regulatory authorities to new demands from society and business. Why is a formalized representation of the effectiveness of the specified resources fulfilling a triune role: communication among the scientific community, educational, transfer of these resources to the economy.

Авторы готовы представить текст на английском языке для сборника материалов мультиконференции, который будет подан для индексирования в Scopus