

На правах рукописи



Переведенцев Денис Алексеевич

**РАЗРАБОТКА МЕТОДИКИ И АЛГОРИТМОВ ПОДДЕРЖКИ
ПРИНЯТИЯ РЕШЕНИЙ ПРИ РЕАЛИЗАЦИИ НАУКОЕМКИХ
ПРОЕКТОВ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ**

Специальность 05.13.01 – Системный анализ, управление и обработка информации
(в науке и технике)

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Ижевск – 2019

Работа выполнена в ФГБОУ ВО «Ижевский государственный технический университет имени М.Т. Калашникова», г. Ижевск

Научный руководитель: кандидат технических наук, доцент,
Благодатский Григорий Александрович,
ФГБОУ ВО «Ижевский государственный
технический университет имени М.Т.
Калашникова»

Официальные оппоненты: доктор физико-математических наук, профессор,
Гитман Михаил Борисович
ФГБОУ ВО «Пермский национальный
исследовательский политехнический университет»,
профессор кафедры вычислительной математики,
механики и биомеханики;

кандидат технических наук, старший научный
сотрудник

Милич Владимир Николаевич
ФГБУН «Удмуртский федеральный
исследовательский центр Уральского отделения
Российской академии наук», старший научный
сотрудник с возложением обязанностей
заведующего отделом исследования и диагностики
пространственных структур.

Ведущая организация: ФГАОУ ВО «Национальный исследовательский
Томский государственный университет».

Защита состоится «20» июня 2019 года в 11 часов на заседании
диссертационного совета Д212.065.06 в Ижевском государственном техническом
университете имени М.Т. Калашникова по адресу 426033, г. Ижевск, ул. 30 лет
Победы, 2, 5 корпус ИжГТУ имени М.Т. Калашникова.

С диссертацией и авторефератом диссертации можно ознакомиться в
библиотеке и на сайте ФГБОУ ВО «Ижевский государственный технический
университет имени М.Т. Калашникова», [http: //istu.ru](http://istu.ru).

Автореферат разослан «__» _____ 2019 г.

Отзывы на автореферат в двух экземплярах, заверенных гербовой
печатью, просим направлять на имя ученого секретаря диссертационного
совета по адресу: 426069, г. Ижевск, Студенческая, д.7, ИжГТУ имени М.Т.
Калашникова.

Ученый секретарь
диссертационного совета Д 212.065.06
кандидат технических наук, доцент



В.Н. Сяктерев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Поддержка принятия решений при разработке и реализации наукоемких проектов сложных технических систем характеризуется высокой сложностью информационной и логической декомпозиции, представления подсистем и отдельных элементов, а также наличием больших объемов данных и интенсивностью информационных потоков. В связи с этим сегодня предъявляются высокие требования к скорости сбора, качеству обработки данных и процессам получения оперативной информации, поскольку от этого зависит эффективность реализации проектов сложных технических систем и внедрения перспективных научных разработок.

Часто этапу реализации проекта предшествует не менее трудоемкая процедура сравнения множества его альтернатив или вариантов реализации, число которых в отдельных случаях может быть десятки и сотни, поэтому необходим учет множества разнородных характеристик и комплексный подход к оценке перспективности и реализуемости наукоемких проектов, предполагая, таким образом, решение задачи многокритериального выбора.

В заданных условиях существующие методические рекомендации и результаты научных трудов отечественных и зарубежных ученых становятся практически неприменимыми к данной предметной области. Требуется детальная разработка теоретических и практических положений по повышению эффективности анализа и оценки перспективности наукоемких проектов на основе специальных информационных технологий. В свою очередь создание проблемно-ориентированной информационной системы требует разработки концептуальной модели предметной области. Сегодня наиболее эффективным решением данной задачи является системный подход, являясь специальной научно-методологической концепцией исследования сложных систем, данный подход позволяет посредством системного анализа и синтеза выявить и описать функциональные связи и характеристики элементов, гармонично объединяемых в более крупные объекты.

Главной задачей разработки систем поддержки принятия решений в рамках научной информационно-технической системы является формирование специализированных баз данных, разработка аналитического и программно-технического инструментария с учетом предметных областей их использования и современных методов обработки информации, способствуя повышению эффективности управления проектами разработки сложных технических систем. В этой связи многомерный анализ данных и алгоритмы машинного обучения являются необходимыми аналитическими компонентами автоматизированной компьютерной системы поддержки наукоемких проектов.

Степень разработанности темы исследования. Фундаментальный вклад в становление методологической основы разработки и сопровождения программных автоматизированных комплексов с использованием онтологий, предназначенных для решения задач моделирования сложных систем, внесли выдающиеся отечественные и зарубежные ученые: Охтилев М.Ю., Соколов

Б.В. и Юсупов Р.М., Соколов Б.В., Поспелов Д.А., Грубер Т.Р., Гуарино Н., Ниренбург С. и Раскин В., Грозоф В., Штааб С. и Штудер Р., развившие базовые элементы таких научных направлений, как онтологии, базы знаний, теория автоматов, теория алгоритмов, теория искусственного интеллекта, математическая логика, теория программирования (Computer Science), теория статистических выводов и решений, теория формальных языков и др.

Применение системного подхода к поддержке принятия решений, в том числе в условиях нечеткой информации в общей теории систем, посвящены работы известных отечественных исследователей, таких как Петров Б.Н., Моисеев Н.Н., Флеров Ю.А., Емельянов С.В., Мелихов А.Н., Ульянов С.В..

Проблемам разработки систем управления проектами в различных областях деятельности посвящены работы Ермолаева Е.Е., Зарницыной К.В., Чжан Юйхуа, Багрий А.Н., Мочалова А.В., Соловьева Д.А., Стрельцина Я.С..

Разработкой подходов и методов к формированию критериев и многокритериальной оценке проектов также занимались многие ученые, среди них: Кирина Л.В., Азизов Ш.М., Шульпин А.Б., Юрковская Г.И., Репецкая Н.В., Жарков И. С., Багрий А.Н., Пащенко Д.С., Гильманова Р.И., Литке М.Г..

Если решение вопросов разработки и внедрения информационных процедур и баз данных в научно-технической сфере получили в научной литературе достаточно полное освещение, то вопросы разработки систем поддержки принятия решений и интеграции интеллектуальных компьютерных технологий в проектную деятельность наукоориентированных предприятий исследованы недостаточно.

Область исследования. Содержание диссертации соответствует следующим пунктам паспорта научной специальности 05.13.01 «Системный анализ, управление и обработка информации (по отраслям)» ВАК РФ: «4. Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений и обработки информации», «6. Методы идентификации систем управления на основе ретроспективной, текущей и экспертной информации», «9. Разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических объектов», «10. Методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических системах».

Объектом исследования являются данные наукоемких проектов разработки сложных технических систем.

Предмет исследования: принципы, методы и алгоритмы обработки информации и принятия управленческих решений в процессе реализации наукоемких проектов разработки сложных технических систем.

Цель диссертационного исследования заключается в разработке методики и алгоритмов, способствующих повышению обоснованности и эффективности принимаемых решений при реализации наукоемких проектов создания сложных технических систем на основе технологий многомерного и интеллектуального анализа данных.

Для достижения поставленной цели необходимо последовательное

решение следующих **задач**:

1. Разработать методику формализации знаний о системной структуре и процессах развития наукоемкого проекта разработки сложной технической системы.

2. Получить знание о системных связях элементов наукоемкого проекта по разработке сложной технической системы и разработать методику многопараметрической оценки его состояния.

3. Провести анализ условий принятия решений при реализации наукоемких проектов по разработке сложных технических систем и формализовать данные условия в форме продукционных моделей специальной базы знаний.

4. Используя разработанные модели и алгоритмы, осуществить программную реализацию совокупности аналитических и интеллектуальных модулей в составе информационно-аналитической системы.

Диссертация содержит следующую **научную новизну**:

1) разработана методика формализации системной структуры наукоемкого проекта в слабоструктурированных предметных областях путем построения концептуально-содержательной модели онтологии на основе профессиональных знаний и опыта экспертов;

2) впервые разработаны методика и математические алгоритмы формализации функциональных связей наукоемких проектов, которые позволяют получить интегральную оценку проекта, построить многомерную модель его данных, а также проводить многокритериальный отбор наукоемких проектов по разработке сложных технических систем;

3) разработана методика проектирования структуры и алгоритмов базы знаний для автоматического формирования рекомендаций по работе с наукоемкими проектами, основанная на формализации знаний эксперта предметной области, а также алгоритм автоматизированного обучения и оптимизации правил этой базы с использованием метода машинного обучения дерева решений (СART);

4) представлена методика проектирования программной архитектуры систем поддержки принятия решений, позволяющая как в рамках монолитных приложений, так и онлайн-платформ реального времени, объединить в единое информационное пространство подсистемы сбора и обработки данных проектов сложных технических систем, инструменты многомерного анализа и модуль экспертных рекомендаций.

Теоретическая значимость диссертационного исследования заключается: в развитии подходов системного анализа к решению информационно-технических задач учета и оценки влияния большого количества параметров наукоемкого проекта на успешность его реализации; в расширении информационно-технологической базы, предназначенной для поддержки принятия решений при планировании и реализации наукоемких проектов; в разработке и применении новых моделей и алгоритмов оценки научного потенциала и реализуемости наукоемких проектов разработки

сложных технических систем, а также автоматизации процедур принятия решений.

Практическая полезность работы. Теоретические выводы и результаты исследования могут использоваться на практике в виде системы поддержки принятия решений, позволяющей извлекать знания из больших массивов накапливаемой в базах данных слабоструктурированной информации в целях поиска путей эффективной реализации, продвижения и внедрения научных результатов и разработок.

Даны практические рекомендации по использованию разработанной экспертной системы, предложена и описана последовательность и содержание этапов подготовки и анализа информации. На основании теоретических расчетов и экспериментальной проверки предложенных алгоритмов показана эффективность применения технологий многомерного анализа и машинного обучения для мониторинга и анализа наукоемких проектов.

Изложенные в диссертации выводы и рекомендации носят практический характер и могут быть применены в проектной деятельности промышленных предприятий и научных учреждений.

Теоретико-методологическая основа исследования. Диссертационная работа основана на фундаментальных и прикладных исследованиях отечественных и зарубежных авторов в следующих областях: управление проектами создания технических системах, повышение эффективности методов ведения научных разработок и оценки их результатов, организация и управление проектной деятельностью, развитие инновационной деятельности предприятий, а также разработка информационных и аналитических систем поддержки принятия решений.

Информационно-эмпирическую базу исследования составили сведения и данные из публикаций отечественных и зарубежных исследователей и специалистов, материалы научно-практических конференций по проблемам управления процессами ведения научных разработок и создания систем поддержки принятия решений в научно-технической деятельности предприятий и государственных учреждений, материалы сети Internet, а также результаты экспертных оценок.

Методы исследования. Задачи, поставленные в рамках исследования, решались на основе: общей теории систем, методов математического моделирования, методов анализа данных и машинного обучения, нечетких множеств, теории принятия решений, методов объектно-ориентированного анализа и программирования, теории баз данных.

Положения, выносимые на защиту:

1. Методика разработки модели онтологии наукоемких проектов на основе концептуально-содержательного подхода к моделированию экспертных знаний о предметной области проекта, позволяющая формализовать его основные атрибуты и системные связи. Сформирована модель данных сущности «наукоемкий проект», а также описаны процессы реализации наукоемких проектов в рамках научной информационно-технической системы.

(Соответствует п. 6 паспорта специальности).

2. Методика и математические алгоритмы, отражающие функциональные зависимости параметров наукоемкого проекта. На основе системы нечетного логического вывода разработана комплексная методика оценки наукоемких проектов, представляющая собой оригинальный набор агрегированных критериев и учитывающая коэффициенты влияния отдельных параметров на интегральный показатель степени привлекательности проекта. Описан алгоритм решения задачи многокритериальной оптимизации отбора наукоемких проектов. Построена многомерная модель данных (MD-модель) наукоемкого проекта, на основе MD-модели разработана логическая структура OLAP-кубов для представления и анализа данных. *(Соответствует п. 4 паспорта специальности).*

3. Методика разработки самообучаемой и адаптивной базы знаний для формирования автоматических рекомендаций по работе с наукоемкими проектами разработки сложных технических систем, описывающая условия и порядок принятия управленческого решения. Данная база знаний позволяет оптимизировать и автоматизировать процессы работы с наукоемкими проектами за счет определения принадлежности текущей ситуации к одной из эталонных и формирования промежуточных выводов и итоговых рекомендаций по принятию управленческого решения. Расширение и оптимизация правил БЗ происходит на основе алгоритма машинного обучения дерева решений (СART). *(Соответствует п. 10 паспорта специальности).*

4. Методика проектирования программной архитектуры и компьютерной реализации разработанных моделей в виде алгоритмического обеспечения системы анализа и обработки информации, включающей серверную часть для ведения базы данных и их обработки, а также системы визуализации информации. Данная ИАС в совокупности с модулем формирования экспертных рекомендаций позволяет повысить обоснованность решений, принимаемых при реализации наукоемких проектов сложных технических систем. *(Соответствует п. 9 паспорта специальности).*

Достоверность и обоснованность теоретических выводов и практических результатов, полученных в работе, подтверждается корректным использованием методов исследования, результатами апробирования разработанных алгоритмов на общедоступных данных, публикацией научных трудов, а также сравнительным анализом результатов работы с известными результатами современных исследований и разработок.

Публикация результатов. Основные научные результаты по теме диссертации опубликованы в 14 научных работах [1-14] в журналах, сборниках научных трудов и материалов всероссийских и международных конференций, монографии [1], в том числе 5 научных трудов в изданиях, рекомендуемых ВАК для публикации основных результатов диссертаций [2-6]. Получены свидетельства на регистрацию базы данных №2016620179 от 08.02.2016 г. [7] (Приложение Г) и программы для ЭВМ №2016615251 от 19.05.2016 г. [8] (Приложение Д).

В работах [9,10] соискателю принадлежат разработка подходов и принципов информационно-аналитического и технического обеспечения процессов управления наукоемкими проектами. В работе [9] соискателю принадлежат результаты исследования направлений развития специальных информационных систем управления научно-техническими проектами. В работах [1-6,11-14] соискателю принадлежат разработка алгоритмов и методик оценки и отбора наукоемких проектов, а также результаты проектирования информационно-аналитической системы.

Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы.

Апробация работы. Основные результаты диссертационного исследования докладывались и обсуждались на региональных, всероссийских и международных форумах, научных и научно-практических конференциях и выставках: International Forum «Education Quality-2012», (February 20-22, 2012, Izhevsk), XVI «XVI Республиканская выставка-сессия студенческих инновационных проектов», (г. Ижевск, 2013 г.), XVII Международный форум по проблемам науки, техники и образования «III тысячелетие – новый мир», (г. Москва, 2013 г.), I Всероссийская научно-практическая конференция «Современные информационные технологии. Теория и практика» (г. Череповец 2014 г.), XXXVI заочная научная конференция Research Journal of International Studies (г. Екатеринбург, 2015 г.), XX Республиканская выставка-сессия студенческих инновационных проектов», ФГБОУ ВПО «ИжГТУ имени М.Т. Калашникова», г. Ижевск, 2015 г., International Conference "Applied Mathematics, Computational Science and Mechanics: Current Problems" (Воронеж, 17-19 декабря 2018 г.), ITNT 2019: V International Conference on Information Technology and Nanotechnology (г. Самара, 2019 г.).

Диссертационное исследование поддержано стипендией Президента Российской Федерации для молодых ученых и аспирантов, осуществляющих перспективные научные исследования и разработки по приоритетным направлениям модернизации российской экономики в 2013-2015 гг.

Личный вклад автора. Автором лично: предложена методика разработки модели онтологии наукоемкого проекта на основе параметров его оценки в целях формализации функциональной структуры проекта; разработаны математические модели и алгоритмы оценки и отбора наукоемких проектов; описаны условия и процедуры многомерного анализа данных и принятия решений; предложены и разработаны алгоритмы реализации и оптимизации баз знаний на основе модели онтологии; разработаны алгоритмы формирования экспертных рекомендаций; описана методика разработки структуры базы данных специальных информационно-аналитических систем на основе онтологии предметной области.

Структура диссертационной работы. Диссертационная работа состоит из введения, пяти глав, заключения, списка литературы, включающего 170 источников, и 7 приложений. Диссертация изложена на 175 страницах основного текста, содержит 38 таблиц и 60 рисунков.

Описаны принципы разработки модели онтологии предметной области для формализации наукоемких проектов, а также показана важность данной модели для разработки базы знаний и проектирования специальной информационно-аналитической системы.

Во **второй главе** представлена методика построения онтологии наукоемких проектов на основе разработки концептуальной и содержательной моделей предметной области сложных технических систем. Согласно данному подходу онтология наукоемкого проекта представляет собой модель предметной области в виде совокупности «концепт – объект – атрибут». За основу методики принято определение ориентированного графа: V – множество вершин, E – множество дуг, представляющих совокупность связей $e = (p_1, p_2)$ вершины p_1 с вершиной p_2 , $M: E \leftarrow L$ – функция разметки дуг, которая каждой дуге сопоставляет элемент из множества меток L , данные метки отображают взаимное отношение объектов.

Содержательная модель предметной области – совокупность условных атрибутов (параметров) и отношений между ними, которым соответствуют сущности из реального мира, реализованная в виде ориентированного помеченного графа (1):

$$G_d = \{V_d, E_d, M_d, L_d\}, \quad (1)$$

где V_d – совокупность объектов $\{p_1, \dots, p_n\}$ с набором меток (атрибутов) $f \in (1, \dots, t)$, где t – количество значений атрибута.

Здесь каждая вершина (p) является параметром, характеризующим проект, и имеет свою метку f – набор значений объекта для каждого проекта; каждая дуга из вершины p_1 в вершину p_2 с меткой l описывает отношение параметра p_1 к параметру p_2 . Таким образом, любое отношение представляет собой форму взаимовлияния условных атрибутов системы.

Концептуальная модель предметной области G_i – ориентированный граф (2):

$$G_i = \{V_i, E_i, M_i, L_i\}, \quad (2)$$

где V_i – совокупность концептов $\{K_1, \dots, K_m\}$ с меткой g_m .

Вершины графа содержательной модели мы будем называть *концептами* (K), которые представляют собой более общие понятия предметной области, чем ее атрибуты, и дают возможность более объективно представлять, характеризовать и сравнивать между собой различные проекты – другими словами это критерии оценки, признанные экспертами данной предметной области.

Тогда *модель предметной области* (МПО) есть объединение графов концептуальной и содержательной моделей (3):

$$G_{out} = \{V_m, E_m, M_m, L_m\} = \{V(m)_d \cup V(m)_i, E_d \cup E_i, M_d \cup M_i, L_d \cup L_i\}, \quad (3)$$

где $M_m(e) = \begin{cases} M_d(e), & \text{если } e \in E_d \\ M_i(e), & \text{если } e \in E_i \end{cases}$.

Данное определение онтологии объединяет две ее составляющих: концептуальную и содержательную. Содержательная часть выражена в терминах предметной области и отношений между ними. Концептуальная модель является базовым звеном в работе с информацией, поскольку она определяет «каркас», на

который крепится содержательная составляющая.

Далее в работе описаны этапы формализации и описания функциональных зависимостей проектов сложных технических систем. Согласно предложенной методике значения параметров проекта p получены путем определения весомости альтернативных вариантов его представления b_i методом решающих матриц. Таким образом, общая задача расчета значений критериев оценки K_n на основе данной модели онтологии сводится к снижению размерности параметрического пространства наукоемкого проекта путем аддитивной свертки его параметров с учетом значений их влияния на критерии оценки (4):

$$K_n = p_1 b_{1m} + \dots + p_j b_{jm}, \quad (4)$$

где K_n – критерий состояния проекта, j – количество параметров, заключенных в критерий оценки проекта, b_m – коэффициент влияния параметра на критерий, p_j – значение параметра проекта.

Для ранжирования параметров в составе агрегирующего их критерия и расчета коэффициентов весомости параметров проведен индивидуальный экспертный опрос. Выявлены характерные особенности основных типов наукоемких проектов. Далее путем аддитивной свертки параметрического пространства получены значения критериев оценки проекта, позволяющие сформировать многомерную информационную модель каждого типа научного проекта.

Для решения задачи комплексной оценки проекта разработана система нечеткой логики на основе алгоритма Мамдани. Входами являются представленные выше критерии, выходом – вероятностные оценки привлекательности проекта, соответствующие значениям входных параметров.

На основе разработанных математических моделей построен алгоритм отбора наукоемких проектов путем решения задачи многокритериальной оптимизации (МКО), заключающейся в задании параметрических и функциональных ограничений множества проектов на основе набора заданных характеристик и предпочтений лица принимающего решения. В рамках данного подхода задача МКО определяется следующим образом: требуется найти значения параметров x_1, x_2, \dots, x_n , удовлетворяющие заданному ограничению $g_i(x_1, x_2, \dots, x_n) \leq b_i$, $i = 1, 2, \dots, m$, для которых функции отбора $z_k = f_k(x_1, x_2, \dots, x_n)$, x_1, x_2, \dots, x_n достигают наибольшего значения.

Все точки множества проектов $X = (x_1, x_2, \dots, x_n)$, удовлетворяющего заданным условиям, образуют допустимую область D , а G – множество точек x , одновременно удовлетворяющих параметрическим и функциональным ограничениям, при этом $G \subseteq D$ (рисунок 2).

Указанные алгоритмы и модели формализации наукоемких проектов позволили сформулировать основные требования к функциональным возможностям информационно-аналитической системы поддержки процесса управления проектом и перейти к ее проектированию и реализации.

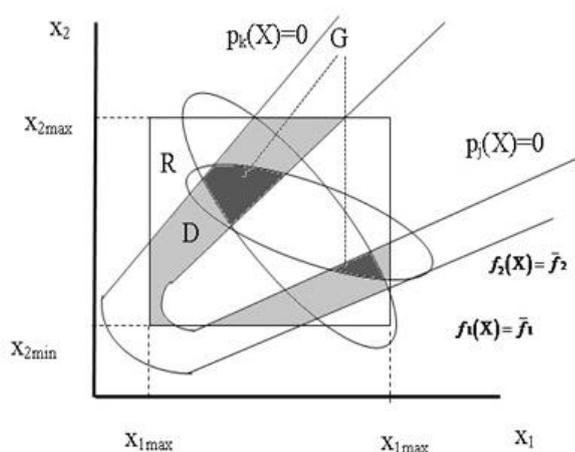


Рисунок 2 – Формирование функциональными ограничениями множества наукоемких проектов СТС согласно запросу ЛПР

В третьей главе показана эффективность современных инструментов многомерного анализа в смежных областях, позволяющая накапливать большие объемы разнородных данных, формировать различные выборки данных, проводить автоматизированный анализ данных и выявлять закономерности, способствуя их самостоятельному развитию.

В рамках современных тенденций развития информационных технологий описана методика разработки специальных информационных систем на основе модели онтологии предметной области, позволяющая сократить время внедрения разработанных алгоритмов в конкретные сферы

деятельности. В этом случае концепты разработанной модели онтологии в зависимости от их информационной нагрузки и целей проектирования ИС могут представлять собой разделы одной базы данных (БД) или же отдельные базы данных G_i_db , а относящиеся к ним параметры p_n получают первичный ключ (идентификаторы) p_n_id и связываются согласно установленным связям ed соответствующего типа. Тогда реляционная база данных RDB , описывающая онтологическую модель, представляет собой кортеж (5):

$$RDB = \langle G_i_db, p_n_id, e_d \rangle, \quad (5)$$

где G_i_db – разделы базы данных, p_n_id – вершины графа, ed – ребра графа.

В целях анализа потока работа и данных в процессе реализации наукоемких проектов проведено моделирование системных связей научно-технической системы, построены процессная и концептуальная модели научно-технической системы, отражающие весь комплекс мероприятий по работе с информацией о наукоемких проектах разработки сложных технических систем.

Формально содержание задачи принятия решения можно представить следующим образом: в проблемной ситуации N , характеризуемой совокупностью параметров $p(n)$, имеющегося времени t и располагаемого опыта и знаний по решению проблем Q необходимо уточнить множество желаемых целей P , имеющихся технологических и технических ограничений U , подобрать варианты решений Y на основе оценки соответствия их параметров $p(y)$ заданной ситуации, далее, согласно алгоритму выбора варианта решения K и предпочтениям лица принимающего решения (ЛПР) – F , найти оптимальное решение $Y(n)$.

В общем виде задача принятия решения может быть приведена к следующей формальной модели (6):

$$(N, p(n), t, Q, P, U, Y, p(y), K, F) \rightarrow Y(n). \quad (6)$$

Описание функциональных зависимостей основных элементов сложных научно-технических систем выявило направления и содержательный аспект

анализа информации о наукоемких проектах, что позволило сформулировать подход к разработке многомерной модели данных. На основе предлагаемой архитектуры ИС разработана методика формирования базы знаний (БЗ) предметной области в виде продукционных моделей, а также алгоритмы многомерного и интеллектуального анализа проектов.

При этом требования к адаптивности и функциональному развитию БЗ экспертной системы удовлетворяются путем использованием модуля самообучения на основе алгоритма дерева решений (*CART, Classification and Regression Tree*). Приемлемым типом критерия останковки обучения было выбрано задание минимального количества примеров в каждом классе, данное решение позволяет экономить ресурсы информационной системы, а также получить необходимый уровень детализации правил базы знаний. В качестве метрики точности алгоритма выбрана *F-мера*, представляющая собой среднее гармоническое метрик точности (*precision*) и полноты (*recall*) анализа данных.

Таким образом, сформулирована методология разработки специальных ИС на основе объектно-когнитивного анализа предметной области, методах объектно-ориентированного анализа, онтологического анализа и семантической сети представления знаний.

В четвертой главе приведен пример реализации разработанной методики. В качестве объекта исследования выбран бизнес-инкубатор ФГБОУ ВО «ИжГТУ имени М.Т. Калашникова».

На основе анализа предметной области сформулированы требования к информационной системе. Далее для построения модели онтологии отобраны эксперты, непосредственно участвующие в процессе организации и управления наукоемкими проектами, ими выступили сотрудники кафедр экономики и менеджмента, а также руководители компаний и специалисты бизнес-инкубатора. Проведена оценка компетенций экспертов, что позволит получить средневзвешенные значения результатов опроса.

На основе описания атрибутов и концептов предметной области построена онтологическая модель, фрагмент которой приведен на рисунке 3.

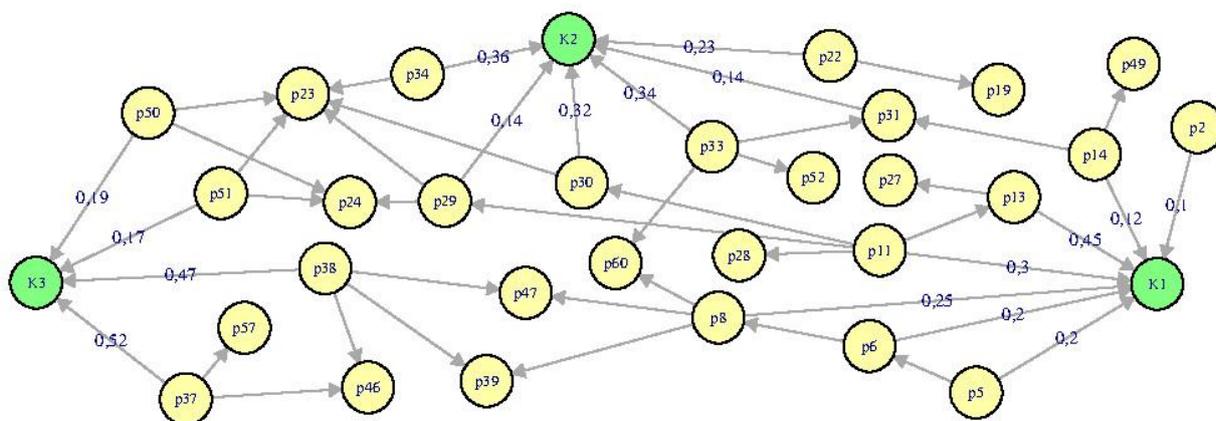


Рисунок 3 – Фрагмент модели онтологии предметной области наукоемких проектов

На схеме использованы следующие обозначения: p_n – параметры проекта, $K1$

– критерий «Профессиональные навыки руководителя проекта», K_2 – критерий «Научная часть проекта», K_3 – критерий «Обеспеченность проекта ресурсами».

Рисунок 3 описывает структуру выбранных атрибутов и концептов исследуемой предметной области, выраженных параметрами проекта и критериями его оценки, а также форму и направление их взаимного влияния. Также на схеме отображены коэффициенты связи параметров проекта с соответствующим критериям его оценки.

Разработка базы правил нечеткой системы велась также при участии выбранных экспертов. База правил отражает функциональные связи между значениями факторов и результативностью проекта, представленных совокупностью нечетких логических условий «если-то» (if – then).

Далее на основе глубокого исследования собранных данных о предметной области сформулированы понятия уровня научно-технического развития наукоемкого проекта и степени его соответствия целевому состоянию, получены их математические модели.

Текущий научно-технического развития проекта можно представить как (7):

$$Ik = z_i + u_i, \quad (7)$$

где Ik – уровень научно-технического развития проекта, z_i – оценка инженерно-технического уровня проекта, u_i – оценка научного развития проекта.

Как показал анализ предметной области, при реализации каждого проекта имеется основной критерий оценки достижения им поставленных целей, например, таковым может являться размер выделенного бюджета, заданные тактико-технические характеристики, показатели надежности и т.п. В данном случае подграф онтологии G'_{ont} , с вершиной, заданной основным критерием (K'_m) и его отношениями $b' = (K'_m, p'_n)$, определяет совокупность параметров проекта для оценки. Рассчитав целевое значение данного критерия оценки проекта на основе выделенной совокупности параметров и сравнив его с текущим значением критерия, получим представление о степени достижения проектом целевого состояния. Тогда значение степени достижения целевого состояния проектом рассчитывается как (8):

$$Q = \frac{K'_m}{L} \times 100, \quad (8)$$

где Q – степень соответствия целевому состоянию, в процентах; L – целевое значение состояния проекта по основному критерию его оценки.

Совокупность данных расчетных характеристик проекта позволяет наиболее точно определить его состояние в рамках специального алгоритма экспертной системы. Таким образом, разработана комплексная методика оценки проектов, отвечающая задачам исследования.

На основе разработанных ранее алгоритмов и требований, предъявляемых к функциональным особенностям информационной системы управления наукоемкими проектами, описан комплекс функциональных модулей СППР, а также выявлены формы и виды принимаемых решений в процессе управления наукоемкими проектами. Формализованы и описаны знания эксперта исследуемой предметной области в виде концептуальной и функциональной моделей поля знаний (рисунок 4).



Рисунок 4 – Функциональная составляющая поля знаний

Данные модели позволили определить условия и порядок принятия решения и разработать базу правил для формирования промежуточных выводов и итоговых решений по работе с конкретным проектом. С целью подробного представления результатов работы алгоритмов ЭС формируемые рекомендации целесообразно отобразить в виде промежуточных выводов и подведения итогового решения, представляя их отдельными блоками информации.

Примем следующие условные обозначения к описанию правил базы знаний: x – проект; OO – общая оценка проекта; $CRП$ – степень реализуемости проекта; $OЗ$ – оценка завершенности проекта; $S(x)$ вывод о состоянии проекта; $F(x)$ – рекомендации по работе проектом x в случае участия в выбранном конкурсе; $H(x)$ – активность руководителя в развитии данного проекта; $L(x)$ – оценка возможности участия в конкурсе; $M(x)$ – результаты многомерного анализа проекта x ; $T(x)$ – рекомендации по подготовке заявки по проекту x на выбранный конкурс; $U(x)$ – оценка возможности автоматического формирования заявки; $N(x)$ – оценка результативности предыдущих заявок по проекту; $R(x)$ – решение по участию проекта x в выбранном конкурсе.

Блок «Состояние проекта» содержит информацию, полученную при обработке рассчитанных оценок проекта, вывод о состоянии проекта строится на основе сопоставления общей оценки, степени реализуемости и оценки завершенности (9):

$$S(x) = OO \wedge CRП \wedge OЗ. \quad (9)$$

Правило базы знаний для данного блока задается следующим условием: «Формирование вывода» = <«Состояние проекта», блок актуализирован, $OO \wedge CRП \wedge OЗ \rightarrow$ решение = {вывод}, иначе = «недостаточно данных», описание состояния проекта».

Блок «Работа с проектом» выводит рекомендации по работе с проектом, основываясь на результатах анализа ситуации развития проекта (10):

$$F(x) = \{ < H, L, M > \}. \quad (10)$$

Правило базы знаний для данного блока задается следующим условием: «Формирование вывода» = <«Работа с проектом», блок актуализирован, $H(x) \vee L(x) \rightarrow$ решение = {рекомендация}, иначе = «недостаточно данных», вывод рекомендации по работе с проектом».

Блок «Работа с заявкой» также выводит рекомендации по работе с проектом,

основываясь на результатах анализа предыдущих результатов подготовки заявок (11):

$$T(x) = \{ \langle U, N, M \rangle \}. \quad (11)$$

Правило базы знаний для данного блока задается следующим условием: «Формирование вывода» = \langle «Работа с заявкой», блок актуализирован, $U(x) \vee N(x) \rightarrow$ вывод = {рекомендация}, иначе = «недостаточно данных», вывод рекомендации по работе с заявкой \rangle

Блок «Решение» выводит итоговые рекомендации исходя из заданных условий, состояния проекта и оценки активности по проекту в виде (12):

$$R(x) = \{ \langle S(x), F(x), T(x) \rangle \}. \quad (12)$$

Правило базы знаний для формирования окончательного решения по проекту задается следующим условием: «Формирование решения» = \langle «Решение», блок актуализирован, $S(x) \wedge F(x) \wedge T(x) \rightarrow$ вывод = {решение}, иначе = «недостаточно данных», решение об участии в конкурсе \rangle .

Далее представлен алгоритм самообучения и оптимизации правил разработанной базы знаний на основе алгоритм CART. Для определения критерия остановки обучения, с помощью разработанной ранее анкеты собраны данные 500 проектов, участвующих в Республиканских выставках-сессиях студенческих инновационных проектов в 2012-2017 г.г., проводимых бизнес-инкубатором ФГБОУ ВО «ИжГТУ имени М.Т. Калашникова». Анализ показал, что выбор уровня критерия остановки равный наличию не менее 4% объектов в каждом классе при незначительном снижении точности классификации позволяет получить оптимальную структуру дерева решения для его формализации.

В качестве примера в работе проведено обучение на выборке проектов, участвовавших в одном конкурсе, однако данный подход применим и к обучению системы и по другим параметрам проектов, что с учетом скорости работы алгоритма позволяет на каждый запрос формировать рекомендации актуальные заданным условиям и накопленным данным.

В работе приведена структура информационной поддержки научно-технической деятельности на основе представленных алгоритмов работы с данными о наукоемких проектах разработки сложных технических систем. Технологическая схема ИС поддержки процессов реализации наукоемких проектов сложных технических систем представлена на рисунке 5.

Описан пример анализа и оценки наукоемких проектов сложных технических систем. На рисунке 6 представлено визуальное сравнение решения конкурсной комиссии и рекомендации ИАС «UNIProject» по каждому проекту в рамках одного конкурса, по оси Y отложены варианты решений экспертной комиссии и разработанной ИАС в порядке возрастания их уровня.

Коэффициент корреляции рангов выделенных уровней экспертных решений равен 83%, что говорит о высокой сопоставимости результатов оценки проектов и экспертных рекомендации ИАС «UNIProject» с результатами экспертной комиссии. Данная экспертная система может использоваться как в качестве самостоятельного программного решения, так и в качестве подсистем для других программных комплексов.

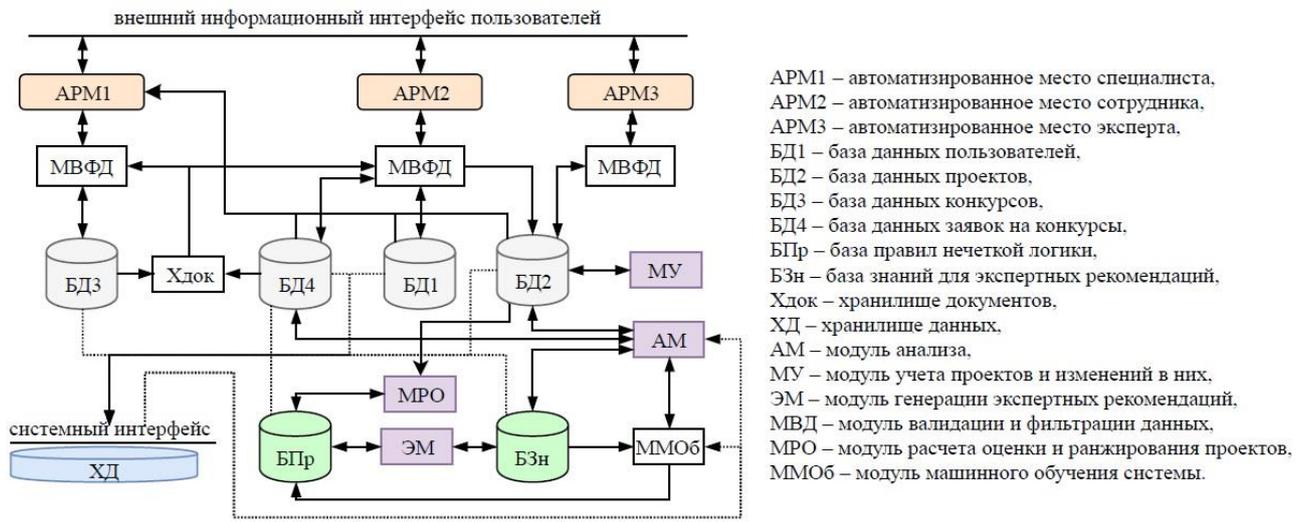


Рисунок 5 – Технологическая схема ИС поддержки наукоемких проектов

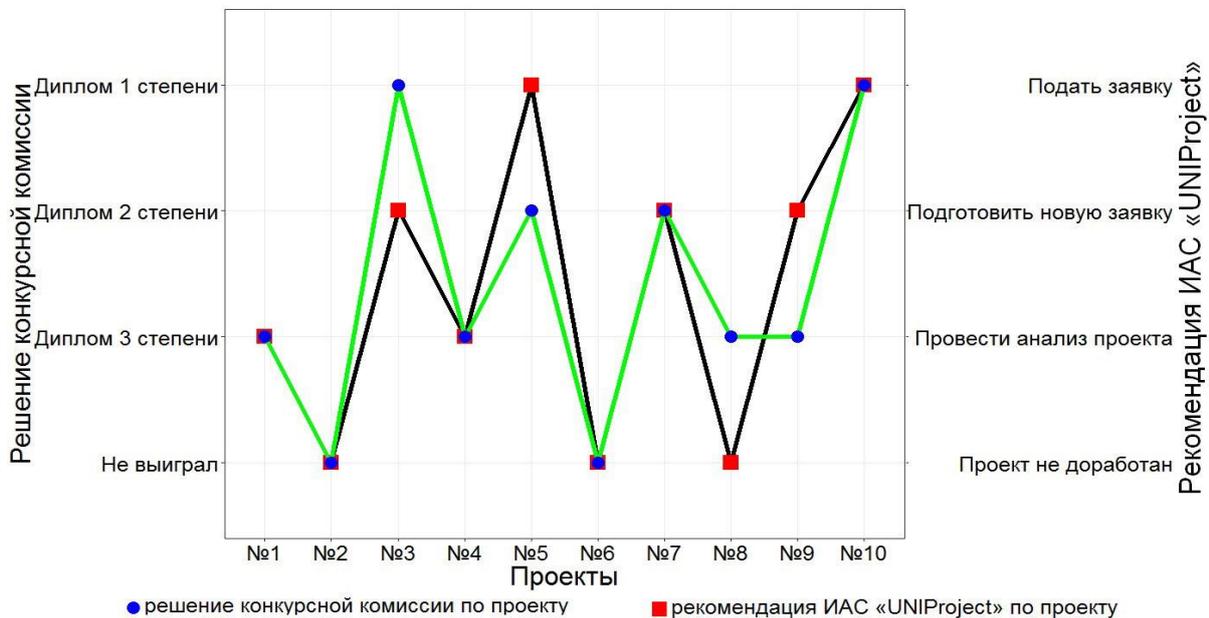


Рисунок 6 – Сравнение экспертных рекомендаций конкурсной комиссии и ИАС «UNIPProject»

В пятой главе представлен пример использования разработанной методики формализации и автоматизации процессов анализа наукоемких проектов в режиме реального времени. В ходе анализа ключевых атрибутов предметной области эксперты установили форму и силу их взаимного влияния, что позволило построить модель онтологии предметной области, затем определены параметры проектов и критерии их оценки.

На основе предложенной методики формализованы функциональные зависимости параметров проекта, что позволило разработать комплексную методику оценки наукоемких проектов, отвечающую задачам исследования. Согласно потребностям онлайн-платформы описаны условия и порядок принятия управленческого решения, далее проведена формализация знаний эксперта исследуемой предметной области. Алгоритм «Random forest» на множестве

решающих деревьев *CART* позволил оценить значимость критериев оценки проекта и выбрать признаки для разработки условий базы знаний. На рисунке 7 представлены результаты анализа признаков модели в порядке убывания их значимости.

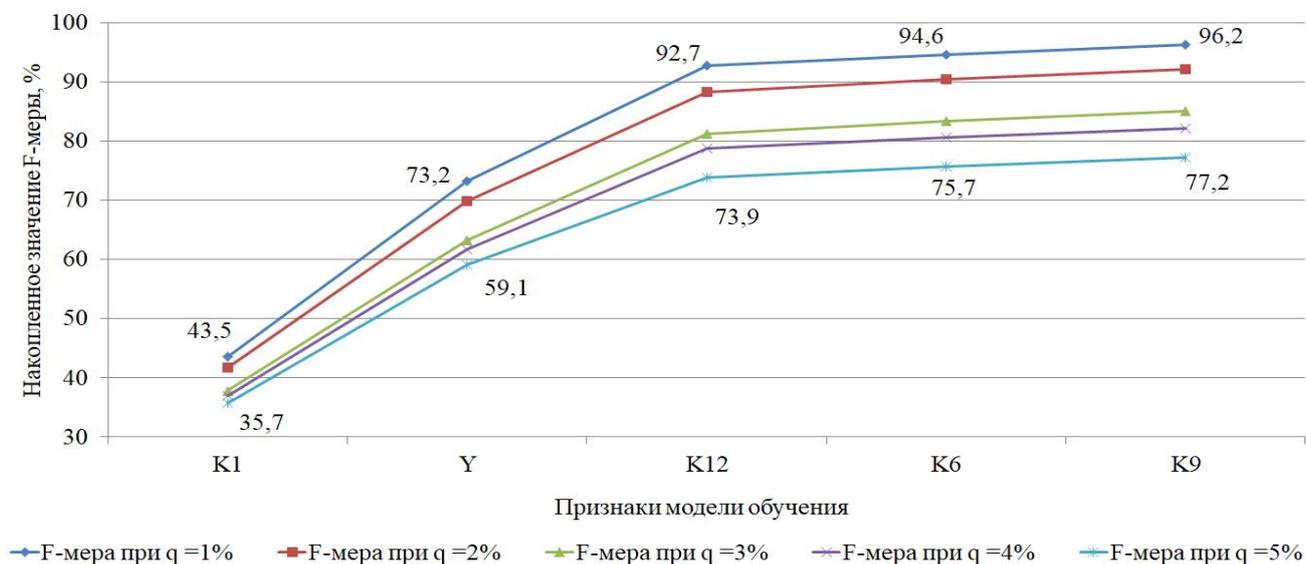


Рисунок 7 – Метрика качества модели в зависимости от уровня критерия остановки обучения

Обозначения на рисунке: *K1* – критерий «Оценка руководителя», *K6* – критерий «Оценка текущих результатов», *K9* – критерий «Рентабельность», *K12* – критерий «Оценка рисков», *Y* – общая оценка проекта, *q* – минимальное количество наблюдений в листе дерева (в % от числа наблюдений), *F-мера* – метрики точности алгоритма: $F\text{-мера} = 2(\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$, где $\text{precision} = TP / (TP + FP)$, $\text{recall} = TP / (TP + FN)$, здесь: *TP* – истинно-положительное решение, *FP* – ложно-положительное решение (ошибка I рода), *FN* – ложно-отрицательное решение (ошибка II рода), *TN* – истинно-отрицательное решение.

Из рисунка 7 видим, что признаки «Оценка руководителя», «Общая оценка» и «Оценка рисков» при уровне остановки обучения 2% от количества наблюдений в совокупности позволяют на уровне 90-процентной эффективности классифицировать проекты в терминах предметной области, выберем данные критерии для формализации условий базы знаний.

Таким образом, окончательное решение по финансированию проекта зависит от общей оценки состояния проекта и уровня принятия рисков инвестором. Пусть *x* – проект, *S(x)* – состояние проекта, *F(x)* – мнение эксперта по проекту *x* и *R(x)* – окончательное решение о поддержке проекта. На основе введенных переменных опишем алгоритмы базы правил и условия их срабатывания.

Блок «Состояние проекта» содержит информацию, полученную при обработке рассчитанных критериев и общей оценки проекта на основе разработанного алгоритма, тогда состояние проекта будет задано функцией (13):

$$S(x) = Y \wedge K1 \wedge K12, \quad (13)$$

где *S(x)* – вывод о состоянии проекта *x*, $\forall x \exists S_x = \{\text{вывод}\}$; *Y* – общая оценка,

$\forall x \exists Y_x \in [0;10]$; $K1$ – уровень оценки опыта руководителя, $\forall x \exists K1_x \in [0;10]$; $K12$ – уровень оценки рисков проекта, $\forall x \exists K12_x \in [0;10]$.

Итак, вывод о состоянии проекта строится на основе сопоставления общей оценки, уровня оценки опыта руководителя и уровня оценки рисков проекта. Тогда правило базы знаний для данного блока выглядит так: «Формирование вывода» = $\langle \text{«Состояние проекта»}, \text{блок актуализирован}, Y \wedge K1 \wedge K12 \rightarrow \text{решение} = \{\text{вывод}\}, \text{иначе} = \langle \text{«недостаточно данных»}, \text{описание состояния проекта} \rangle$.

Блок «Решение» выводит итоговые рекомендации исходя из состояния проекта и мнения эксперта. Обозначим ее функцией (14):

$$R(x) = \{ \langle S(x), F(x) \rangle \}, \quad (14)$$

где $R(x)$ – решение о финансировании проекта x , $R(x) = \{ \text{решение} \}$.

Алгоритм базы правил для формирования окончательного решения по проекту представлен в следующем виде: «Формирование решения» = $\langle \text{«Решение»}, \text{блок актуализирован}, S(x) \wedge F(x) \rightarrow \text{вывод} = \{\text{решение}\}, \text{иначе} = \langle \text{«недостаточно данных»}, \text{решение об участии в конкурсе} \rangle$.

Таким образом, создана база знаний онлайн-платформы, выраженная на естественном языке и в терминах предметной области, данная БЗ является ядром модуля принятия решений, описаны условия ее оптимизации с использованием разработанного алгоритма самообучения.

В качестве примера реализации разработанной методики проведен анализ и оценка научно-технических проектов, по результатам которого получены автоматические экспертные рекомендации СППР по работе с наукоемкими проектами разработки сложных технических систем (таблица 1).

Таблица 1 – Классификация проектов на основе разработанных алгоритмов аналитического модуля онлайн-платформы «ScientificCoin»

Название проекта	$K1$	$K12$	Y	Автоматическая рекомендация системы
Безмасляный двигатель с подвижным цилиндром	6,35	2,33	6,38	Имеет высокий потенциал реализации и низкие риски
Тандемный ускоритель с вакуумной изоляцией	3,85	2,67	5,92	Перспективный проект, имеет приемлемый риск
Радиографическая установка	5,00	4,79	5,18	Проект имеет научный потенциал, средний риск

Согласно сформулированным ранее правилам базы знаний из рассматриваемой совокупности проектов наиболее привлекательным на данный момент является проект разработки безмасляного двигателя (проект №1).

Таким образом, реализация разработанных интеллектуальных алгоритмов позволяет в реальном времени формировать рекомендации актуальные накопленным данным и предпочтениям пользователей. Эффективность разработанного аналитического модуля подтверждается актом внедрения в бизнес-модель онлайн-платформы.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

В ходе диссертационного исследования получены следующие результаты:

1. Дано определение наукоемкого проекта разработки сложной технической системы, показана его структура, функциональные связи и механизмы взаимодействия с системами из внешнего окружения. Проведен комплексный анализ процесса принятия решений о способах и методах реализации проектов. Описаны принципы разработки модели онтологии предметной области для определения совокупности понятий, характеризующих наукоемкие проекты, а также показана важность данной модели для проектирования и реализации автоматизированной информационной системы на основе базы знаний.

2. Концептуально-содержательный подход к описанию ключевых понятий предметной области наукоемких проектов позволил выявить структуру параметров проекта и формализовать их взаимосвязи в форме онтологической модели предметной области. Разработаны авторские алгоритмы оценки и отбора наукоемких проектов разработки сложных технических систем на основе системы нечеткого логического вывода для получения интегральной оценки привлекательности проектов и многокритериальной оптимизации параметрического пространства проектов, способствующие решению многокритериальной задачи выбора наиболее приоритетного проекта в соответствии с предпочтениями лица принимающего решения. Приведены примеры реализации алгоритмов при расчете интегрального показателя привлекательности наукоемкого проекта в рамках ИАС «UNIProject» на основе 8 входных параметров и онлайн-платформы «ScientificCoin» на основе 12 входных параметров.

3. В рамках методики формализации наукоемких проектов разработаны модели анализа и управления данными проектов разработки сложных технических систем. На основе модели онтологии предметной области описан подход к разработке структуры базы данных, способной поддерживать функции аккумуляции, хранения, анализа и обработки больших объемов данных. Описаны формы представления экспертных знаний, выявлены направления и содержание анализа данных о проектах.

4. Представлены примеры реализации разработанной методики формализации наукоемких проектов в двух различных аналитических системах: ИАС «UNIProject», содержащей 500 проектами в БД и поддерживающей их анализ по 60 параметрам, а также аналитического модуля онлайн-платформы «ScientificCoin», аккумулирующей данные более чем 2500 проектов и поддерживающей их анализ по 75 параметрам.

5. Представлена методика формализации и описания структуры знаний эксперта в виде поля знаний исследуемой предметной области, определены условия и порядок принятия управленческого решения, что позволяет разрабатывать специальные базы знаний для формирования рекомендаций пользователям ИАС по работе с наукоемкими проектами СТС. Описан алгоритм самообучения и оптимизации базы знаний с использованием алгоритма деревьев

решений, в качестве критерий качества алгоритма выбрана F-мера, показавшая качество обучения на собранных данных на уровне 90%.

6. В рамках реализации методики сформирован и описан комплекс функциональных модулей СППР для бизнес-инкубатора и онлайн-платформы реального времени. Реализован графический интерфейс доступа к функционалу ИАС «UNIProject», реализуемому особенностями базы данных и входящих в ее структуру функциональных модулей. Показано, что разработанная СППР способствует повышению оперативности принимаемых решений в 12,5 раз за счет увеличения скорости анализа и обработки информации. Представлены примеры формирования автоматических рекомендаций СППР для наукоемких проектов по разработке сложных технических систем в ИАС «UNIProject» и проектов на онлайн-платформе «ScientificCoin».

Научная новизна выводов и результатов диссертации состоит в разработке алгоритмов принятия решений при реализации наукоемких проектов на основе их многомерного анализа, комплексной оценки и разработки соответствующего программного обеспечения, способствующих повышению эффективности проектирования и реализации сложных технических систем.

ПУБЛИКАЦИИ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

I. Монография

1. Благодатский Г.А., Горохов М.М., Переведенцев Д.А. Методы и инструменты многомерного анализа баз данных перспективных научных разработок: монография. – Ижевск: Изд-во ИжГТУ имени М. Т. Калашникова, 2018. – 180 с.

II. Статьи в журналах, рекомендованных ВАК РФ

2. Благодатский Г.А., Горохов М.М., Переведенцев Д.А. Моделирование системы нечеткого логического вывода оценки наукоемких проектов // Автоматизация процессов управления. – 2017. – №2(48). – С.84-91.
3. Горохов М.М., Переведенцев Д.А. Проектирование базы знаний экспертной системы управления научными проектами // Вестник Томского Государственного университета. Управление, вычислительная техника и информатика. – 2017. – №1(38). – С.47-51.
4. Горохов М.М., Переведенцев Д.А. Формирование многомерной модели данных для целей OLAP – анализа в информационно-аналитической системе управления научными проектами // Вестник Воронежского университета. – 2016. – №3. – 124-130.
5. Переведенцев Д.А. Разработка методики параметрической оценки научных и инновационных проектов // Auditorium: электронный научный журнал Курского государственного университета. – 2015. – № 3 (07). – С. 56-65.
6. Переведенцев Д.А. Разработка UML – модели информационно-аналитической системы перспективных научных проектов // «Вестник ИжГТУ имени М. Т. Калашникова». – 2015. – № 4. – С. 58-60.

III. Свидетельства о регистрации базы данных и программы для ЭВМ

7. Благодатский Г.А., Горохов М.М., Переведенцев Д.А. База данных научных проектов «IProjects». Свидетельство о регистрации базы данных № 2016620179, дата государственной регистрации 08.02.2016 г. / правообладатель ФГБОУ ВПО ИжГТУ им. М.Т. Калашникова (RU), - Реестр баз данных. – 1 с.
8. Переведенцев Д.А. Информационно-аналитическая система «UNIProject». Свидетельство о регистрации программы для ЭВМ №2016615251, дата государственной регистрации 19.05.2016 г. / правообладатель ФГБОУ ВПО ИжГТУ им. М.Т. Калашникова (RU), - Реестр программ для ЭВМ. – 1 с.

IV. Материалы международных конференций

9. Переведенцев Д.А., Переведенцев К.А., Романов К.А. Принципы информационно-аналитического обеспечения процесса коммерциализации научно-технических работ // Труды XVII Международного форума по проблемам науки, техники и образования. / Под редакцией В.А. Малинникова, В.В. Вишневого – М.: Академия наук о Земле, 2013 г. – С. 122-123.
10. Переведенцев Д.А., Переведенцев К.А., Романов К.А. Информационно-технические ресурсы инновационной инфраструктуры бюджетного учреждения (на примере университета) // Сборник по результатам XXVI заочной научной конференции Research Journal of International Studies. Часть 2. – 2014. – №4 (23). – С. 50-51.
11. Perevedencev D.A. The capabilities system operational analysis of data for intensification the process of commercialization innovation / Meždunarodnyj naučno-issledovatel'skij žurnal. – №2 (33) 2015. – p. 63-64.
12. Переведенцев Д.А. Система поддержки принятия решения для организации и управления научными проектами предприятий // Международная научно-практическая конференция «Развитие экономического образования в современном мире» Ижевск: Изд-во ИжГТУ им. М.Т. Калашникова, – 2015. – С.45-51.

V. Публикации в других изданиях

13. Переведенцев Д.А. Программные средства анализа баз данных перспективных научных разработок на основе технологии многомерной обработки данных / сборник трудов IV Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых «Актуальные проблемы современной науки: взгляд молодежи»: сборник трудов, 21 апреля 2015 г. / [науч. ред. О.С. Нагорная]. – Челябинск: Центр оперативной полиграфии “Violitprint”. – С. 569-572.
14. Благодатский Г.А., Вологдин С.В., Горохов М.М., Переведенцев Д.А. Методика автоматизированной оптимизации базы знаний экспертной системы на основе алгоритма решающих деревьев (CART) / Актуальные проблемы прикладной математики, информатики и механики: сборник трудов Международной научной конференции, Воронеж, 17–19 декабря 2018 г. – Воронеж: Издательство «Научно-исследовательские публикации», 2019. – С. 1541-1547.