

д.т.н. В.М. Лазарев

к.т.н. А.Е. Любимов.

**Предложения по использованию информационно-аналитических систем
в информационно-правовом обеспечении органов законодательной и
исполнительной власти федерального, регионального и местного уровней**

Аннотация: В статье описан реализованный подход к комплексной обработке неструктурированной – текстовой и аудиовизуальной – информации в целях поддержки принятия решения пользователей правовой информации.

Ключевые слова: неструктурированная информация, полнотекстовый поиск, анализ текста, обработка аудиовизуальной информации, моделирование на основе когнитивных карт, поддержка принятия решений

Введение

В соответствии с Концепцией информатизации Министерства юстиции Российской Федерации на Министерство юстиции Российской Федерации федеральными законами, указами Президента Российской Федерации, постановлениями Правительства Российской Федерации возложены задачи сбора, обработки, хранения, анализа правовой информации и организации ее использования. При этом основными информационными ресурсами, находящимися в распоряжении Министерства юстиции Российской Федерации и подведомственных ему организаций, являются компьютерные базы данных по различным аспектам права. Кроме того, Минюст России обладает уникальными информационными ресурсами на бумажных носителях - это фонды правовых актов НЦПИ, Минюста России, бывшего Минюста СССР, содержащие около 3,5 млн. документов. Важная роль в использовании этих информационных ресурсов принадлежит Научному центру правовой информации (НЦПИ) при Министерстве юстиции Российской Федерации

Созданная система распределенной обработки правовой информации позволяет обрабатывать в НЦПИ информацию, поступающую из федеральных органов государственной власти, а также органов государственной власти Москвы и Московской области. Правовая информация органов государственной власти субъектов Российской Федерации обрабатывается территориальными органами Минюста России или центрами правовой информатизации и передается в НЦПИ как в центральный узел системы по коммутируемым каналам связи, в том числе с использованием сети Интернет.

Дальнейшее развитие системы информатизации Минюста России требует создания высокоинтеллектуальных систем, выполняющих не только функции справочно-информационного обслуживания пользователей, но и обеспечивающих эффективную поддержку законотворческой деятельности.

Указанные системы должны обеспечивать законотворческую деятельность Минюста России на основе решения следующих задач:

обработки запросов, заданных в виде свободных наборов терминов или произвольных текстов на естественном языке;

выделения в полнотекстовой базе данных множества релевантных документов (или их фрагментов) с автоматическим ранжированием текстов по уровню релевантности;

динамического формирования текстов документов, актуальных на задаваемый пользователем период времени;

навигации в правовой базе данных не только по заранее установленным гиперссылкам, но и по любому другому признаку, выбираемому пользователем (например, по органам, источникам опубликования, позициям общеправового классификатора отраслей права и т.д.).

Информационное пространство, в котором функционируют организации, осуществляющие правовое обеспечение существенно изменилось и имеет следующие особенности:

– неструктурированные данные (тексты, изображения, видео, аудио и пр.) составляют большую часть накопленной информации: их доля может составлять более 95% накопленных данных;

– увеличивается разнообразие и количество технических средств получения, преобразования и представления информации (телевидение, радио, интернет, электронная почта, видео конференцсвязь, сотовая связь, видеозапись и пр.);

– увеличивается количество источников информации (газеты, журналы, телевизионные и радиоканалы, интернет-сайты, форумы, блоги и т.п.);

– скорость распространения неструктурированной информации непрерывно растет;

– современные международные связи и деловое общение, как никогда ранее, предполагают многоязыковость.

Все это приводит к резкому возрастанию роли информационно-аналитических систем при возрастании требований к ним, а именно: сокращаются сроки представления данных эксперту, аналитику; увеличивается количество факторов учитываемых системой; все более важное значение приобретает прогнозирование развития событий в интересах государственных органов.

В связи с этим крайне актуальной является проблема автоматизации процессов обработки неструктурированной информации в информационно-аналитических системах. В настоящей статье описан реализованный подход к комплексной обработке неструктурированной информации в целях поддержки принятия решений.

Предлагается в качестве одного из путей решения этой проблемы рассматривать разработку и внедрение информационно-аналитических систем, позволяющих круглосуточно обрабатывать огромные разнородные информационные потоки путем создания единого информационного пространства и применения интеллектуальных технологий работы с информацией.

Неструктурированная информация: проблемы обработки

Как показали исследования, проведенные ведущими независимыми аналитическими агентствами, объем всей накопленной в мире информации

превысил 281 экзабайт (или 281 млрд. гигабайт)/1/. Особого внимания заслуживает тот факт, что более 95 % всей созданной и накопленной в мире информации носит неструктурированный характер, т.е. имеет форму текстов произвольного содержания и аудиовизуальную форму. Неструктурированная информация включает в себя: телефонные разговоры, голосовую почту, электронную почту, текстовые документы, бумажные документы, изображения, веб-страницы, видео и сотни других различных видов и форматов файлов и данных, причем доля информации, представленной в аудиовизуальной форме, стремительно растет.

К сожалению, в настоящее время попытки выгодно использовать этот огромный и стратегический ресурс, часто оказывались безуспешными, т.к. у многих организаций, как частных, так и государственных, нет необходимой технологии, чтобы понять и эффективно работать с информацией, которая находится вне упорядоченных баз данных, и тем более, с информацией в аудиовизуальной форме. При этом работа со структурированной информацией остается также крайне важной в деятельности любой организации.

Все это требует наличия информационной инфраструктурной платформы, которая может работать со всеми типами информации, методом, принятым для структурированных баз данных. Аналогично системе управления реляционными базами данных, которая явилась революцией в компьютерном мире в 1960х годах, эта инновационная платформа должна давать компьютерам возможность обрабатывать не только структурированные данные, но также громадные количества частично структурированной и неструктурированной информации, используя реляционный индекс.

Анализ положений об органах государственной власти (ОГВ) показал, что, как правило, последним, на законодательном уровне, вменяется в обязанность осуществлять мониторинг информационного пространства в сфере своих компетенций, разрабатывать предложения по улучшению тех или иных ситуаций в сфере своих компетенций. Кроме того, специализированные службы органов государственной власти ежедневно занимаются информационным обеспечением рутинной и интеллектуальной деятельности персонала ОГВ.

Анализ официальных документов (Концепция формирования в РФ электронного правительства до 2010 года, Концепция использования информационных технологий в деятельности федеральных органов государственной власти до 2010 года, Концепция формирования и развития единого информационного пространства России и соответствующих государственных информационных ресурсов и др.) показал, что существует ряд общих проблем, значительно усложняющих внедрение информационных систем в органах государственной власти. К таким проблемам относятся:

- внутриведомственный характер информационно-коммуникационных технологий в органах государственной власти, препятствующий эффективному межведомственному взаимодействию и повышению качества государственных услуг, предоставляемых гражданам;

- нестандартизированность действующих государственных информационных систем и их недоступность другим органам государственной власти для оперативного использования

- противоречивость данных, содержащихся в государственных информационных системах;

- использование недокументированных форматов данных, протоколов обмена, иных закрытых информационных технологий и отсутствие единых классификаторов, справочников и схем данных, что ограничивает возможность применения автоматизированных средств поиска и аналитической обработки информации, содержащейся в различных системах, и затрудняет доступ организаций и граждан к государственным информационным системам;

- неавтоматизированность процедур сбора и обработки информации, необходимой для определения и контроля целевых показателей результативности деятельности органов государственной власти.

Кроме того, существуют и объективные тенденции и проблемы, характерные для внешней среды любой современной организации, которые оказывают существенное влияние на то, как организации работают с информацией. Среди них можно отметить тенденцию к приданию электронной информации определенного правового статуса, а также значительное снижение времени, отводимого на подготовку и принятие решений.

Таким образом, можно в общем виде сформулировать основные проблемы, с которыми сталкиваются органы государственной власти при обработке информации:

- необходимость работать с информацией разного типа (структурированная и неструктурированная);

- разнородность источников и форматов данных (телевидение, радио, печатные издания, интернет, базы данных и пр.);

- большие объемы данных;

- необходимость гибко и оперативно настраивать систему на различные задачи в соответствии с меняющейся обстановкой;

- необходимость синхронизации мощностей системы с нарастающими потоками данных (масштабируемости);

- необходимость эффективно анализировать данные в распределенной среде;

- необходимость прогнозирования развития ситуаций, например, по модели «что, если...»;

- необходимость мониторинга открытого информационного пространства (интернет/СМИ);

- необходимость применять максимально стандартизованные решения.

Описание предлагаемого подхода

Решение перечисленных проблем требует использования современных аналитических систем. В статье представлено описание одной из таких систем. Информационно-аналитическая система (ИАС) "Лавина" предназначена для сбора, обработки и консолидации разнородной неструктурированной информации – текстовой и аудиовизуальной – из внутренних и внешних источников (базы данных, интернет, файловые системы, корпоративные информационные системы, телевизионный и радио эфир и др.) и ее автоматической аналитической обработки в режиме, близком к реальному времени.

Применение ИАС "Лавина" позволяет:

- оперативно отслеживать появление новых информационных поводов;
- осуществлять непрерывный мониторинг и анализ развития различных ситуаций;
- оценивать "информационный портрет" персоны, организации, бренда и пр. в СМИ и социальных медиа;
- прогнозировать развитие ситуаций.

Принципиальной особенностью системы "Лавина" является то, что, во-первых, система способна обрабатывать как текстовую, так и аудиовизуальную информацию и, во-вторых, система поддерживает полный цикл обработки данных, т.е. преобразование данных в информацию и извлечение знаний из информации посредством проведения анализа текста и ситуационного моделирования.

В качестве источников данных для ИАС могут выступать любые известные источники информации (интернет, файловые системы, базы данных, аудио- и видеоканалы). ИАС способна работать как с потоком, так и с файлами практически всех известных форматов. Укрупненная схема работы системы представлена на рис. 1.

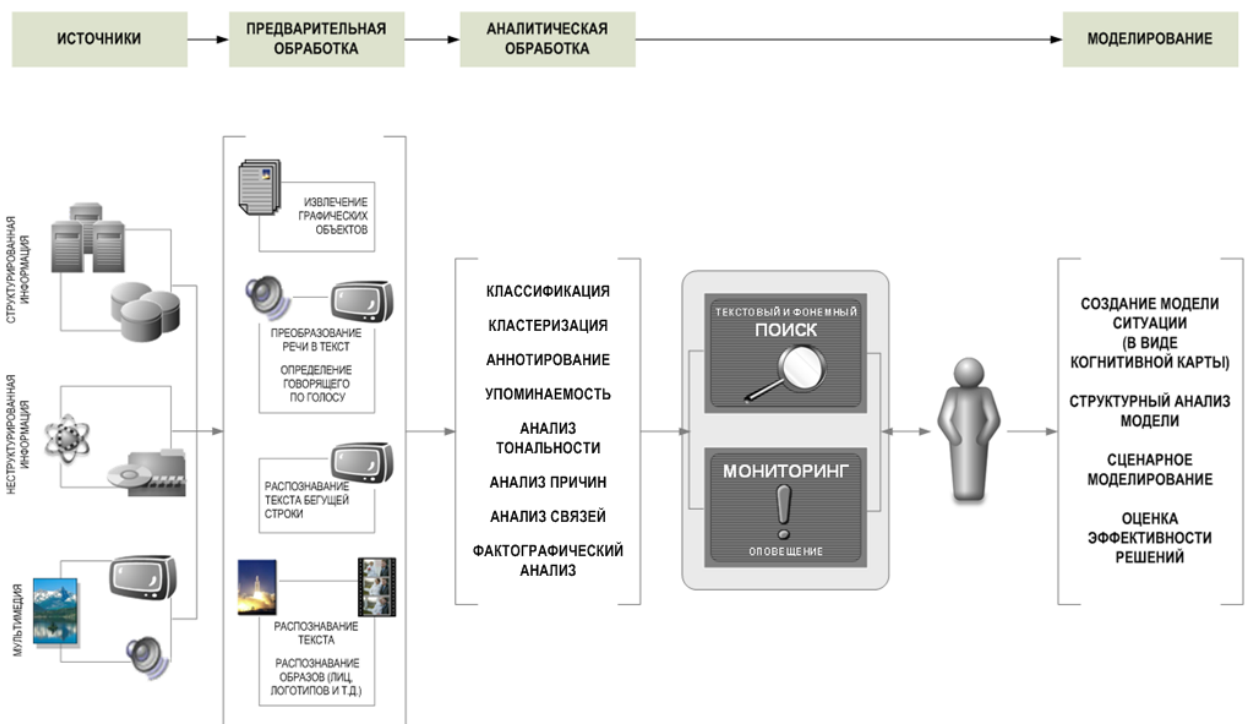


Рис.1. Укрупненная схема работы системы

Каждый входящий файл подвергается предварительной обработке. Речь, содержащаяся в аудиофайлах и звуковых дорожках видеофайлов, преобразовывается в текст, и определяется принадлежность голоса говорящего (диктора). Бегущая строка и титры в видеофайлах преобразовываются в текст. В графических файлах, извлеченных из текста, а также кадрах, на которые разбиваются видеофайлы, производится поиск и распознавание печатного текста и образов (логотипов, силуэтов и т.п.) и идентификация лиц. Такая информация, так же как и текстовое содержание документов, индексируется и становится доступной для поиска, мониторинга и дальнейшей аналитической обработки. Результаты

обработки текстов на примере анализа деятельности госкорпораций представлены на рис.2.

Аналитическая обработка неструктурированной информации заключается в ее

- классификации;
- кластеризации;
- аннотировании;
- формировании рядов данных по
 - упоминаемости,
 - основным тенденциям,
 - оценке критики,
 - негативного и позитивного окраса тех или иных объектов, событий
 - и пр.

На основе производимого синтаксического и семантического анализа текстов возможно решения ряда прикладных задач, таких, как: контент-анализ, ивент-анализ, поддержка ситуационного моделирования и пр.

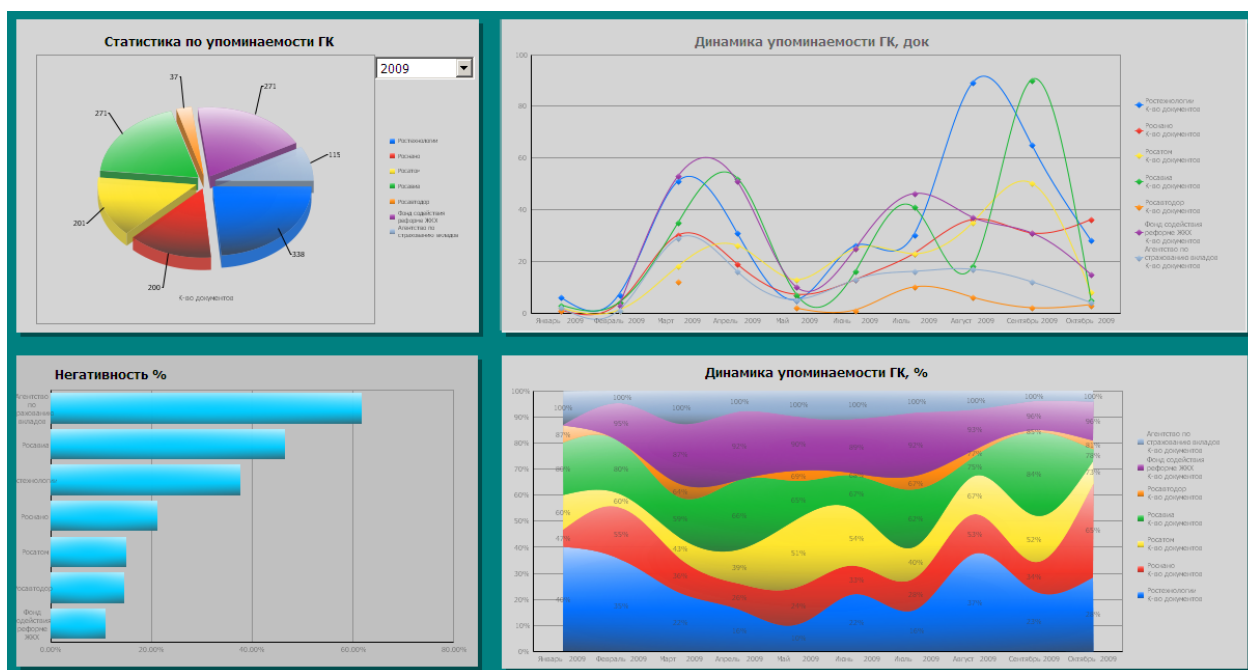


Рис.2. Результаты аналитической обработки текстов.
На примере анализа деятельности госкорпораций

Когнитивное моделирование

При решении задач управления слабоструктурированными объектами (СО) и ситуациями, часто возникают трудности по структурированию и переработке информации, необходимой для подготовки и принятия управленческих решений.

Одной из причин является недостаток информации о состоянии СО в условиях изменяющейся и слабо контролируемой внешней среды. Отсутствие

достаточных знаний о системе, относительно которой принимается решение, не является единственной неопределенностью, обусловленной субъективными причинами. Также можно выделить неопределенность целей развития СО и критериев выбора управленческого решения /2/. Как правило, неудовлетворенность текущим состоянием системы осознается субъектом управления, но его представления о причинах и возможных способах изменения ситуации в СО размыты, нечетки и противоречивы. Формализация нечетких представлений – одна из главных задач, которую необходимо решить при разработке моделей и методов поддержки принятия решений в слабоструктурированных ситуациях /3/.

Другая трудность связана с тем, что субъекту управления приходится манипулировать качественной информацией в виде гипотез (предположений), интуитивных понятий и смысловых образов. Многочисленные исследования процессов принятия решений подтверждают, что субъекту управления несвойственно мыслить и принимать решения только в количественных характеристиках. Он мыслит, прежде всего, качественно, и для него поиск решения – это, поиск, в первую очередь, замысла решения, где количественные оценки играют вспомогательную роль /4/. Поэтому “структуры знания в мышлении субъекта (лица, принимающего решения), оказываются важнейшими элементами ситуации, неустранимыми из модели принятия решений”.

В практике подготовки и принятия решений класс управленческих задач, связанных с выявлением и диагностированием слабоструктурированных проблем и постановки целей развития СО, слабо поддержан формализованными методами, которые ориентированы на преодоление перечисленных выше трудностей.

Для решения этих задач на практике, как правило, применяются эвристические экспертные методы (мозговой штурм, интервьюирование и т.п.). При этом для структуризации проблемы нередко используется иерархическое представление – в виде дерева. Традиционные теоретические методы концентрируют внимание на процессах выбора альтернативы из фиксированного набора решений, но чаще всего множество альтернатив не может быть представлено эксперту в полном объеме. Поэтому важен не только процесс поиска рационального решения, но и собственно процесс формирования его допустимых вариантов /5/.

В качестве формализованного метода, ориентированного на работу с неоднородной информацией (качественной и количественной), с экспертными знаниями, когда исследуемая система, ситуация, проблема не могут быть описаны субъектом управления точно, сегодня все шире применяется когнитивный подход, который позволяет поддержать ранние этапы процесса подготовки и принятия решений – этапы выявления и диагностирования слабоструктурированных проблем и постановки целей развития СО/6/.

Типичным классом СО, для решения проблем развития которых целесообразно применение когнитивного моделирования, является социально-экономические системы.

Основные понятия в современном когнитивном моделировании

Ключевые понятия, сложившиеся и широко используемые в рамках когнитивного подхода и его различных школ, в публикациях зачастую не определяются; при этом нередко возникает неоднозначность понимания вплоть до противоречий как из-за различий понимания в разных школах, так и, в некоторых случаях, в рамках одной школы. Авторами статьи предпринята попытка уточнить некоторые основные понятия когнитивного подхода к решению задач анализа и управления СО.

Когнитивное моделирование – это исследование функционирования и развития слабоструктурированных систем и ситуаций *посредством построения когнитивной карты СО.*

Когнитивная карта /7/ отражает субъективные представления (индивидуальные или коллективные) исследуемой проблемы, ситуации, связанной с функционированием и развитием СО. Основными элементами когнитивной карты являются базисные факторы/8/ (или просто факторы) и причинно-следственные связи между ними.

Содержательно, *базисные факторы* – это факторы, которые (1) определяют и ограничивают наблюдаемые явления и процессы в СО и окружающей его среде и (2) интерпретированы субъектом управления как существенные, ключевые параметры, признаки этих явлений и процессов.

Метод формирования стратегических проблем развития социально-экономической системы

В общем виде *управление развитием социально-экономической системы (СЭС)* можно представлять как построение стратегии развития системы, которая определяет основные цели и общие направления их достижения, и ее реализацию.

Одним из ключевых этапов построения стратегии управления СЭС является выявление проблем развития системы, оказывающих негативное влияние на достижение стратегических целей управления.

В данном разделе кратко представлен метод формирования стратегии решения слабоструктурированных проблем на основе линейных динамических когнитивных моделей применительно к СЭС.

Задача управления состоит в переводе СЭС в одно из состояний, соответствующих *целевому образу* системы. Динамика изменения факторов когнитивной модели, характеризующих состояние системы в некоторый момент времени, определяется моделью вида (1).

Целевой образ СЭС задает, что необходимо изменить с позиции субъекта управления, и формально представляется как

$$C = (X^C, R(X^C)), \quad (2)$$

где X^C – подмножество целевых факторов, $X^C \subseteq X$ (X – множество факторов когнитивной модели);

$R(X^C)$ – вектор оценок динамики факторов (ОДФ), определяющий желательные направления изменения целевых факторов

$$R(x_i^C) = \begin{cases} +1, & \text{если желательно увеличение значения фактора } x_i^C \\ -1, & \text{если желательно уменьшение значения фактора } x_i^C \end{cases}$$

В рамках предлагаемого метода рассматривается коллективный субъект управления – руководство СЭС, принимающее управленческие решения, и его окружение (системные аналитики и эксперты), осуществляющие подготовку решений. На стратегию управления развитием СЭС могут влиять и другие субъекты, которые имеют собственные цели, интересы, представления и установки, определяющие выбор решений по развитию СЭС. В этом случае ту же задачу управления нужно решать с позиции каждого выделенного субъекта.

Стратегия решения проблем развития СЭС состоит из w стратегических шагов, которые задают последовательность изменений состояний системы

$$S^0 \rightarrow S^1 \rightarrow S^2 \dots \rightarrow S^m \rightarrow S^C,$$

где S^0 – исходное состояние,

S^C – целевое состояние (соответствующее целевому образу (2)),

$S^i \rightarrow S^{i+1}$ – стратегический шаг, на котором выявляется проблема и на основе ее анализа на множестве факторов X когнитивной модели выделяется подмножество локальных целей (целевых факторов) и подмножество управлений (управляющих факторов), изменение которых приводит к желательному изменению целевых факторов.

Каждый стратегический шаг $S^i \rightarrow S^{i+1}$ включает

- выявление проблемы на базе моделирования саморазвития начального состояния i -го шага, в результате которого проблема уточняется в виде P^i – подмножества факторов, изменение которых не соответствует целевому образу;
- диагностирование проблемы путем построения “подграфа причин” и структурно-целевого анализа с целью выделения из P^i подмножества локальных непротиворечивых целевых факторов Y^i и поиска вариантов управлений (подмножеств управляющих факторов) U_j^i , способствующих изменению Y^i в желательном направлении. Непротиворечивость целевых факторов означает, что желательное изменение любого целевого фактора из Y^i не приводит к нежелательному изменению остальных факторов из Y^i ;
- моделирование управляемого развития системы, на основании которого формируются различные сценарии.

Заключение

В заключении нам представляется целесообразным сформулировать направления дальнейших исследований. Опыт применения различных моделей и методов на базе когнитивного подхода (в России и за рубежом), повышающийся интерес управленцев–практиков к разработкам в данном направлении показывают

целесообразность развития данного подхода в управлении. При этом следует отметить наличие нерешенных (или отчасти решенных) проблем. Выделим некоторые направления исследований, которыми авторы статьи будут заниматься в рамках дальнейшего развития когнитивного подхода в моделировании и управлении.

Разработка научно-методического обеспечения. В этом направлении проводятся работы, связанные с разработкой методик выявления типичных рисков из-за человеческого фактора при формировании и формализации знаний о развитии слабоструктурированной системы, ее проблемах, целях развития и стратегиях разрешения проблем, включая формирование системы критериев, ориентированной на достоверность формализации первичных знаний (представлений); разработку принципов и методов, направленных на повышение уровня надежности и точности измерений при построении когнитивных карт и интерпретации результатов.

Разрабатывается подход к формализации первичных представлений о слабоструктурированной проблеме в виде коллективной когнитивной карты с целью обобщения и согласования разных представлений у носителей проблемы, компетентных в различных предметных областях знаний. Решение этой задачи опирается на разработанные методы концептуальной структуризации и критерии и частные технологии формирования и согласования коллективных понятий.

Планируется проведение цикла работ по интеграции когнитивного подхода и методов теории активных систем, поскольку в обоих научных направлениях большое значение отводится исследованию сложных систем, в которых одним из основных элементов являются активные субъекты, существенно влияющие на эффективность управления системой.

Повышение прикладной значимости результатов исследований. Продолжается работа по созданию технологии когнитивного моделирования слабоструктурированных ситуаций и систем с развитым научно-методическим и инструментальным обеспечением для внедрения в практику управленческой деятельности по планированию и управлению развитием социально-экономических систем.

В настоящее время разработан программно-аналитический комплекс, в котором реализованы функции построения когнитивных моделей, структурно-целевого анализа, сценарного моделирования и сравнительной оценки сценариев.

Модульная архитектура разработанного комплекса позволяет наращивать его другими инструментальными средствами решения различных задач управления, а также взаимодействовать с современными информационно-аналитическими системами (например, системами сбора и анализа информации, ERP – системами).

Литература

1. The Expanding Digital Universe. A Forecast of Worldwide Information Growth Through 2010 / D. Reinsel, Ch. Chute, W. Schlichting, et al. – Framingham, MA: IDC, 2007. – 24 p.; The Diverse and Exploding Digital Universe. An Updated Forecast of Worldwide Information Growth Through 2011 / Ch. Chute, A. Manfrediz, S. Minton, et al. – Framingham, MA: IDC, 2008. – 16 p.
2. Диев В.С. Нечеткость в принятии решений // Философия науки. – 1998. – № 1(4). – С. 45–52.;
3. Трахтенгерц Э.А. Субъективность в компьютерной поддержке решений. – М.: СИНТЕГ, 2001. – 256 с.
4. Сергеев В.М. Когнитивные методы в социальных исследованиях // Язык и моделирование социального взаимодействия. – г. Благовещенск: БГК им. И.А. Бодуэна де Куртенэ, 1998. – С. 3–19.
5. Компьютерная поддержка сложных организационно-технических систем / В.В. Борисов, И.А. Бычков, А.В. Дементьев и др. – М.: Горячая линия – Телеком, 2002. – 154 с.;
6. Дёрнер Д. Логика неудачи. Стратегическое мышление в сложных ситуациях. – М.: Смысл, 1997. – 243 с.;
7. Федулов А.С. Нечеткие реляционные когнитивные карты // Теория и системы управления. – 2005. – №1. – С. 120–132.
8. Лазарев В.М., Свиридов А.П. НЕЙРОСЕТИ И НЕЙРОКОМПЬЮТЕРЫ. Монография.- М.: 2011.- 131 с. МИРЭА, 2011.