

УДК 004.8:001.8:620.9

АРХИТЕКТУРА ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ПОДДЕРЖКИ ЭКСПЕРТНЫХ РЕШЕНИЙ ПО СТРАТЕГИЧЕСКОМУ ИННОВАЦИОННОМУ РАЗВИТИЮ ЭНЕРГЕТИКИ

Копайгородский Алексей Николаевич

К.т.н., ведущий специалист отдела Систем искусственного интеллекта в энергетике,

e-mail: kopaygorodsky@isem.irk.ru

Мамедов Тимур Габилевич

инженер отдела Систем искусственного интеллекта в энергетике,

e-mail: mamedowtymur@yandex.ru

Институт систем энергетики им. Л.А. Мелентьева СО РАН,

664033 г. Иркутск, ул. Лермонтова 130

Аннотация. В статье рассмотрены методы построения интеллектуальной информационной системы для поддержки экспертных решений по стратегическому инновационному развитию энергетики. Обоснована необходимость применения методов анализа Больших данных (Big Data). Представлена архитектура интегрированного хранилища интеллектуальной информационной системы, основным компонентом которой является система онтологий, объединяющая данные и знания из различных источников.

Ключевые слова: научно-технологическое прогнозирование, семантический анализ, поддержка научных исследований, архитектура информационной системы.

Цитирование: Копайгородский А.Н., Мамедов Т.Г. Архитектура интеллектуальной информационной системы для поддержки экспертных решений по стратегическому инновационному развитию энергетики // Информационные и математические технологии в науке и управлении. 2020. № 4 (20). С. 168-176. DOI: 10.38028/ESI.2020.20.4.015

Введение. Активное развитие информационно-телекоммуникационных технологий в значительной мере повлияло на сокращение времени выхода новых инновационных разработок на рынок. В первую очередь такой эффект наблюдается благодаря ускорению передачи научных знаний и разработке на их основе новых производственных технологий с последующим выпуском продукции. При этом некоторые вещи, казавшиеся еще не так давно фантастикой, входят в повседневную жизнь, существенно меняя среду обитания человека. Подобным образом новые разработки и технологии влияют и на традиционные отрасли народного хозяйства, к которым относится энергетика, являющаяся базовой и от состояния которой зависит многие отрасли экономики. Научно-технологическое прогнозирование и организация мониторинга инновационных технологических решений в области энергетики требуют разработки методов семантического анализа Big Data для выработки оценок и опережающих рекомендаций, а также создания новых инструментальных средств для поддержки этой деятельности. Необходимость развития методов анализа и обработки Big Data с помощью интеллектуальных информационных систем подчеркивается Национальной технологической инициативой (НТИ) России, а их применение в области энергетики

соотносится с рынком EnergyNet, что находит отражение в «Дорожной карте», одобренной в 2016 году Президиумом Совета при Президенте Российской Федерации по модернизации экономики и инновационному развитию России.

Управление инновациями и поиск технологических решений. В последние десятилетия наблюдается активное развитие целого спектра интеллектуальных подходов и методов для поддержки принятия решений по планированию и управлению инновационным развитием как отдельных секторов, так и экономики в целом [9, 12]. С начала 2000-х гг. была сформирована международная рабочая группа из ведущих ученых США, Европы и стран Восточной Азии, которая координирует исследования по перспективному анализу научно-технологического развития (Future Oriented Technology Analysis) [8]. Целью подобных исследований является разработка средств интеллектуальной поддержки систематического процесса обоснования возможных путей развития науки и технологий в различных областях, оценки перспективного влияния новых технологий на общество и окружающий мир, в том числе и на конкретные отраслевые инфраструктуры, а также поддержка выработки «скользящих» стратегических решений по инновационному развитию как отдельных технологических компаний, так и отраслей мировой экономики. Традиционно обоснованные научно-технологические прогнозы и программы инновационного развития являлись одной из важнейших устоявшихся форм регулирования экономики в таких странах, как США и Великобритания. Однако, в связи со стремительным развитием информационных технологий и наступлением «эпохи больших данных» (Big Data Age), в последние несколько лет особенно активно выполняются научные исследования в этой области коллективами из Китая, которые проспонсированы Национальным фондом естественных наук (National Natural Science Foundation of China – NSFC). Значительная часть методов определения новых технологических решений основана на использовании интеллектуальных семантических технологий поиска, извлечения и анализа гетерогенных данных из электронных источников информации в соответствии с концепцией Big Data [4, 5]. Методика Tech Mining применяется, как форма статистического контекстного анализа текстовых документов на базе научно-технических ресурсов для выявления прорывных технологий и разработок, оценки их инновационного потенциала [10]. Эта методика, основанная на подходе Text Mining, позволяет выявлять в фрагментах текстовых данных корреляции, связи и ранее неизвестные, но полезные с практической точки зрения интерпретируемые знания [11, 13, 14], применение которых возможно в различных сферах.

В России для решения задач по поддержке принятия решений в научно-технологическом прогнозировании энергетики используются две группы подходов. Первая группа основана на методах научно-технологического форсайта: все обзоры, оценки, прогнозы и рекомендации формируются на основе консолидированного мнения экспертов [3]. Этот подход часто критикуется за субъективность, слабую обоснованность и отсутствие ответственности экспертов, хотя и является достаточно простым, оперативным и малозатратным. Методология системного анализа относится ко второй группе и является наиболее строго научно обоснованным подходом, поскольку использует в качестве основных инструментов методы системного оценивания и сопоставления технологий и системное технологическое моделирование (математические модели технологий и их структуры). В

математических моделях учитываются не только энергетические технологии по всей совокупности их основных характеристик, но и целый комплекс внешних факторов (потребности, ресурсы, экология, функционирование, институты и т.п.) [1]. Применение указанных инструментов для разработки программ инновационного развития в энергетике требует выполнения полномасштабных научно-исследовательских работ силами высококвалифицированных коллективов, больших финансовых затрат и достаточное количество времени. Результаты таких исследований, как правило, очень высокого качества и предлагают отличные решения для конкретных отраслевых научно-технических задач и задач формирования обобщенных стратегий развития энергетики [2]. Поскольку исследования выполняются значительное время, возникает проблема отсутствия эффективных механизмов оперативной адаптации аппарата системного технологического моделирования к меняющимся условиям технологического развития. Особенно это проявляется в части учета факторов научно-технического прогресса в смежных или слабосвязанных областях, в сложности учета междисциплинарных и инфраструктурных технологических эффектов и в возросшей информационной, технической и организационной комплексности объектов исследования.

Предлагаемые методы и подходы для поддержки прогнозирования инновационного развития энергетики. Методы прогнозирования развития энергетической отрасли, основанные на традиционных математических моделях и программных комплексах, не всегда эффективны в условиях неопределенности и отсутствия необходимой достоверной информации для имеющихся моделей. Для решения поставленных задач предлагается использовать интеллектуальные методы семантического анализа, машинного обучения и технологии Big Data для создания инструментария, облегчающего работу экспертных групп, и инструментария, выполняющего предварительную обработку информации, анализируемой экспертами. Источниками информации могут выступать Открытые данные (Open Data) и Большие данные (Big Data). Кроме того, могут эффективно использоваться методы семантического моделирования для «экспресс-анализа» собранной информации [7], разработка которых ведется в отделе Систем искусственного интеллекта в энергетике Института систем энергетики им Л.А. Мелентьева Сибирского отделения РАН (ИСЭМ СО РАН)¹.

Авторами предлагается использовать методы построения онтологического пространства знаний в энергетике и семантического моделирования: когнитивного (для описания причинно-следственных связей), событийного (для моделирования последствий принимаемых решений) и вероятностного (для оценки перспектив и рисков) моделирования. Предлагается использование паттернов (шаблонов), основанных на экспертных суждениях (прогнозных гипотезах) об инновационных путях развития энергетики и перспективных технологических решениях в энергетике, которые могут использоваться для поиска информации и проверки гипотез. В общем случае гипотезы формируются на основе анализа массива собранных данных. Поскольку данные собираются из открытых источников, они носят явно статистический характер и к ним применимы статистические методы, в том числе методы статистической оценки гипотез и критериев. Ввиду большого объема собираемых

¹ <http://isem.irk.ru/institute/departments/>

данных целесообразно применение второго уровня статистической значимости (0,01 или 1%), что задает уровень статистической ошибки первого рода не более 1% и достоверность получаемых результатов при проверке гипотез не менее 99%.

Использование в качестве источников данных для анализа и последующего научно-технологического прогнозирования развития энергетики исключительно российских информационных систем и тематических баз данных является ошибочным, поскольку из-за глобализации инновационные разработки и результаты научных исследований не ограничены экономикой одной отдельно взятой страны или макрорегиона. Аналогичным образом обосновывается и необходимость анализа информации не только на русском, но и на английском языке. Наличие множества публикаций на русском языке является одним из индикаторов готовности к практическому внедрению новой технологии в производственные процессы на предприятиях и организациях России.

Кроме того, является целесообразным применение методов и математических моделей исследования направлений развития энергетики, которые неоднократно использовались в ИСЭМ СО РАН при разработке стратегий развития энергетики России. Коллективом отдела Систем искусственного интеллекта в энергетике под руководством д.т.н., проф. Л.В. Массель предложена и реализована двухуровневая технология исследований проблем энергетической безопасности [6], в которой на первом, верхнем уровне экспертами выполняется качественный анализ вариантов развития энергетики с использованием семантических технологий, а на втором (нижнем) уровне для количественного обоснования выбранных вариантов используются экономико-математические модели ТЭК России большой размерности и традиционные программные комплексы для многовариантных расчетов (например, программный комплекс ИНТЭК-М).

Архитектура интеллектуальной информационной системы. Одним из основных компонентов интеллектуальной информационной системы для поддержки экспертных решений является интегрированное хранилище (база знаний), включающее семантические модели, а также методы семантического анализа Big Data в виде программных компонентов (агентов). Архитектура интегрированного хранилища интеллектуальной информационной системы и основные его компоненты представлены на рисунке 1. Важным компонентом хранилища является система онтологий, которая включает онтологии отраслей энергетики и отдельных энергетических технологий, онтологию исследований энергетики, онтологии задач, методов исследований и др., используется для интеграции знаний и позволяет унифицировать описания методов анализа, форматов, используемых в них данных, спецификации реализующих эти методы программных компонентов. Подсистема внесения информации инкапсулирует программы-краулеры и адаптеры источников данных, используемые для поиска, построения описания и загрузки потенциально интересной информации в хранилище. Подсистема поиска является точкой получения информации (метаданных, ссылок на внешние ресурсы и документов исследователей, описанных в понятиях системы онтологий), которая после извлечения может быть представлена пользователю или обработана другими программными агентами. Средства коллективной работы, научные прототипы которых были реализованы под руководством Копайгородского А.Н., позволяют разрабатывать, описывать и управлять документами, создаваемыми

исследователями и отражающими некоторые полученные научные результаты. В качестве редактора онтологий предполагается применение инструментального средства OnotMap, которое имеет расширяемую архитектуру и может быть использовано как средство конфигурирования, задающее специфику предметной области и решаемой задачи. Указанная особенность интеллектуальной информационной системы является ключевой и позволяет использовать некоторые ее компоненты для решения задач, например, в области научного обоснования стратегических решений по цифровой трансформации энергетики. Это достигается за счет более позднего наложения ограничений на предметную область, что значительно увеличивает гибкость системы. Кроме того, предлагается применение сервис-ориентированного подхода, упрощающего использование как оригинальных авторских, так и имеющихся в свободном доступе готовых реализаций методов, что позволяет снять ряд проблем, связанных с различием платформ, стандартов, форматов данных, а также с удаленным размещением реализаций этих методов.



Рис. 1. Архитектура интегрированного хранилища интеллектуальной информационной системы.

Исходя из опыта реализации других научных проектов, связанных со сбором и обработкой большого объема данных, было принято решение о реализации хранилища метаданных в виде распределенной базы данных (рис. 2). Для улучшения производительности в хранилище метаданных индексы выделяются в отдельные структуры. Доступ к файлам, содержащим текстовые данные, которые хранятся на отдельных носителях, осуществляется через дескрипторы, размещенные в базах метаданных.

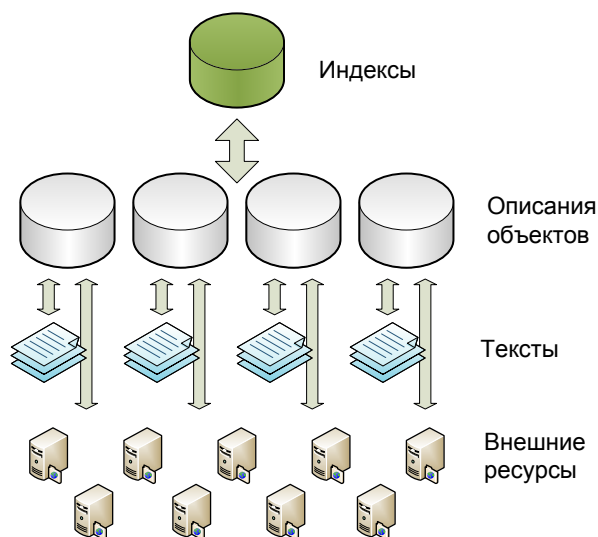


Рис. 2. Распределенная архитектура хранилища метаданных.

На рисунке 3 показана архитектура интеллектуальной информационной системы, включающая средства семантического анализа тактовых данных, онтологического, когнитивного и событийного моделирования, средства проверки гипотез и визуализации результатов поиска. При обработке результатов поисковых запросов во внешних, по отношению к интегрированному хранилищу, программных средствах извлекаемых данных исследователи получают новые знания, которые могут быть представлены в явной форме и загружены в интегрированное хранилище с использованием средств коллективной работы.

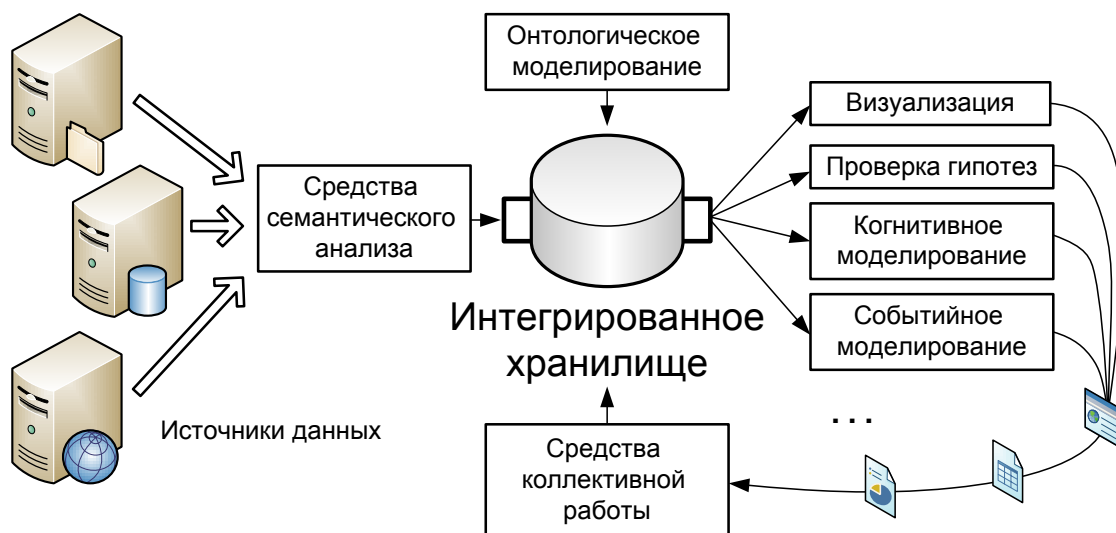


Рис. 3. Архитектура интеллектуальной информационной системы.

Применение для связи с внешними компонентами стандартизированных интерфейсов и сетевых протоколов Интернет, в частности JSON и HTTP, обеспечивает интеграцию методов при решении нетривиальных задач и позволяет получить доступ к информации и функциональным компонентам, размещенным на удаленных серверах.

Заключение. Реализация интеллектуальной информационной системы с применением сервис-ориентированного подхода и интеграцией существующих сторонних и авторских компонентов позволит эффективно решать научно-практические задачи, как в

области поддержки экспертных решений по стратегическому инновационному развитию энергетики, так и в области научного обоснования стратегических решений по цифровой трансформации энергетики. Применение предложенных методов и подходов для проектирования и реализации отдельных компонентов интеллектуальной информационной системы обеспечивает достаточный уровень ее гибкости и расширяемости.

Работа выполнена при частичной финансовой поддержке грантов РФФИ №20-07-00994 и № 19-07-00351.

СПИСОК ЛИТЕРАТУРЫ

1. Системные исследования в энергетике: ретроспектива научных направлений СЭИ-ИСЭМ / отв. ред. Н.И. Воропай. Новосибирск: Наука. 2010. 686 с.
2. Энергетика XXI века. Условия развития. Технологии. Прогнозы. / Отв. ред. Н.И. Воропай. Новосибирск: Наука. 2010. 386 с.
3. Прогноз научно-технологического развития России: 2030 / под. ред. Л.М. Гохберга, С.П. Филиппова. Москва: Министерство образования и науки Российской Федерации, Национальный исследовательский университет «Высшая школа экономики». 2014. 244 с.
4. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. М.: Манн, Иванов и Фербер. 2014. 230 с.
5. Н. Марц, Дж. Уоррен. Большие данные. Принципы и практика построения масштабируемых систем обработки данных в реальном времени. М.: Вильямс. 2016. 368 с. ISBN: 978-5-8459-2075-1
6. Массель Л.В. Создание и интеграция интеллектуальных информационных технологий и ресурсов для комплексных исследований в энергетике // Вестник РФФИ. 2012. №4 С. 74-81
7. Массель Л.В., Массель А.Г. Интеллектуальные вычисления в исследованиях направлений развития энергетики // Известия Томского политехнического университета. 2012. Т. 321. № 5. Управление, вычислительная техника и информатика. С. 135-141.
8. Future-Oriented Technology Analysis. Strategic Intelligence for an Innovative Economy. Springer. 2008. 170 p. DOI: 10.1007/978-3-540-68811-2.
9. V. Coates, et al. On the future of technological forecasting // Technol. Forecast. Soc. Change. 67 (1). 2001. Pp. 1 – 17.
10. Cunningham S.W., Porter, A.L., and Newman, N.C. Tech Mining, special issue of Technological Forecasting & Social Change. 73 (8). 2006. Pp. 915-1060.
11. Mirhosseini M.A Clustering Approach using a Combination of Gravitational Search Algorithm and k-Harmonic Means and its Application in Text Document Clustering // Turkish Journal of Electrical Engineering and Computer Sciences. Vol. 25. № 2. 2017. Pp. 1251-1262. DOI: 10.3906/elk-1508-31.
12. Zhang Y., et al. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. Technol. Forecast. Soc. Change. 2016. DOI: 10.1016/j.techfore.2016.01.015.
13. Zhang M.L, Zhou Z.H. Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization // IEEE Transactions on Knowledge and Data Engineering. Vol. 18. № 10. 2006. Pp. 1338-1351. DOI:10.1109/TKDE.2006.162.
14. Zheng L., Noroozi V., Yu P.S. Joint Deep Modeling of Users and Items using Reviews for Recommendation. 2017. DOI:10.1145/3018661.3018665.

UDK 004.8:001.8:620.9

**ARCHITECTURE OF THE INTELLECTUAL INFORMATION SYSTEM TO SUPPORT
EXPERT DECISIONS ON STRATEGIC INNOVATIVE ENERGY DEVELOPMENT**

Alex N. Kopaygorodsky

Ph.D., Leading specialist, Department of Artificial Intelligence Systems,

e-mail: kopaygorodsky@isem.irk.ru,

Timur G. Mamedov

Graduate student, Department of Artificial Intelligence Systems,

e-mail: mamedowtymur@yandex.ru,

Melentiev Energy Systems Institute

Siberian Branch of the Russian Academy of Sciences

130, Lermontov Str., 664033, Irkutsk, Russia.

Abstract. The article discusses methods of building an intelligent information system to support expert decisions on strategic innovative development of the energy sector. The necessity of using Big Data analysis methods has been substantiated. Architecture of an integrated repository of an intelligent information system is presented, in which the main component is a system of ontologies on the basis of which information, data and knowledge from various sources are combined.

Keywords: scientific and technological forecasting, semantic analysis, support for scientific research, information system architecture

References

1. Sistemnye issledovaniya v jenergetike: retrospektiva nauchnyh napravlenij SEI-ISEM [Systems research in energy: a retrospective of the scientific directions of SEI-ISEM] / Ed. N.I. Voropay. Novosibirsk: Nauka = Science. 2010. 686 p.
2. Energetika XXI veka. Uslovija razvitija. Tehnologii. Prognozy. [Energy of the XXI century. Development conditions. Technologies. Forecasts.] / Ed. N.I. Voropay. Novosibirsk: Nauka = Science. 2010. 386 p.
3. Prognoz nauchno-tehnologicheskogo razvitija Rossii: 2030 [Forecast of scientific and technological development of Russia: 2030] / Ed. L.M. Gohberg, S.P. Filippov. Moscow: Ministerstvo obrazovaniya i nauki Rossijskoj Federacii, Nacional'nyj issledovatel'skij universitet «Vysshaja shkola jekonomiki» = Ministry of Education and Science of the Russian Federation, National Research University Higher School of Economic. 2014. 244 p.
4. Mayer-Shenberger V., Cukier K. Bol'shie dannye. Revoljucija, kotoraja izmenit to, kak my zhivem, rabotaem i myslim [Big Data. A Revolution That Will Transform How We Live, Work and Think]/ Trans. from Eng. I. Gayduk. Moscow: Mani, Ivanov and Ferber. 2014. 230 p.
5. Nathan Marz, James Warren. Bol'shie dannye. Principy i praktika postroenija masshtabiruemyh sistem obrabotki dannyh v real'nom vremeni [Big Data: Principles and best practices of scalable realtime data systems]. Moscow: Wilyams = Moscow.: Williams. 2016. 368 p. ISBN: 978-5-8459-2075-1
6. Massel' L.V. Sozdanie i integracija intellektual'nyh informacionnyh tehnologij i resursov dlja kompleksnyh issledovanij v jenergetike [Creation and integration of intelligent information

technologies and resources for complex research in energy] // RFBR Bulletin. 2012. №4. Pp. 74-81

7. Massel' L.V., Massel' A.G. Intellektual'nye vychislenija v issledovanijah napravlenij razvitija jenergetiki [Intelligent computing in studies of energy development directions] // Bulletin of the Tomsk Polytechnic University. 2012. Vol. 321. № 5. Management, computer technology and informatics. Pp. 135-141.

8. Future-Oriented Technology Analysis. Strategic Intelligence for an Innovative Economy. Springer. 2008. 170 p. DOI: 10.1007/978-3-540-68811-2.

9. V. Coates, et al. On the future of technological forecasting // Technol. Forecast. Soc. Change. 67 (1). 2001. Pp. 1 – 17.

10. Cunningham S.W., Porter A.L., and Newman, N.C. Tech Mining, special issue of Technological Forecasting & Social Change. 73 (8). 2006. Pp. 915-1060.

11. Mirhosseini M.A Clustering Approach using a Combination of Gravitational Search Algorithm and k-Harmonic Means and its Application in Text Document Clustering // Turkish Journal of Electrical Engineering and Computer Sciences. Vol. 25. № 2. 2017. Pp. 1251-1262. DOI: 10.3906/elk-1508-31.

12. Zhang Y., et al. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. Technol. Forecast. Soc. Change. 2016. DOI: 10.1016/j.techfore.2016.01.015.

13. Zhang M.L, Zhou Z.H. Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization // IEEE Transactions on Knowledge and Data Engineering. Vol. 18. № 10. 2006. Pp. 1338-1351. DOI:10.1109/TKDE.2006.162.

14. Zheng L., Noroozi V., Yu P.S. Joint Deep Modeling of Users and Items using Reviews for Recommendation. 2017. DOI:10.1145/3018661.3018665.