

XV Международная научно-практическая конференция студентов аспирантов и молодых учёных
«Молодёжь и современные информационные технологии»

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ЭКСПЕРТНЫХ РЕШЕНИЙ НА ОСНОВЕ АНАЛИЗА ОТКРЫТЫХ ИСТОЧНИКОВ

А.Н. Копайгородский
Институт систем энергетики им. Л.А. Мелентьева СО РАН
kopaygorodsky@mail.ru

Введение

Ввиду сложных мировых экономических условий исследования перспектив развития отраслей становятся особенно актуальными. В первую очередь это связано с желанием предсказать с высокой достоверностью перспективность инновационных проектов и желанием сэкономить. Энергетика является одной из базовых отраслей экономики и от ее развития во многом зависит производство конечной продукции. В Институте систем энергетики им. Л.А. Мелентьева СО РАН разрабатывается интеллектуальная информационная система для поддержки коллективной экспертной деятельности для прогнозирования технологи-ческого развития энергетики России.

Прогнозирование развития энергетики

Существуют несколько основных методов технологического прогнозирования, которые используются для исследований энергетики России:

1. Экспертные опросы являются простыми дешевыми и быстрыми решениями, но эти методы имеют такие плохие качества как субъективность, низкая обоснованность, отсутствие ответственности.
2. Технологический анализ на уровне отдельных энергетических компаний основывается на эконометрике, мониторинге, базах данных по разработкам, оперативном анализе проблем и предложений.
3. Системный анализ технологий требует достаточно много времени и усилий высококвалифицированных научных коллективов.

Исследователи из ИСЭМ не обеспечены в должной мере данными и актуальными знаниями для прогнозирования энергетики на высоком уровне. Технологический прогноз энергетики РФ до 2035 года рассматривает три разных сценария развития мировой энергетики [2], при этом вероятность каждого достаточно сложно оценить. Данное обстоятельство приводит к высокой неопределенности в стратегическом планировании инновационного развития энергетики РФ из-за высоких инвестиционных рисков.

Применение семантических технологий облегчает научно-технологическое прогнозирование в энергетике. Для этого могут быть использованы такие технологии как:

- семантический информационный поиск;

- анализ слабоструктурированных текстов и данных, включая извлечение терминов, концептов, их согласование и агрегирование;
- статистическая обработка и анализ полученных семантически структурированных данных, включая расчет количественных показателей и индикаторов развития технологий;
- когнитивные технологии визуализации;
- онтологическое моделирование.

Задача разработки инструментальных средств для поддержки прогнозирования развития энергетики состоит из двух основных частей:

- создание Интеллектуальной информационной среды для совместной работы экспертов и накопления знаний;
- совершенствование существующих и создание новых методов поиска информации, знаний и прогнозирования.

Поиск новых технологических решений в энергетике

Новые подходы к извлечению и анализу знаний и данных позволяют улучшить качество получаемых научных результатов. Возможно уменьшение неопределенности с помощью анализа больших данных (Large Data Analytics – LDA), который может быть применен к глобальным источникам научных и технологических данных и инноваций. При применении LDA можно определить уже существующие и спрогнозировать новые тенденции развития, а также предусмотреть технологические прорывы путем всестороннего понимания протекающих инновационных процессов [3].

В качестве источников информации для прогнозирования развития исследователи могут использовать базы данных по патентам и изобретениям, новостные ленты, информацию из государственных информационных систем, которые содержат данные о научно-технических проектах и разработках (ЦИТИС, РФФИ, ФИПС и пр.), а также из различных коммерческих систем (SCOPUS, Web of Science, РИНЦ, Science Index).

Задача обеспечения актуальной информацией исследователей является наиболее важной. В процессе работы исследователи используют некоторые уже существующие модели и создают новые. Новая информация и знания позволяют выполнить актуализацию существующих моделей. Большой поток слабоструктурированной информации существенно замедляет и усложняет построение и обновление моделей. В этом случае

исследователи вынуждены тратить достаточно много времени на поиск и выявление интересных их фактов, что является крайне малоэффективным. Для эффективного извлечения знаний из больших наборов данных требуется их автоматическая классификация.

Методический подход к сбору знаний из текстовых источников состоит в решении некоторых задач:

- извлечение необработанных полуструктурированных текстовых данных;
- предварительная очистка, идентификация и извлечение текста;
- классификация текста, составление дайджеста для каждого обработанного документа.

Полученные элементы дайджеста могут быть сопоставлены с понятиями единой системы

онтологий и сохранены в базе знаний со ссылками на исходные документы.

Интеллектуальная информационная система

На рисунке 1 представлена архитектура Интеллектуальной информационной системы прогнозирования инновационного развития энергетики (ИИСПИР). Подсистемы поиска и анализа получают новые документы, выполняют их классификацию и сопоставление с понятиями системы онтологий. Дайджесты документов размещаются в хранилище, к которому через Web-сервисы имеют доступ исследователи. Доступ к существующим в ИСЭМ базам данных осуществляется через интерфейсы хранилища данных.

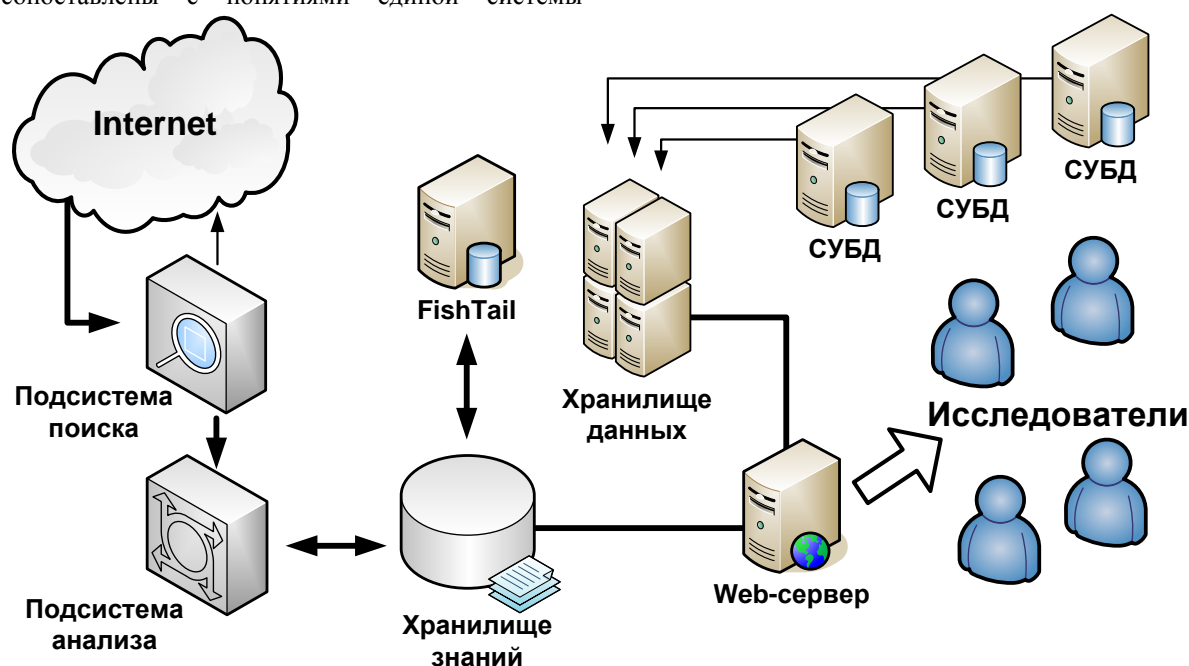


Рис. 1. Архитектура Интеллектуальной информационной системы прогнозирования инновационного развития энергетики.

Заключение

Одной из основных проблем поддержки прогнозирования развития энергетики России является необходимость обработки достаточного большого объема разнородной информации. Разрабатываемая в ИСЭМ интеллектуальная информационная система осуществляет сбор и обработку информации, получаемой открытых источников. В процессе поиска информации выполняется ее классификация на основе выделяемых ключевых слов, сопоставленных с понятиями, которые введены в системе онтологий. Собранный в автоматическом режиме информация размещается в специализированном хранилище, маркируется и агрегируется. Предложенный подход способствует улучшению качества прогнозов в исследованиях развития энергетики

Список использованных источников

1. Прогноз научно-технологического развития отраслей топливно-энергетического комплекса России на период до 2035 года. Министерство энергетики РФ [Электронный ресурс]. – Режим доступа: <http://minenergo.gov.ru/system/download-pdf/6365/66647>, свободный (дата обращения: 26.09.2017).
2. Daim T.U., Chiavetta D., Porter A.L., Saritas O., “Anticipating future innovation pathways through large data analysis”, In: Innovation, technology, and knowledge management, Springer, 2016, 360 p., DOI 10.1007/978-3-319-39056-7.