

Некоторые научно-методические вопросы создания Фонда аналитических технологий и программ

Аннотация. В статье обсуждаются вопросы интеграции усилий разработчиков информационно-аналитических систем в рамках национальной интеллектуальной аналитической платформы (АП). Предлагается создать Фонд аналитических алгоритмов и программ – информационную систему, предназначенную для сбора, систематизации и регистрации аналитических технологий и программ для электронных вычислительных машин, а также для обеспечения доступа к ним заинтересованных физических и юридических лиц. Критерии включения АП в Фонд будут учитывать научный уровень разработок, степень их реализации и внедрения, выполнение ряда требований к программному обеспечению и документации.

Ключевые слова: фонд аналитических алгоритмов и программ, национальная интеллектуальная аналитическая платформа.

Введение

Возрастание значимости информационных технологий (ИТ) практически во всех сферах деятельности общества сегодня неоспоримый факт. В нашей стране развитие информационных технологий определяется целым рядом основополагающих документов на период до 2030 г. [1-3].

В самом общем виде следует выделить три определяющих группы факторов:

- объективная потребность в обеспечении глобальной информационной доступности (Интернет, Интернет вещей, мобильные сети, социальные сети, информационные услуги широкого класса и др.);

- необходимость обеспечения не только информационной, но и национальной безопасности в условиях резко возросшего перечня угроз в информационной сфере;

- широкое применение информационных технологий в сфере производства (САПР, системы ERP, АСУ производством, АСУ предприятиями, математическое моделирование, мотивация высококвалифицированных кадров в области ИТ); на повестке дня – цифровая экономика.

Интенсивное развитие ИТ, с одной стороны, и огромные объемы накопленных данных в различных областях науки и производства, с другой, объективно внесли существенные изменения в подходы к организации научных исследований, приводящих к выдающимся научным результатам и открытиям. На смену традиционным подходам в научных исследованиях приходят методы с интенсивным использованием данных, объединяющие наблюдения (эксперимент), теорию и моделирование для извлечения знаний из больших объемов накопленных данных. Такой подход получил название четвертой парадигмы, в соответствии с которой проведение научных исследований, основанных на сборе, систематизации больших объемов данных, их анализе с целью извлечения знаний (Data intensive domains - DID) и становится неотъемлемой частью различных областей науки, экономики, бизнеса [4]. Полный набор данных о предметной области становится ее моделью, предназначенной для анализа и получения новых знаний.

Большие объемы данных накоплены в астрономии, физике высоких энергий, биоинформатике, мониторинге и моделировании климата в геонауках, в исследованиях на основе численного моделирования и др. В последнее вре-

мя часто различные области научного знания в науках с интенсивным использованием данных называют “Х-информатикой” (например, Биоинформатика, Геоинформатика, Астроинформатика, Нейроинформатика, Медицинская информатика, и др.) [5]. Заметим, что появляются и новые научные направления, в частности, *геоника (геомиметика)* – наука о законах развития объектов неорганического мира [6].

Совершенно очевидно, что новая парадигма научных исследований требует новых методов анализа данных с целью извлечения знаний и соответственно реализующих эти методы информационных технологий и программ. Естественно, имеет смысл вести речь о них как об *аналитических технологиях и программах*. Заметим, что именно на основе новых подходов к проведению научных исследований сегодня сформирована программа развития Федерального исследовательского центра «Информатика и управление» (ФИЦ ИУ РАН).

Учитывая принципиальную научно-практическую направленность деятельности ФИЦ ИУ РАН, на эффективную реализацию результатов научных исследований в практике создания информационно-аналитических систем преимущественно в сфере органов государственной власти, представляется крайне важной постановка задачи формирования и поддержания Фонда аналитических алгоритмов и программ. Кроме того, создание такого Фонда, безусловно, будет способствовать решению задач формирования национальной интеллектуальной аналитической платформы для интеграции существующих и разработки перспективных систем поддержки принятия решений в АСУ, создаваемых в интересах обороны, безопасности и охраны правопорядка.

Понятия «Фонд алгоритмов», «Фонд алгоритмов и программ» и другие, им подобные, не являются новыми. Так Постановлением Правительства от 30 января 2013г. №62 «О национальном фонде алгоритмов и программ для ЭВМ» создана федеральная государственная информационная система по сбору, обработке и хранению алгоритмов и программ для электронных вычислительных машин - Национальный фонд алгоритмов и программ для электронных вычислительных машин (НФАП).

Вместе с тем, как показывает анализ, создаются ведомственные фонды, в рамках которых конкретные заказчики пытаются накапливать

алгоритмы и программы, разработанные в ходе проводимых ОКР, с целью исключения, по возможности, дублирования разработок, а также обеспечения преемственности в ходе перманентно ведущихся работ по развитию информационных систем. Естественно, что такие фонды имеют закрытый характер. Неким антиподом таким фондам являются общероссийские фонды, доступные практически всем [7]. По содержанию в таких фондах собираются и хранятся компоненты информационного и лингвистического обеспечения (классификаторы, словари и т.п.), общего, общесистемного и прикладного программного обеспечения.

1. Цели и задачи Фонда аналитических технологий и программ

На сегодняшний день с учетом изложенных выше тенденций развития ИТ и кардинальных изменений в сфере научных исследований особую актуальность приобретают именно аналитические технологии и программы, реализующие новые методы извлечения необходимых знаний в заданной предметной области из огромных массивов накопленных данных. Излишне говорить об актуальности таких технологий в сфере военного управления.

Под Фондом аналитических технологий и программ (ФАТП) мы понимаем информационную систему, предназначенную для сбора, систематизации и регистрации аналитических технологий и программ для электронных вычислительных машин, а также для обеспечения доступа к ним заинтересованных физических и юридических лиц. Объектом учета Фонда является Аналитическая технология и программный продукт (АТиП), ее реализующий, с комплектом документации.

При этом под *аналитической технологией* мы понимаем регламентированный процесс получения информационного продукта:

- созданного на основе специально собранных и подготовленных достоверных данных предметной области анализа;
- содержащего количественные или качественные оценки анализируемых процессов, полученные с помощью определенных математических, эвристических, экспертных и других методов анализа, как правило, с использованием программного обеспечения;

- характеризующего скрытые процессы, закономерности, факторы и параметры анализируемой предметной области;
- содержательная интерпретация которого позволяет принять результативные управленческие решения.

Целями создания Фонда являются:

- формирование национальной интеллектуальной аналитической платформы в части АТиП;
- решение научно-технических задач систематизации разработки и применения созданных АТиП в различных сферах жизнедеятельности;
- упорядочение рынка АТиП с целью получения коммерческой выгоды учредителями и участниками на основе анализа спроса и формирования инновационных предложений, а также в интересах потенциальных пользователей.

Предполагается, что **Фонд будет решать следующие задачи:**

- систематизация аналитической проблематики различных предметных областей и выявление тематических сфер и направлений перспективных разработок;
- создание методической базы применения АТиП Фонда в различных отраслях народного хозяйства и в специальных применениях;
- унификацию АТиП (документация, интерфейсы, программное обеспечение) – создание принципов «хорошей практики» разработки и применения АТиП;
- пропаганду, популяризацию и продвижение АТиП (информационный портал, издание, выставки, конференции);
- разработку методик очного и дистанционного обучения и подготовки квалифицированных пользователей.

2. Методическое и информационное обеспечение деятельности ФАТП

Деятельность Фонда определяется нормативно-правовым и научно-методическим обеспечением (положение, типовые документы, методики и т.п.)

В составе Фонда предусмотрено создание следующих компонентов:

- Реестр АТиП:
- апробированные и перспективные аналитические технологии и реализующие их про-

граммные продукты, включая демо-версии, обучающие программы и рекламные ролики и т.д.;

- программная документация.
- Техническая база (узел Интернета, презентационный и тестовый стенд, средства хранения данных Фонда).
- Информационное обеспечение (классификаторы методов и задач, образцы документов).
- Интернет-ресурсы участников – организаций, собственников АТиП, с подробной информацией об АТиП.
- Организационная структура (руководитель Фонда, экспертные группы, группа эксплуатации).

Определены следующие **требования к АТиП** для включения в состав Фонда:

1. АТиП после настройки на конкретную предметную область с использованием апробированных методов анализа осуществляет решение значимой (комплексной) аналитической проблемы.

2. АТиП удовлетворяет требованиям Фонда к программному обеспечению, программной и технологической документации.

3. АТиП внедрен и промышленно эксплуатируется в ряде государственных и коммерческих организаций.

4. Собственник АТиП предоставил для рассмотрения в экспертную группу Фонда все необходимые материалы и продемонстрировал функциональные возможности регистрируемой версии.

5. Собственник АТиП заключил соглашение с учредителями Фонда и согласен разместить определенную информацию об АТиП на информационном портале Фонда, а также поддерживать ее в актуальном состоянии.

С целью выявления новых перспективных разработок в составе Фонда организуется раздел перспективных АТиП, в котором учитываются предложения по разработке и внедрению АТиП, основанных на новых методах анализа или применения апробированных методов анализа для решения новых аналитических задач. Данные предложения учитываются в Фонде на основе экспертной оценки.

Следует отметить, Фонд не будет являться частью Национального фонда НФАП, так как в него поступает интеллектуальная собственность РФ, а в Фонде в основном регистрируются АТиП, находящиеся в интеллектуальной собственности организаций разработчиков или

коллективов разработчиков. Также отличие Фонда от НФАП – в специализации, т.е. в наличии более развитых средств подбора аналитических технологий под конкретные аналитические задачи.

В рамках Фонда, несомненно, надо систематизировать аналитические методы и перечень аналитических задач для всех уровней управления (стратегический, тактический, базовый) [1]. Сюда относятся:

1. *Анализ фактов, формализация данных, извлечение знаний.*

1.1. Анализ текстов на естественном языке и слабоструктурированных данных с целью выделения из них контролируемых информационных объектов предметной области и их взаимосвязей.

1.2. Когнитивный анализ.

1.3. Методы искусственного интеллекта.

1.4. Автоматизированный учет, рубрицирование, реферирование и перевод документов.

2. *Анализ взаимосвязей на дискретной модели.*

2.1. Визуализация и навигация по семантической сети.

2.2. Поиск неочевидных взаимосвязей (путей на графе).

2.3. Поиск схожих пространственно-временных конфигураций событий (идентификация ситуаций).

3. *Анализ массовых потоков событий.*

3.1. Формирование выборок, рядов и OLAP-кубов.

3.2. Экспресс-анализ ситуаций.

3.3. Динамическое моделирование ситуаций.

3.4. Прогнозирование ситуаций, сценарное прогнозирование.

3.5. Анализ временных рядов.

3.6. Статистическая оценка характеристик потоков событий.

4. *Информационно-расчетные задачи.*

4.1. Картометрические ИРЗ (моделирование и прогнозирование).

4.2. Расчет тематических зон.

4.3. Расчет зон достижимости и оптимальных маршрутов.

4.4. Расчет вероятностей выполнения задач.

4.5. Логистические задачи.

5. *Моделирование и математические методы.*

5.1. Математическое моделирование.

5.2. Моделирование логических структур.

5.3. Моделирование процессов управления.

5.4. Адаптивные и обучающие системы.

5.5. Самоорганизующиеся системы.

5.6. Оптимизация надежности.

5.7. Методы оценки надежности.

5.8. Теория полезности и принятия решений.

5.9. Теория игр и ее применение в кибернетических системах.

5.10. Теория массового обслуживания и ее приложения.

5.11. Методы исследования операций.

5.12. Технические приложения вероятностно-статистических методов.

5.13. Программно-целевое планирование.

5.14. Теория планирования эксперимента.

Первоочередной задачей создания Фонда является разработка проекта Положения о ФАТП, включая следующие разделы:

- Общие положения.
- Критерии включения АТиП в Фонд.
 - Требования к изделиям: требование к программному обеспечению, требования к документации.
- Процедура включения АТиП в Фонд.
- Правовые отношения между разработчиком изделия и Учредителями фонда.
- Портал Фонда.
- Издательская и выставочная деятельность Фонда.
- Типовые документы Фонда.
 - Типовое соглашение.
 - Типовые методические рекомендации по направлениям применения изделия для создания СППР в АСУ военного и специального назначения.
 - Типовые форматы обмена аналитическими данными.
 - Типовые формы обучения.

Заключение

Несомненно, при включении АТиП в Фонд будут учитываться научный уровень разработок, степень их реализации и внедрения, выполнение ряда требований к программному обеспечению и документации. Важнейшую роль должен играть портал Фонда, который должен стать площадкой для взаимодействия потенциальных пользователей и разработчиков,

а также обеспечивать полную информационную поддержку процессов регистрации АТиП, ведения учетов математических методов анализа, перспективных направлений исследований.

Важнейшие вопросы интеграции аналитических данных, получаемых от различных АТиП при создании интегрированных систем поддержки принятия решений должны изучаться на основе существующего зарубежного и отечественного опыта. В [2, 3] описаны способы формализации аналитических данных с учетом всех составляющих контролируемого пространства распределенной системы управления, включая, целевую обстановку, контролируемые объекты и ресурсы, неконтролируемые факторы. При этом учитываются и пять стадий аналитической обработки данных: оценка параметров ненаблюдаемых (скрытых) элементов обстановки, оперативный анализ обстановки путем ее сравнения с прошедшим периодом, оценка ситуации с целью определения степени ее сложности – штатная, критическая, чрезвычайная, прогнозирование изменения обстановки, и непосредственно поддержка процессов принятия управленческих решений.

Считается, что Фонд будет играть важную роль во внедрении результатов передовых научных разработок в промышленный оборот, что будет способствовать решению актуальных

задач, стоящих перед экономикой и государством в целом.

Литература

1. Зацаринный А.А., Сучков А.П. Некоторые подходы к ситуационному анализу потоков событий // Открытое образование, 2012. № 1, С. 39–45.
2. Сучков А.П. Некоторые подходы к интеграции аналитических данных существующих и перспективных систем поддержки принятия решений // Системы и средства информатики, 2015, т. 25, № 3, С. 201–211.
3. Сучков А.П. Ситуационный подход и информационная модель предметной области в правоохранительной сфере // Методы построения и технологии функционирования ситуационных центров. – М.: ИПИ РАН, 2011. С.76–88.
4. Tansley S. et al. (ed.). The fourth paradigm: data-intensive scientific discovery. – Redmond, WA: Microsoft research. – Т. 1. 2009.
5. Вовченко А.Е., Калинин Л.А., Ковалев Д.Ю. Методы разрешения сущностей и слияния данных в ETL-процессе и их реализация в среде Nadoop // Информатика и ее применения. – Т. 8. – №. 4. – 2014. С. 94–109.
6. Лесовик В.С. Геоника (Геомиметика) как трансдисциплинарное направление исследований // Высшее образование в России. – №3. 2014. С. 77–83.
7. Постановление Правительства Российской Федерации от 30 января 2013г. № 62, г. Москва "О национальном фонде алгоритмов и программ для электронных вычислительных машин" // [Электронный ресурс] URL: <http://rg.ru/2013/02/04/elektrofond-site-dok.html> (дата обращения 04.04.2016).

Зацаринный Александр Алексеевич. Заместитель директора ФИЦ ИУ РАН. Окончил Киевское высшее военное инженерное училище связи в 1973 году, Академию Генерального штаба Вооруженных Сил РФ в 2000 году, Школу управления им. Дж.Ф. Кеннеди при Гарвардском университете (США) в 2003 году. Доктор технических наук, профессор. Автор 170 печатных работ. Область научных интересов: информационно-телекоммуникационные и управляющие сети, методы оценки эффективности информационных сетей, методы системного проектирования информационных сетей. E-mail: azatsarinny@ipiran.ru

Соченков Илья Владимирович. Научный сотрудник, заместитель заведующего лабораторией ИСА ФИЦ ИУ РАН, доцент кафедры информационных технологий РУДН. Окончил Российский университет дружбы народов в 2009 году. Кандидат физико-математических наук. Автор 50 печатных работ. Область научных интересов: интеллектуальные методы поиска и анализа информации, обработка больших массивов данных, контентная фильтрация, компьютерная лингвистика, распознавание образов. E-mail: sochenkov@isa.ru

Сучков Александр Павлович. Ведущий научный сотрудник ИПИ ФИЦ ИУ РАН. Окончил МГУ им. М.В. Ломоносова в 1976 году. Доктор физико-математических наук. Автор около 100 печатных работ. Область научных интересов: системный анализ, ситуационный анализ, системы поддержки принятия решений в системах управления. E-mail: asuchkov@ipiran.ru