

# ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ

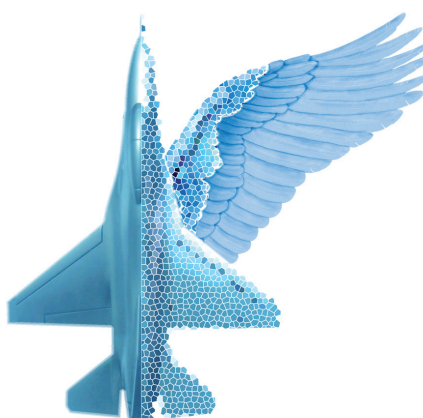


# ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ

Научный журнал

Том 9

№ 1(31)



## EDITORIAL BOARD – РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Anatoly I. Belousov	Белоусов Анатолий Иванович, д.т.н., профессор, Самарский университет, Самара
Nikolay M. Borgest	Боргест Николай Михайлович*, к.т.н., профессор, Самарский университет, член ИАОА, Самара
Stanislav N. Vasiliev	Васильев Станислав Николаевич*, д.ф.-м.н., профессор, академик РАН, ИПУ РАН, Москва
Vladimir G. Gainutdinov	Гайнутдинов Владимир Григорьевич, д.т.н., профессор, КНИТУ-КАИ, Казань
Vladimir V. Golenkov	Голенков Владимир Васильевич*, д.т.н., профессор, БГУИР, Минск
Vladimir I. Gorodetsky	Городецкий Владимир Иванович*, д.т.н., профессор, ООО «ИнфоВингс», Санкт-Петербург
Valeriya V. Gribova	Грибова Валерия Викторовна*, д.т.н., с.н.с., ИАПУ ДВО РАН, Владивосток
Yury A. Zagorulko	Загорюлько Юрий Алексеевич*, к.т.н., с.н.с., ИСИ СО РАН, Новосибирск
Anton V. Ivaschenko	Иващенко Антон Владимирович, д.т.н., профессор, СамГТУ, Самара
Valery A. Komarov	Комаров Валерий Андреевич, д.т.н., профессор, Самарский университет, Самара
Victor M. Kureichik	Курейчик Виктор Михайлович*, д.т.н., профессор, Южный федеральный университет, Таганрог
Dmitry V. Lande	Ландэ Дмитрий Владимирович*, д.т.н., с.н.с., ИПРИ НАН Украины, Киев
Lyudmila V. Massel	Массель Людмила Васильевна*, д.т.н., профессор, ИСЭМ СО РАН, Иркутск
Aleksandr Yu. Nesterov	Нестеров Александр Юрьевич, д.филос.н., профессор, Самарский университет, Самара
Dmitry A. Novikov	Новиков Дмитрий Александрович, д.т.н., проф., член-корреспондент РАН, ИПУ РАН, Москва
Alexander V. Palagin	Палагин Александр Васильевич, д.т.н., проф., академик НАН Украины, Ин-т кибернетики, Киев
Semyon A. Piyavsky	Пиявский Семён Авраамович, д.т.н., профессор, ПГУТИ, Самара
Yury M. Reznik	Резник Юрий Михайлович, д.филос.н., профессор, Институт философии РАН, Москва
George Rzevski	Ржевский Георгий, профессор, Открытый университет, Лондон
Roman O. Samsonov	Самсонов Роман Олегович, д.т.н., Самарский университет, Самара
Peter O. Skobelev	Скобелев Петр Олегович, д.т.н., НПКи «Разумные решения», Самара
Sergey V. Smirnov	Смирнов Сергей Викторович*, д.т.н., ИПУСС РАН, член ИАОА, Самара
Peter I. Sosnin	Соснин Петр Иванович*, д.т.н., профессор, УлГТУ, Ульяновск
Dzhavdet S. Suleymanov	Сулейманов Джавдет Шевкетович*, д.т.н., профессор, академик АН РТ, Казань
Boris E. Fedunov	Федунов Борис Евгеньевич*, д.т.н., профессор, ГосНИИ авиационных систем, Москва
Altynbek Sharipbay	Шарипбай Алтынбек*, д.т.н., профессор, Институт искусственного интеллекта, Астана
Boris Ya. Shvedin	Шведин Борис Яковлевич, к.психол.н., ООО «Дан Роуз», член ИАОА, Ростов-на-Дону

\* - члены Российской ассоциации искусственного интеллекта - [http://www.raai.org/about/about.shtml?raai\\_list](http://www.raai.org/about/about.shtml?raai_list)

## Executive Editorial Board - Исполнительная редакция

Lead Editor	Samsonov R.O.	Шеф-редактор	Самсонов Р.О.	Первый проректор Самарского университета
Chief Editor	Skobelev P.O.	Главный редактор	Скобелев П.О.	директор НПКи «Разумные решения»
Deputy Chief Editor	Smirnov S.V.	Зам. главного редактора	Смирнов С.В.	зам. директора ИПУСС РАН по науке
Executive Editor	Borgest N.M.	Выпускающий редактор	Боргест Н.М.	директор издательства «Новая техника»
Editor	Kozlov D.M.	Редактор	Козлов Д.М.	доцент Самарского университета
Technical Editor	Simonova A.U.	Технический редактор	Симонова А.Ю.	редактор издательства «Новая техника»
Translation Editor	Korovin M.D.	Редактор перевода	Коровин М.Д.	инженер Самарского университета

## CONTACTS FOUNDERS – КОНТАКТЫ УЧРЕДИТЕЛЕЙ

### ИПУСС РАН

443020, Самара, ул. Садовая, 61  
тел./факс.: +7 (846) 333 27 70

Смирнов С.В.  
[smirnov@iccs.ru](mailto:smirnov@iccs.ru)

### Самарский университет

443086, Самара, Московское шоссе 34, корп. 10  
тел.: +7 (846) 267 46 47, факс.: +7 (846) 267 46 46

Боргест Н.М.  
[borgest@yandex.ru](mailto:borgest@yandex.ru)

### ООО «Новая техника» (издательство)

Адрес редакции: 443010, Самара, ул. Фрунзе, 145, тел.: +7 (846) 332 67 84, факс.: +7 (846) 332 67 81

The journal has entered into an electronic licensing relationship with EBSCO Publishing, the world's leading aggregator of full text journals, magazines and eBooks. The full text of JOURNAL can be found in the EBSCOhost™ databases.

Журнал размещен в коллекции «Издания по естественным наукам» на платформе EastView.



The journal has been successfully evaluated in the evaluation procedure for the ICI Journals Master List 2014-2017 and journal received the ICV (Index Copernicus Value) of 67.46 points (2014), 67.64 (2015), 77.98 (2016), 87.78. (2017).

Журнал включён в перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёной степени кандидата наук, на соискание учёной степени доктора наук (Письмо Департамента аттестации научных и научно-педагогических работников Минобрнауки РФ от 01.12.2015 № 13-6518) по группам научных специальностей 05.13.00 и 05.07.00.

Журнал включён в РИНЦ. Пятилетний импакт-фактор РИНЦ 0,968 (2013), 0,895 (2014), 1,305 (2015), 1,055 (2016), 0,841 (2017).

Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство ПИ № ФС 77-70157 от 16.06.2017 г. (ранее выданное свидетельство ПИ № ФС 77-46447 от 07.09.2011 г.)

[http://agora.guru.ru/scientific\\_journal/](http://agora.guru.ru/scientific_journal/)



Отпечатано в ООО «Новая техника», г. Самара, пр. К. Маркса, 24-76.  
Дата выхода 29.03.2019. Тираж 300 экз. Свободная цена. (6+).

© Все права принадлежат авторам публикуемых статей  
© ООО «Новая техника» - «New Engineering» Ltd., 2011-2019  
© Самарский университет - Samara University, 2015-2019  
© ИПУСС РАН - ICCS RAS, 2015-2019

## СОДЕРЖАНИЕ

### ОТ РЕДАКЦИИ

- ВЫХОД ЗА ГРАНИЦЫ ИНДУСТРИИ 5-6

### ОБЩИЕ ВОПРОСЫ ФОРМАЛИЗАЦИИ ПРОЕКТИРОВАНИЯ: ОНТОЛОГИЧЕСКИЕ АСПЕКТЫ

- Г.Б. Евгеньев 7-23  
ИНДУСТРИЯ 5.0 КАК ИНТЕГРАЦИЯ ИНТЕРНЕТА ЗНАНИЙ И ИНТЕРНЕТА ВЕЩЕЙ

- Ю.П. Похабов 24-35  
ПРОЕКТИРОВАНИЕ СЛОЖНЫХ ИЗДЕЛИЙ С МАЛОЙ ВЕРОЯТНОСТЬЮ  
ОТКАЗОВ В УСЛОВИЯХ ИНДУСТРИИ 4.0

### ПРИКЛАДНЫЕ ОНТОЛОГИИ ПРОЕКТИРОВАНИЯ

- А.А. Шарипбай, Ж.С. Сауханова, Г.Б. Шахметова, М.С. Сауханова 36-49  
ОНТОЛОГИЯ КОНЕЧНО-АВТОМАТНОЙ КРИПТОГРАФИИ

- А.А. Муромский, Н.П. Тучкова 50-69  
ПРЕДСТАВЛЕНИЕ МАТЕМАТИЧЕСКИХ ПОНЯТИЙ  
В ОНТОЛОГИИ НАУЧНЫХ ЗНАНИЙ

- Ю.В. Рогушина 70-84  
ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА  
ДЛЯ СОЗДАНИЯ СОВРЕМЕННЫХ ЭНЦИКЛОПЕДИЧЕСКИХ ПОРТАЛОВ

- Н.М. Боргест, А.С. Галахарь, М.В. Овсянников, Р.О. Самсонов 85-100  
ПРЕДПРОЕКТНЫЙ АНАЛИЗ ИНТЕЛЛЕКТУАЛЬНОГО ЖИЛОГО ДОМА  
ДЛЯ УСЛОВИЙ АРКТИКИ

### ИНЖИНИРИНГ ОНТОЛОГИЙ

- С.В. Лебедев 101-116  
МОДЕЛЬ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ПОСТРОЕНИЮ  
СВЯЗАННЫХ ДАННЫХ НА ОСНОВЕ РАЗНОРОДНЫХ ИСТОЧНИКОВ

- О.Н. Тушканова, В.В. Самойлов 117-131  
KNOWLEDGE NET: МОДЕЛЬ И СИСТЕМА НАКОПЛЕНИЯ,  
ПРЕДСТАВЛЕНИЯ И ИСПОЛЬЗОВАНИЯ ЗНАНИЙ И ДАННЫХ

- В.В. Голенков, Н.А. Гулякина, И.Т. Давыденко, Д.В. Шункевич, А.П. Еремеев 132-151  
ОНТОЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ ГИБРИДНЫХ СЕМАНТИЧЕСКИ  
СОВМЕСТИМЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ  
НА ОСНОВЕ СМЫСЛОВОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

### МЕТОДЫ И ТЕХНОЛОГИИ ПРИНЯТИЯ РЕШЕНИЙ

- С.М. Ковалев, А.Е. Колоденкова, В. Снасель 152-168  
ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ СЛИЯНИЯ ДАННЫХ  
ПРИ ДИАГНОСТИРОВАНИИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ



## CONTENT

### FROM THE EDITORS

- GOING BEYOND THE BOUNDARIES OF THE INDUSTRY 5-6

### GENERAL ISSUES OF FORMALIZATION IN THE DESIGNING: ONTOLOGICAL ASPECTS

- G.B. Evgenev 7-23  
INDUSTRY 5.0 AS AN INTEGRATION OF THE INTERNET OF KNOWLEDGE  
AND THE INTERNET OF THINGS

- Yu.P. Pokhabov 24-35  
DESIGNING COMPLEX PRODUCTS WITH SMALL PROBABILITY OF FAILURE  
IN THE CONTEXT OF INDUSTRY 4.0

### APPLIED ONTOLOGY OF DESIGNING

- A.A. Sharipbay, Zh.S. Saukhanova, G.B. Shakhmetova, M.S. Saukhanova 36-49  
ONTOLOGY OF FINITE-AUTOMATION CRYPTOGRAPHY

- A.A. Muromskiy, N.P. Tuchkova 50-69  
REPRESENTATION OF MATHEMATICAL CONCEPTS IN THE ONTOLOGY  
OF SCIENTIFIC KNOWLEDGE

- J.V. Rogushina 70-84  
USE OF ONTOLOGICAL ANALYSIS FOR CREATION OF ENCYCLOPEDIA PORTALS

- N.M. Borgest, A.S. Galakhar, M.V. Ovsyannikov, R.O. Samsonov 85-100  
PRE-PROJECT ANALYSIS OF AN INTELLIGENT RESIDENTIAL BUILDING  
FOR THE CONDITIONS OF THE ARCTIC

### ONTOLOGY ENGINEERING

- S.V. Lebedev 101-116  
MODEL-DRIVEN TRANSFORMATION OF HETEROGENEOUS SOURCES  
INTO LINKED DATA

- O.N. Tushkanova, V.V. Samoylov 117-131  
KNOWLEDGE NET: MODEL AND SYSTEM FOR ACCUMULATION,  
REPRESENTATION AND USE OF KNOWLEDGE AND DATA

- V.V. Golenkov, N.A. Guliakina, I.T. Davydenko, D.V. Shunkevich, A.P. Ereemeev 132-151  
ONTOLOGICAL DESIGN OF HYBRID SEMANTICALLY COMPATIBLE INTELLIGENT  
SYSTEMS BASED ON SENSE REPRESENTATION OF KNOWLEDGE

### METHODS AND TECHNOLOGIES OF DECISION MAKING

- S.M. Kovalev, A.E. Kolodenkova, V. Snasel 152-168  
INTELLECTUAL TECHNOLOGIES OF DATA FUSION  
FOR DIAGNOSTICS TECHNICAL OBJECTS



## ВЫХОД ЗА ГРАНИЦЫ ИНДУСТРИИ GOING BEYOND THE BOUNDARIES OF THE INDUSTRY

**Дорогой наш читатель,  
уважаемые авторы и члены редакционной коллегии!**

Цифровая трансформация промышленности и общества в целом активно меняет нашу жизнь. Японская модель *Общества 5.0* рассматривается прагматиками и футурологами как модель ближайшего будущего, в котором «искусственный интеллект» (ИИ) на равных будет сосуществовать и развиваться, становиться частью социума. Видимо ещё рано говорить о возможном «восстании машин», ведь ИИ - это пока лишь наш с вами не всегда до конца верифицированный, но уже формализованный опыт. Однако недавние трагедии, происшедшие с новейшими самолётами компании Боинг<sup>1</sup>, в которых пилот способен лишь наблюдать за ходом выполнения полета и в критический момент сбоя программы полёта не готов или не имеет возможности исправить ситуацию, - это сигнал для всех, кто занимается формализацией своих неполных знаний о предметной области.

О растущем влиянии ИИ говорится и в отчете, подготовленном в конце 2018 года в подкомитете по информационным технологиям Конгресса США. В 16-ти страничном документе, озаглавленном «Восстание машин: искусственный интеллект и его растущее влияние на политику США»<sup>2</sup>, говорится, что «ИИ имеет потенциал, чтобы разрушить слои общества как ожидаемым, так и непредвиденным образом». ИИ определяется как вычислительная технология, которая работает и реагирует по-человечески. ИИ, как правило, делится на две категории: «узкий» ИИ и «общий» ИИ. Узкий ИИ решает конкретные задачи, такие как стратегические игры, языковой перевод, автоматическое управление транспортом, распознавание изображений и др. Общий ИИ может выполнять задачи в различных предметных областях, обучаясь в ходе решения задач на основе рассуждений. Сегодня активно начал использоваться узкий ИИ, а общий ИИ пока недостаточно развит.

В глоссарии «Европейской этической хартии об использовании искусственного интеллекта в судебных системах и их окружении»<sup>3</sup> ИИ определяется как совокупность научных методов, теорий и техник, целью которых является воспроизведение с помощью машины когнитивных способностей человека. В этом юридическом документе по аналогии с американским документом различают «сильный» ИИ, способный контекстуализировать специализированные и разнообразные проблемы полностью автономно, и «слабый» ИИ, обладающий высокой эффективностью в своей узкой области обучения. Признавая важность ИИ в современных обществах и ожидаемые выгоды, Европейская этическая хартия официально приняла пять фундаментальных принципиальных положений об использовании ИИ в судебных системах и их окружении: уважение фундаментальных прав; недискриминация; качество и безопасность; прозрачность, беспристрастность и честность; контроль пользователем.

<sup>1</sup> Президент и генеральный директор компании Боинг Дэннис Муиленбург (*Dennis Muilenburg*) в своём обращении от 19 марта 2019 года к авиакомпаниям, пассажирам и авиационному сообществу заявил о *выпуске обновления программного обеспечения* и соответствующей подготовке пилотов для 737 MAX, которые будут решать проблемы, обнаруженные после авиакатастроф - <http://www.boeing.com/commercial/737max/737-max-update.page#/letter>.

<sup>2</sup> Hurd, Will, Kelly, Robin L. Rise of the Machines: Artificial Intelligence and Its Growing Impact on U.S. Policy. 2018-09. - <https://www.hsdl.org/?abstract&did=816362>.

<sup>3</sup> European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment adopted by the CEPEJ during its 31st Plenary meeting (Strasbourg, 3-4 December 2018) - <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

В феврале 2019 года журнал «Профиль» вышел с заголовком на обложке «Стремительное развитие ИИ пугает своей непредсказуемостью»<sup>4</sup>. В статьях, опубликованных в этом журнале, обсуждается программа цифровизации китайского общества, тотальный мониторинг граждан средствами ИИ, использование интеллектуальных камер и программ оценки общественной полезности граждан. Ещё в 2017 году Государственный совет КНР опубликовал «План развития ИИ следующего поколения», в котором изложена стратегия Китая стать мировым лидером и к 2020 году догнать США в области разработки технологий и приложений, где используется ИИ.

В ежегодном отчёте Стэнфордского университета<sup>5</sup> отмечается, что лидерство Европы и США в издании статей по ИИ уже не столь очевидно из-за бурного роста публикаций в Китае. Замечено также, что частные компании играют центральную роль в разработке и инвестировании ИИ в США. В 2017 году частные технологические компании, такие как Amazon и Alphabet, инвестировали 16,1 и 13,9 млрд долл. США соответственно в исследования и разработки с ИИ. Для сравнения: общий бюджет Национального научного фонда США вместе с инвестициями DARPA и DOT в разработку автономных и беспилотных систем на 2019 год составляет 5,3 млрд долларов.

В России также создаётся своя национальная стратегия в области ИИ, о которой объявил Президент РФ<sup>6</sup> в феврале 2019 года. Правительству РФ поручено разработать её до 15 июня 2019 года, а к 1 июля 2019 года осуществить реализацию дополнительных мер, направленных на стимулирование роста инвестиций в высокотехнологичные проекты в области ИИ, Интернета вещей, робототехники и обработки больших массивов данных, осуществляемые субъектами малого и среднего предпринимательства.

Тема цифровой экономики уже затрагивалась в редакционной статье во втором номере нашего журнала за 2018 год<sup>7</sup>. Модный тренд подхвачен и организаторами XXI-ой международной научной конференции «Проблемы управления и моделирования в сложных системах» (ПУМСС-2019), которая состоится 3-6 сентября 2019 года в Самаре. Первым пунктом в направлениях работы конференции ПУМСС-2019 обозначена тема «Industry 5.0 и Super Smart Society»<sup>8</sup>.

Этот номер журнала открывают статьи, которые уже в своих названиях используют термины Industry 5.0 (Г.Б. Евгеньев) и Industry 4.0 (Ю.П. Похабов). Выйти за границы Индустрии стремятся авторы и других статей, представляя результаты онтологического моделирования криптографии (А.А. Шарипбай и др.), научных знаний (А.А. Муромский, Н.П. Тучкова), энциклопедических порталов (Ю.В. Рогушина), интеллектуальных систем (В.В. Голенков, и др.), а также интеллектуальные технологии слияния (С.М. Ковалев и др.), связывания (С.В. Лебедев) и накопления (О.Н. Тушканова, В.В. Самойлов) данных.

*Уважаемый читатель, приятного Вам прочтения!*

А статья автором нашего журнала - очень просто для тех, у кого есть *новые результаты* в области формализации знаний на основе онтологического моделирования предметных областей и процессов в них!

*Ontologists and designers of all countries and subject areas, join us!*

<sup>4</sup> См., например, статьи Ивана Дмитриенко «Распознавай и властвуй: развитие нейросетей ведёт наш мир к невиданному до сих пор уровню тоталитарного контроля над обществом», и «Капкан из нейросетей: доверяя свою судьбу искусственному интеллекту, человечество сильно рискует» в еженедельном деловом журнале «Профиль» №5(69) 11.02.2019. – с.15-26.

<sup>5</sup> Yoav Shoham, Raymond Perrault, Erik Brynjolfsson, Jack Clark, James Manyika, Juan Carlos Nieves, Terah Lyons, John Etchemendy, Barbara Grosz and Zoe Bauer, "The AI Index 2018 Annual Report", AI Index Steering Committee, Human-Centered AI Initiative, Stanford University, Stanford, CA, December 2018.

<sup>6</sup> Перечень поручений по реализации послания Президента Федеральному Собранию (27 февраля 2019 г. ПР-294).

<sup>7</sup> Модный тренд: цифровая экономика. *Онтология проектирования*. №2(28), 2018. – с.173-174.

<sup>8</sup> Complex systems: control and modelling problems. XXI International Scientific Conference. - <http://www.cscmp.ru/>.

УДК 621

## ИНДУСТРИЯ 5.0 КАК ИНТЕГРАЦИЯ ИНТЕРНЕТА ЗНАНИЙ И ИНТЕРНЕТА ВЕЩЕЙ

Г.Б. Евгениев

Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия  
g.evgeniev@mail.ru

### Аннотация

Предложена методология создания систем класса «Индустрия 5.0» с использованием технологий искусственного интеллекта. Методология основана на многоагентных методах создания баз знаний и пригодна для разработки систем проектирования и управления для цифровых интеллектуальных производств. Разработана интегрированная структура Интернета знаний и Интернета вещей. Проанализирован жизненный цикл изделий машиностроения и предложены методы применения Интернета знаний и Интернета вещей на различных этапах этого цикла. Приведена функциональная декомпозиция основных этапов жизненного цикла. Даны концептуальные основы Интернета знаний. Разработаны многоагентные методы создания баз знаний. Предложена метаонтология инженерных агентов. Описаны принципы построения многоагентных систем полуавтоматического проектирования изделий. Приведено описание возможностей интеллектуальных систем программирования обработки на оборудовании с ЧПУ в части формирования траектории и областей переходов. Описаны возможности интеллектуальных систем проектирования и нормирования технологических процессов. Предложено использование стандарта IDEF3 для создания метамodelей технологических процессов и модифицированных маршрутных карт для формирования баз знаний. Дано описание интеллектуальной системы оперативного управления машиностроительным производством.

**Ключевые слова:** Индустрия 4.0, Индустрия 5.0, цифровые производства, Интернет знаний, Интернет вещей, интеллектуальные системы.

**Цитирование:** Евгениев, Г.Б. Индустрия 5.0 как интеграция Интернета знаний и Интернета вещей / Г.Б. Евгениев // Онтология проектирования. – 2019. – Т.9, №1(31). – С.7-23. – DOI: 10.18287/2223-9537-2019-9-1-7-23.

### Введение

В настоящее время происходит четвёртая промышленная революция (4ПР), для обозначения которой используется термин Индустрия 4.0. [1-5]. В развитии производительных сил человечества принято выделять ряд этапов. Этапы 1.0 и 2.0 были связаны с развитием сельского хозяйства. Третья ступень – это индустриальная революция, начало массового использования промышленного производства. Этап 4.0 охватывает период всеобщей компьютеризации, включая технологическое оборудование и этапы изготовления изделий. Суть новой революции заключается в том, что материальный мир соединяется с виртуальным, в результате чего рождаются новые киберфизические комплексы, объединённые в одну цифровую экосистему. Этап 5.0 должен охватить все стадии жизненного цикла (ЖЦ) изделий.

В статье [6] приводится интервью с представителем японской компании о развитии цифрового общества. Отмечено, что Индустрия 4.0 – лишь часть процесса с переходом к Обществу 5.0, которое представляет собой ступень, следующую за информационным обществом [7]. Общество 5.0 представляет собой оптимизацию ресурсов социума в целом через интеграцию физического и киберпространства.

В материале [8] рассматривается интеграционная концепция Индустрии 5.0 по трём основным направлениям: проектное управление и цифровая экономика, развитие личности и компетенций, промышленный и домашний Интернет. Индустрия 5.0 описывается как информационная система, обеспечивая планирование, регистрацию, переработку и отражение информации как минимум по трём измерениям элементов влияния: в сферах, поддерживающих и создающих продукт, в стадиях осуществления деятельности и в составе элементов управленческой деятельности.

Фундаментом для 4ПР является Интернет вещей (*Internet of Things, IoT*) — концепция вычислительной сети физических предметов («вещей»), оснащённых встроенными технологиями для взаимодействия друг с другом или с внешней средой, рассматривающая организацию таких сетей как явление, способное перестроить экономические и общественные процессы, исключаящее из части действий и операций необходимость участия человека [8]. 4ПР должна охватывать автоматизацию абсолютно всех этапов и процессов, включая цифровое проектирование изделия, создание его виртуальной копии, совместную работу инженеров-конструкторов и технологов в едином цифровом пространстве и т.д., когда изделия ещё не являются вещами, а существуют в виртуальном мире в виде информационных моделей.

Цифровая революция должна дать возможность непрограммирующему носителю знаний вводить их в компьютер без посредников. Это стало возможным, используя методологию *экспертного программирования* [9]. В этой методологии знания описываются на языке *деловой прозы*, максимально приближенном к литературному языку, но формализованному настолько, что имеется возможность автоматической генерации программных средств, соответствующих исходным текстам.

Интегрирование промышленной и цифровой революций необходимо рассматривать в совокупности двух миров: мира виртуального, реализуемого Интернетом знаний (*Internet of Knowledge, IoK*), и мира реального, реализуемого Интернетом вещей. Интернет знаний строится на онтологической основе, коренным объектом которой является метаонтология. С точки зрения искусственного интеллекта (ИИ), *онтология* — это эксплицитная (явная) спецификация концептуализации знаний. Метаонтология оперирует общими концептами и отношениями, которые не зависят от конкретной предметной области. Метаонтология должна содержать концепты и отношения, необходимые как для предметной онтологии, так и для онтологий задач и оптимизации. Метаонтология включает в себя три компоненты: предметную онтологию, онтологию задач и онтологию оптимизации (рисунок 1).

Предметная онтология состоит из иерархии понятий, их определений и атрибутов, а также связанных с ними аксиом и правил вывода. Предметная онтология на основе использования задач обеспечивает на основе технических заданий генерацию 3D моделей изделий, которые адекватно представляют изделия в виртуальном мире и удовлетворяют требованиям заданий. Онтология задач включает задачи структурного и параметрического синтеза моделей изделий и процессов. С её помощью обеспечивается генерация цифровых моделей процессов и производства [10].

Третью компоненту метаонтологии составляет онтология оптимизации, включающая компоненты однокритериальной и многокритериальной оптимизации.

В описанном подходе мир вещей состоит из изделий, изготовленных с помощью 3D принтеров или станков и роботов, составляющих оборудование в составе средств технологического оснащения конкретного производства и имеющейся в нём оснастки. Процесс изготовления регламентируется цифровыми моделями, сгенерированными с помощью соответствующих систем, включающих необходимые задачи.

Проектирование процессов производится с использованием данных о состоянии производства, получаемых из Интернета вещей.



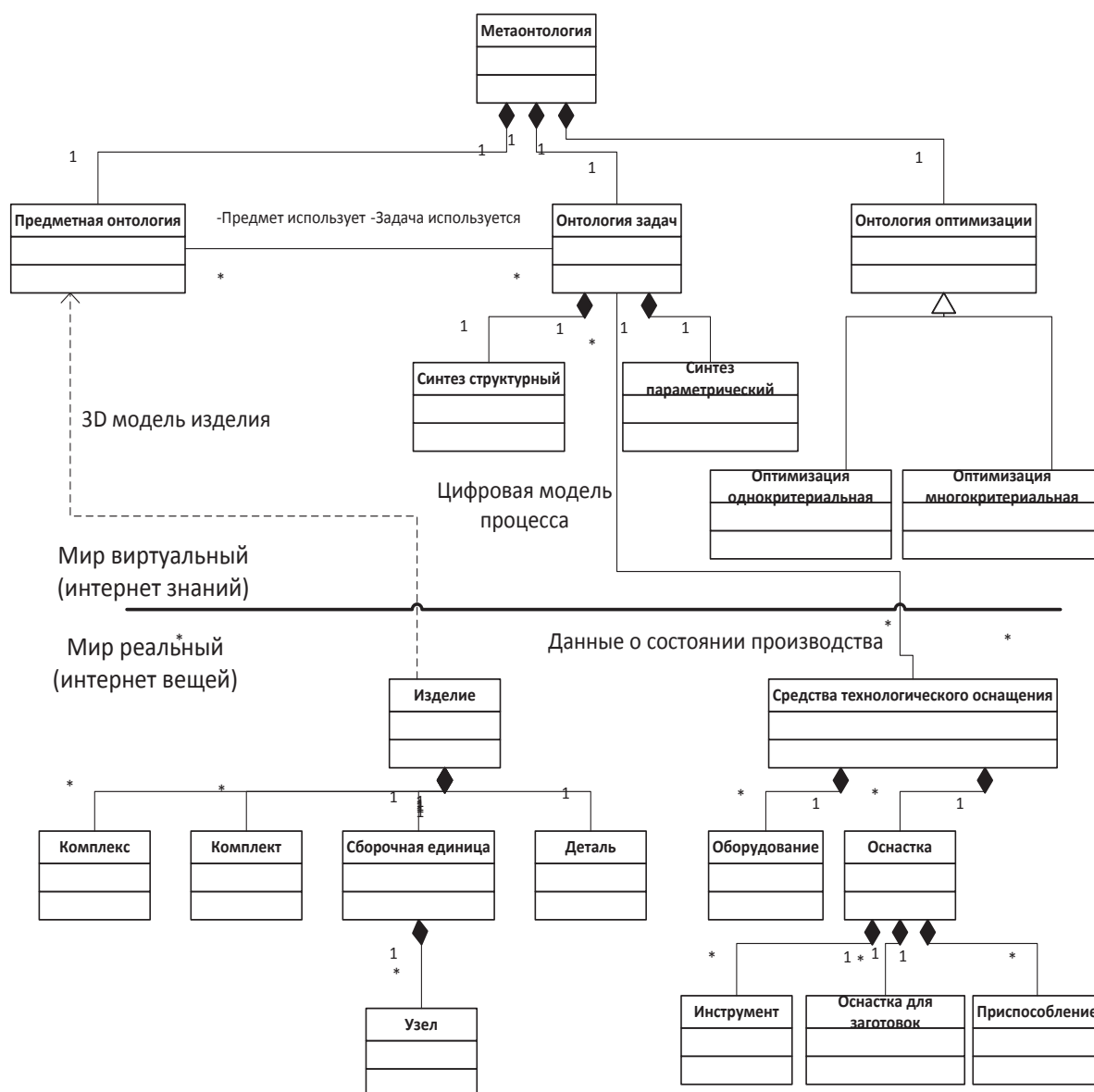


Рисунок 1 - Интегрированная структура Интернета знаний и Интернета вещей

## 1 ЖЦ изделий машиностроения

На рисунке 2 приведён ЖЦ изделий и интегрированная структура Интернета знаний и Интернета вещей. ЖЦ изделия содержит этапы: проектирование, производство, эксплуатация и утилизация. На этапах конструкторского и технологического проектирования изделия и процессы существуют в виртуальном мире, в связи с чем применяется Интернет знаний. В процессе производства изделия возникают как вещи, поэтому появляется возможность совместного использования Интернета вещей и Интернета знаний.

На рисунке 3 раскрыт этап проектирования, который начинается с маркетинга. В качестве выходов этого блока в соответствии с двумя концепциями маркетинга отмечены «техническое задание (совершенствование товара)» и «совершенствование производства». Техническое задание является входом блока «Проектирование изделий», а совершенствование производства – входом блока «Проектирование технологических процессов и оснастки».

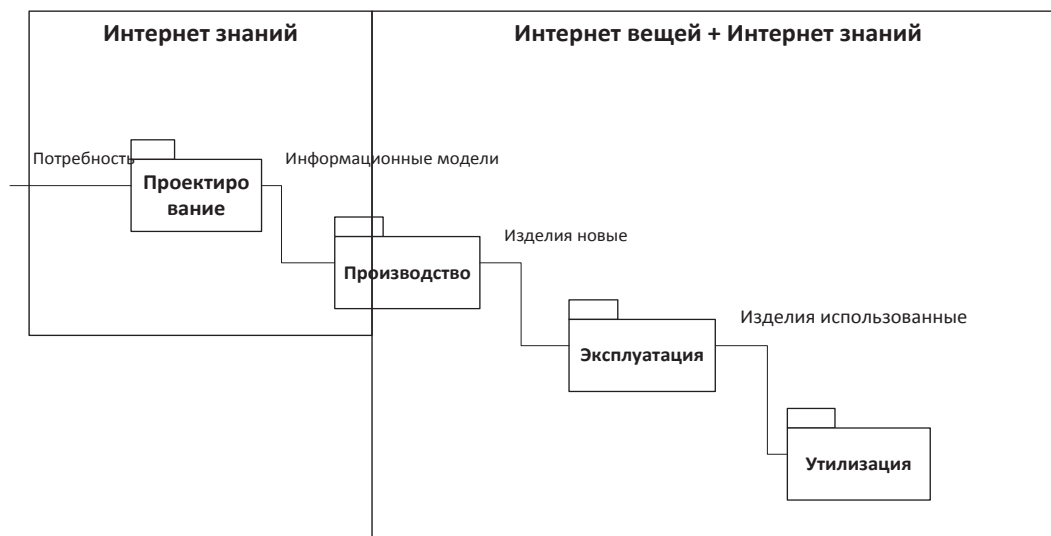


Рисунок 2 - ЖЦ изделий и интегрированная структура Интернета знаний и Интернета вещей

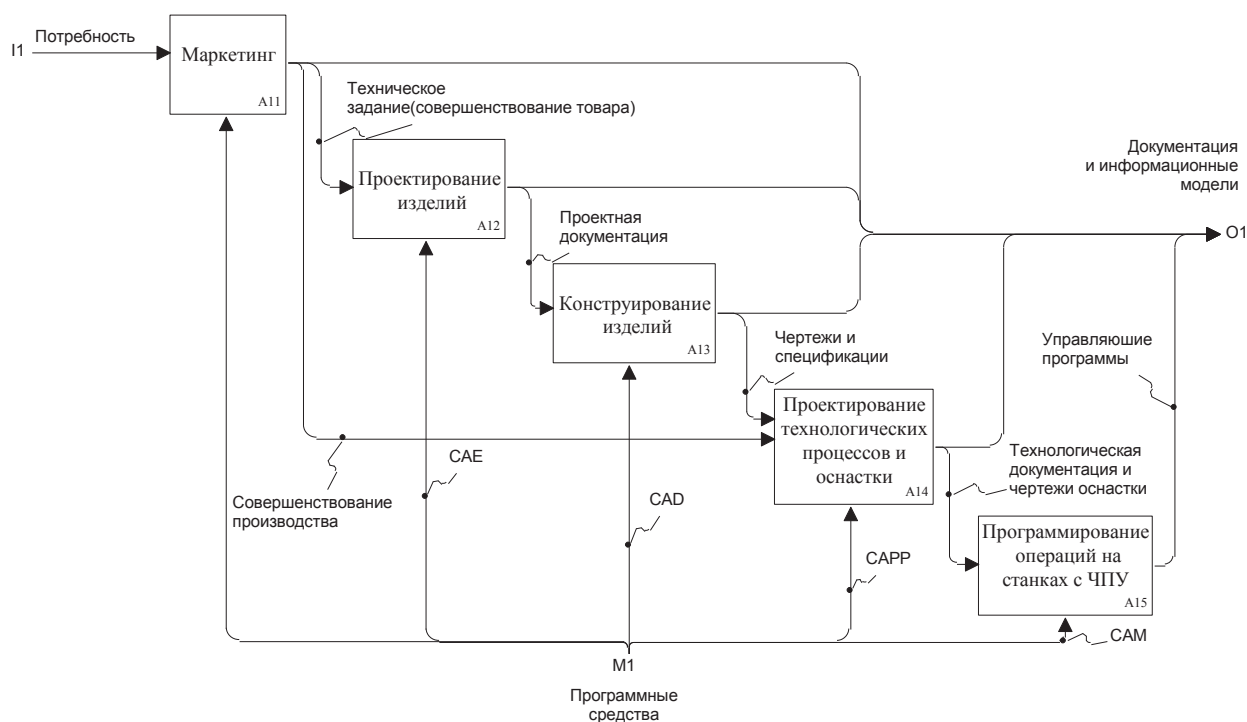


Рисунок 3 - Декомпозиция этапа проектирования

На этапе проектирования изделие представляется как формальная система с оформлением соответствующей проектной документации [10, 11].

Описанные функции составляют содержание инженерной подготовки производства. Декомпозиция этапа «Производство» (рисунок 2) приведена на рисунке 4. Первым функциональным блоком этой декомпозиции является «Планирование ресурсов предприятия». Средством реализации этого блока являются системы класса ERP (*Enterprise Resource Planning*).

Данные о приобретённых материалах и комплектующих совместно с документацией и информационными моделями изделий передаются на вход функционального блока «Планирование и диспетчирование производства», средством реализации которого являются системы класса MES (от *Manufacturing Execution System*, производственная исполнительная си-



стема) — специализированное прикладное программное обеспечение, предназначенное для решения задач синхронизации, координации, анализа и оптимизации выпуска продукции в рамках какого-либо производства.

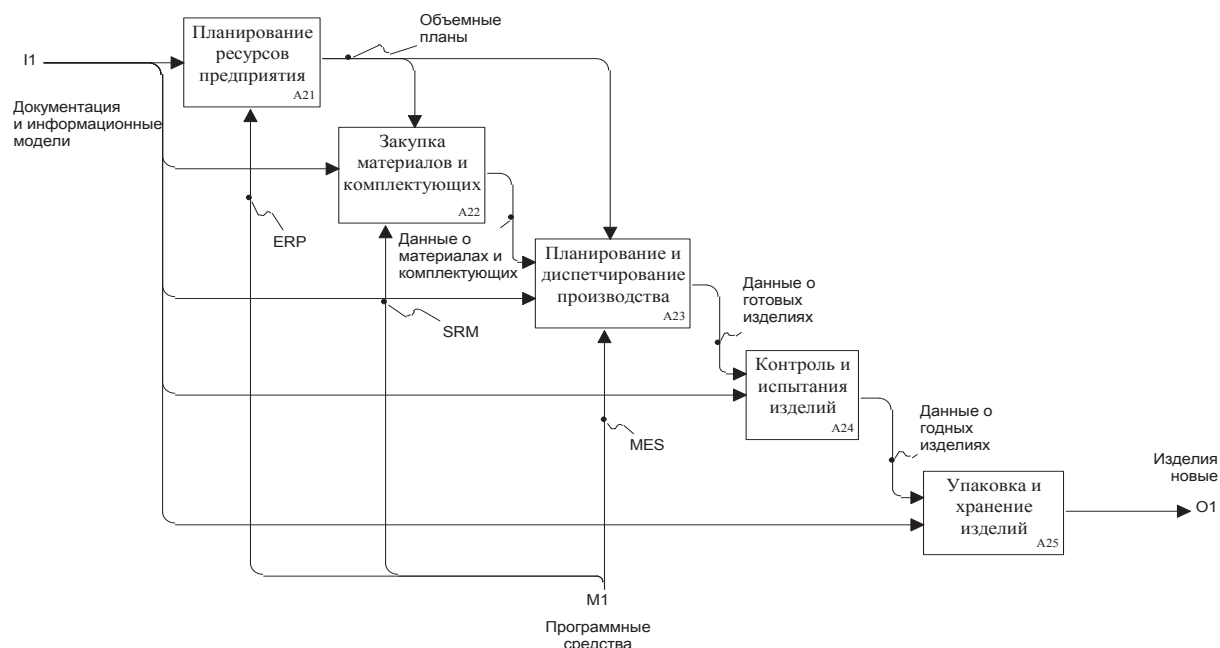


Рисунок 4 - Декомпозиция этапа производство

## 2 Концептуальные основы Интернета знаний

Семантика Интернета знаний определяется метаонтологией (см. рисунок 5). Коренным метаобъектом в этом случае является «База знаний» (БЗ), которая имеет своё наименование, имя-идентификатор и версию. Методы, подключённые к БЗ, позволяют производить сортировку и поиск модулей знаний, из которых она состоит. Подмножество модулей знаний, объединённое в семантическую сеть, является методом, который можно экспортировать в другие БЗ, а также импортировать и производить слияние в единую семантическую сеть.

Согласно технологии вики (Wiki) БЗ представляет собой страницу. При этом БЗ может рассматриваться как модуль в составе интегрированной БЗ, которая в этом случае состоит из набора страниц. Структурными компонентами БЗ являются словарь БЗ и множество модулей. Словарь имеет своё имя-идентификатор и состоит из слов. Имеется возможность сортировки и поиска слов, а также импорт слов из текстовых документов. Слова имеют имя-идентификатор и содержательное наименование на одном из языков мира, а также тип слова из наборов чисел и символов. Слова могут иметь ассоциативные списки, содержащие допустимые наименования слов. Ассоциативный список имеет имя-идентификатор, наименование и тип значений из числа названных. Применительно к спискам имеются возможности добавления, удаления, сортировки значений и поиска списка. Возможность замены словаря БЗ позволяет обеспечить выполнение концептуального требования вики-систем: использование любых языков мира.

Вторым структурным компонентом БЗ является множество модулей знаний. Модули знаний выполняют основное функциональное назначение БЗ – преобразование текущего состояния данных с целью получения новых объектов, удовлетворяющих целям проектирова-

ния. Каждый модуль имеет наименование, а также имена-идентификаторы самого модуля, предусловия его исполнения и версию.

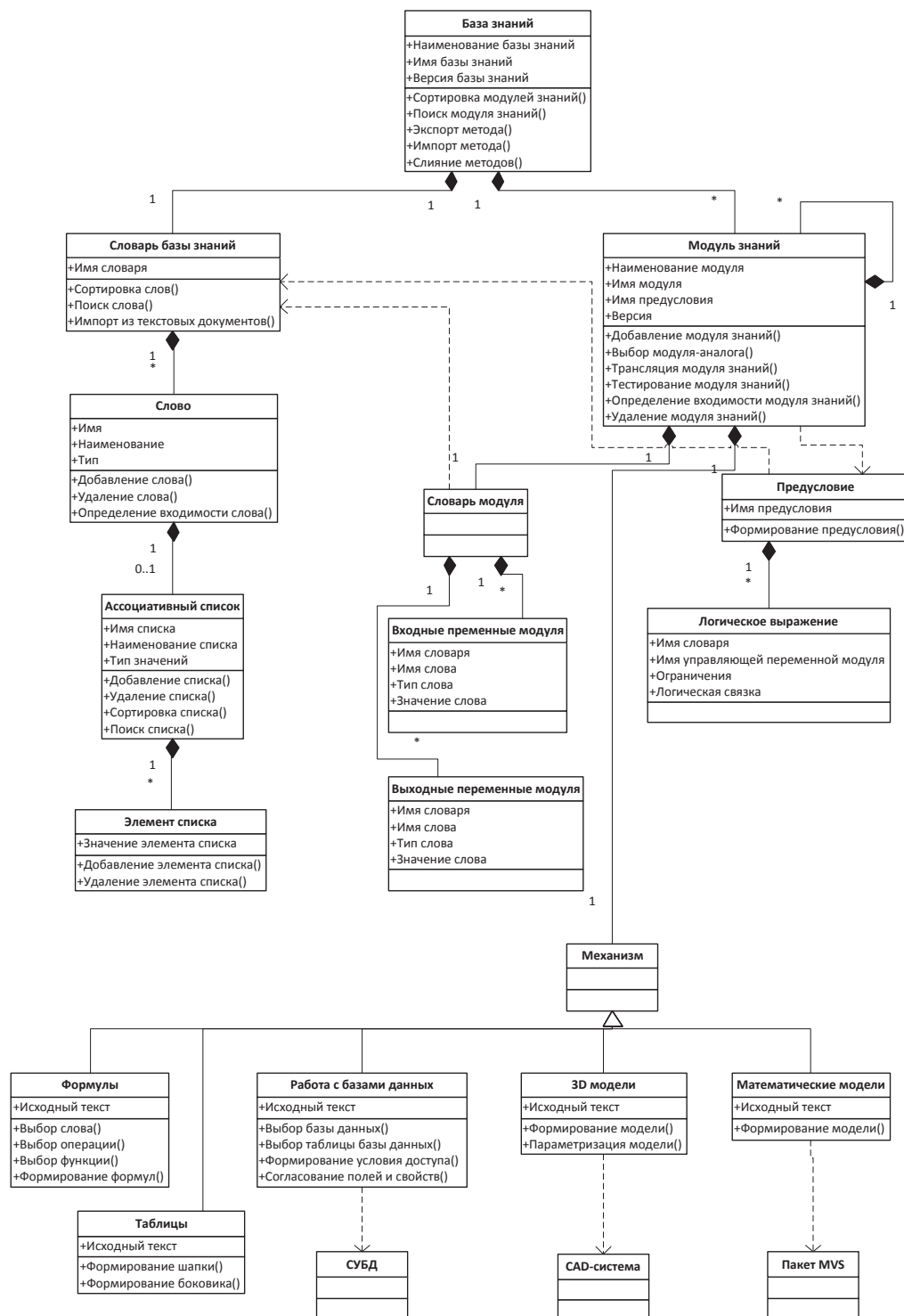


Рисунок 5 - Метаонтология Интернета знаний

В БЗ имеются возможности добавления модулей, выбора модуля-аналога, трансляции модуля знаний на один из языков программирования, тестирования полученного результата, а также определения входимости модуля в БЗ и удаления выбранного модуля. Эти действия могут выполняться любым оператором, что позволяет реализовать вторую концептуальную

основу вики-систем: обеспечение возможности независимого пополнения и корректирования содержания системы. Каждый модуль знаний имеет свой словарь, являющийся подмножеством словаря БЗ. В связи с тем, что модули представляют собой объект-функции, словарь имеет два подмножества: входные и выходные переменные модуля. Модули знаний могут иметь предусловия, определяющие возможности его выполнения. Предусловия представляют собой логические выражения, составленные из слов словаря БЗ, ограничений значения переменных и логических связей.

### 3 Многоагентные методы создания БЗ

Важнейшей методической основой для реализации концепций *интеграция, интеллектуализация и индивидуализация* является теория многоагентных систем (МАС).

Рассмотрим применение МАС на примере интеллектуальных систем (ИС) проектирования и управления [9]. В машиностроительных системах в качестве агентов выступают комплексы, сборочные единицы, детали и их элементы, а также технологические процессы (ТП) их изготовления. Активность агентов заключается в необходимости решения двух категорий задач: структурного синтеза и синтеза параметрического. Структурный синтез заключается в выборе структуры подчинённых объектов, а параметрический – в генерации значений собственных свойств, в результате чего из класса объектов, представленных в форме агента, генерируется один экземпляр, который и включается в проект.

Реактивность агентов обеспечивает решение упомянутых задач за счёт обмена информацией между агентами непосредственно или через БЗ [12]. Автономность агентов основывается на встроенных в них методах, в которых содержатся инженерные знания по различного рода расчётам, а также геометрические и графические знания в форме параметризованных моделей, обеспечивающих генерацию трёхмерных образов и чертежей. Общительность агентов имеет как вертикальную, так и горизонтальную составляющие. Вертикальная составляющая включает обмен данными по иерархии «целое-часть» и «род-вид», а горизонтальная – обмен между конструктивно сопряжёнными, но не подчинёнными друг другу по иерархии агентами. Целенаправленность агентов определяется необходимостью реализации проекта, удовлетворяющего техническим требованиям заказчика, а также другим требованиям, накладываемым разработчиком.

Обобщённая модель класса инженерных агентов приведена на рисунке 6. Любой агент представляет собой открытую систему, помещённую в некоторую среду. В случае САПР этой средой является проект, формируемый в базах данных (БД), в качестве которых целесообразно использовать БД объектного типа для представления модели изделия (*внутренняя среда*), и реляционную БД для поиска стандартных и покупных изделий, свойств материалов и т.п. информации (*внешняя среда*). Внешняя среда, как правило, является сетевой.

Свойства агента могут принадлежать трём различным категориям: импортируемые, экспортируемые и внутренние. Импортируемые свойства являются *рецепторами* агента, формирующими его систему *восприятия*. Экспортируемые свойства агента являются его *эффекторами*, функция которых состоит в воздействии на среду, то есть на состояние проекта. Свойства всех трёх категорий агента образуют его *память*, в которой хранится текущее состояние агента.

*Процессор* агента формируют его методы, обеспечивающие объединение и переработку разнородных данных, выработку соответствующих реакций на информацию о состоянии среды (проекта), принятие решений о выполнении тех или иных действий. В целом процессор определяет *поведение* агента, которое можно наблюдать, например, в графическом окне, в котором отображаются сгенерированные чертежи и другая геометрическая информация.

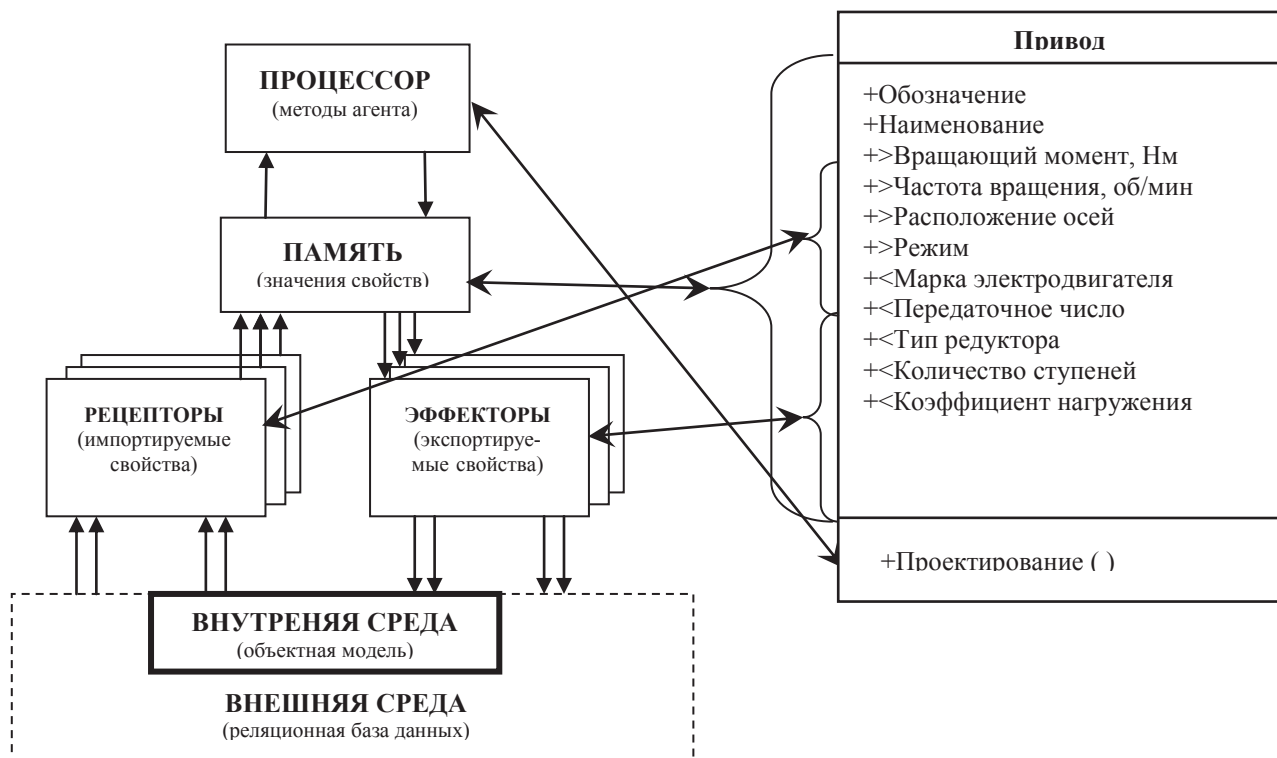


Рисунок 6 – Обобщённая модель класса инженерных агентов

При создании ИС первой фазой проектирования является формирование номенклатуры и оргструктуры агентов. Номенклатура агентов определена в словаре системы. Формирование оргструктуры агентов осуществляется в форме построения диаграммы классов UML. Применительно к созданию программных средств диаграммы классов эквивалентны конструкторским сборочным чертежам.

На рисунке 7 приведена методология инженерных агентов проектирования электромеханических приводов и технологических процессов их изготовления. Механизмы реализации представлены в виде методов, являющихся БЗ по расчёту, проектированию и конструированию изделия и его узлов, а также проектированию и нормированию ТП. Управление обеспечивается предусловиями. Концептуально искусственный агент представляет собой объект-функцию. Модель функционального блока, связанного с агентом, определена стандартом IDEF0. Важной частью объектного подхода является то, что объектно-ориентированное программирование использует в качестве основных логических конструктивных элементов объекты, а не алгоритмы. Именно на это направлена многоагентная методология.

#### 4 МАС полуавтоматического проектирования изделий

Имеются два основных способа использования технологий ИИ при автоматизации проектирования: встраивание и надстраивание. Первый подход используется в ряде CAD-систем, например SolidWorks [13]. Недостатком такого подхода является то, что инженеры, не занимающиеся конструированием, например технологи, не могут использовать интеллектуальные возможности таких систем. В статье предлагается использовать подход надстраивания, который позволяет применять методы ИИ на всех этапах ЖЦ изделий.

МАС обладают статическими и динамическими свойствами. В UML для представления статического и динамического вида системы используются различные диаграммы. Основной статической диаграммой UML является диаграмма классов. Представление классов объектов

в ней не обладает функциональностью, необходимой для описания МАС. Во-первых, не предусмотрено деление свойств на входные и выходные, что необходимо для представления объект-функций. Во-вторых, классы объектов не имеют предусловий, необходимых для управления процессом поведения системы. Эти проблемы можно решить, не выходя за рамки средств, принятых для представления диаграмм классов.

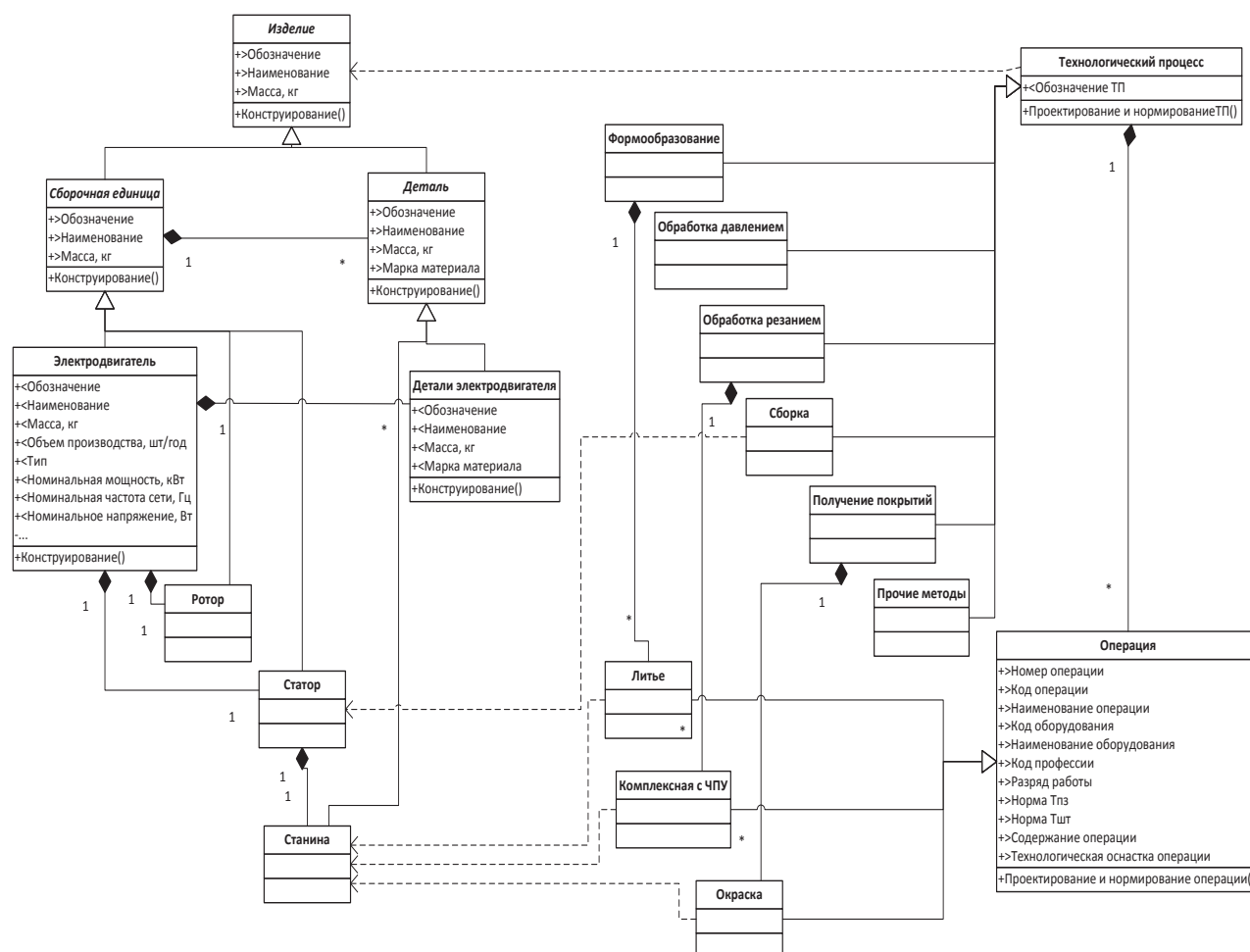


Рисунок 7 - Метаонтология инженерных агентов проектирования и изготовления электромеханических приводов

На рисунке 8 представлена модифицированная диаграмма классов для МАС. На ней на примере описания редуктора входящие свойства редуктора обозначены символом  $>$ , а выходящие, получаемые в результате проектирования, – символом  $<$ . Предусловие для узла вала промежуточного представлены в виде ограничения, в котором записано, что этот узел появляется при передаточном числе более 6.3.

Что касается динамических свойств, то методы классов объектов представляются в виде диаграмм состояний, в которых предусловия включены внутрь модулей.

На рисунке 9 приведена схема функционирования полуавтоматической системы в процессе проектирования одноступенчатого редуктора.

На входе в систему имеется техническое задание (ТЗ), которое формируется с использованием соответствующего окна. Имеется возможность формирования ТЗ разного состава. На рисунке 9 приведён вариант ТЗ, включающего величины момента вращающего на выходном валу в  $H\cdot m$ , частоты вращения на выходе в  $об/мин$ , ресурса передачи в  $часах$  и количества изделий.

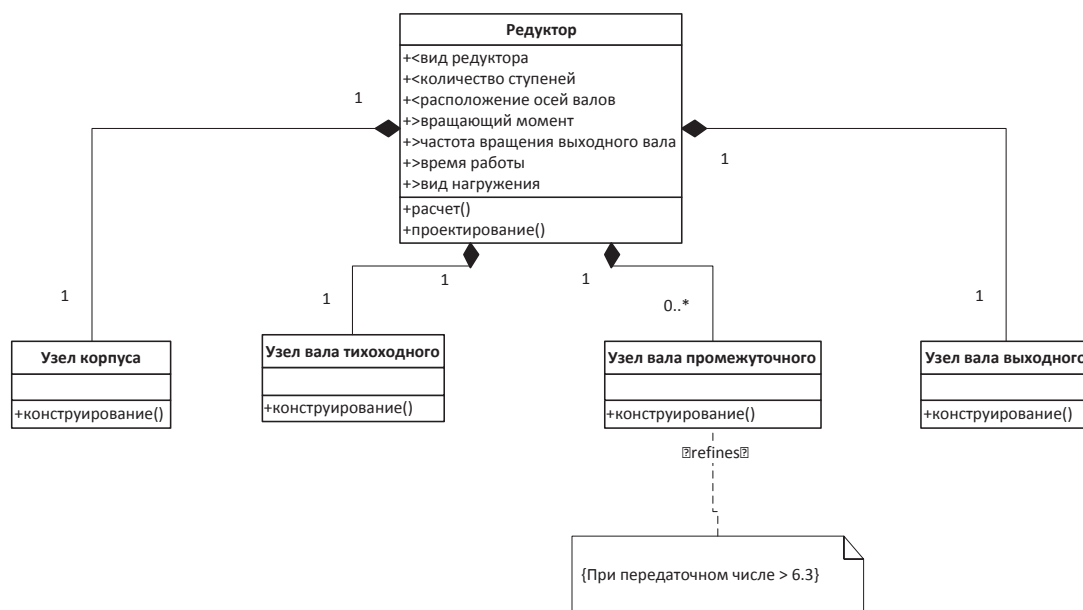


Рисунок 8 - Модифицированная диаграмма классов для МАС

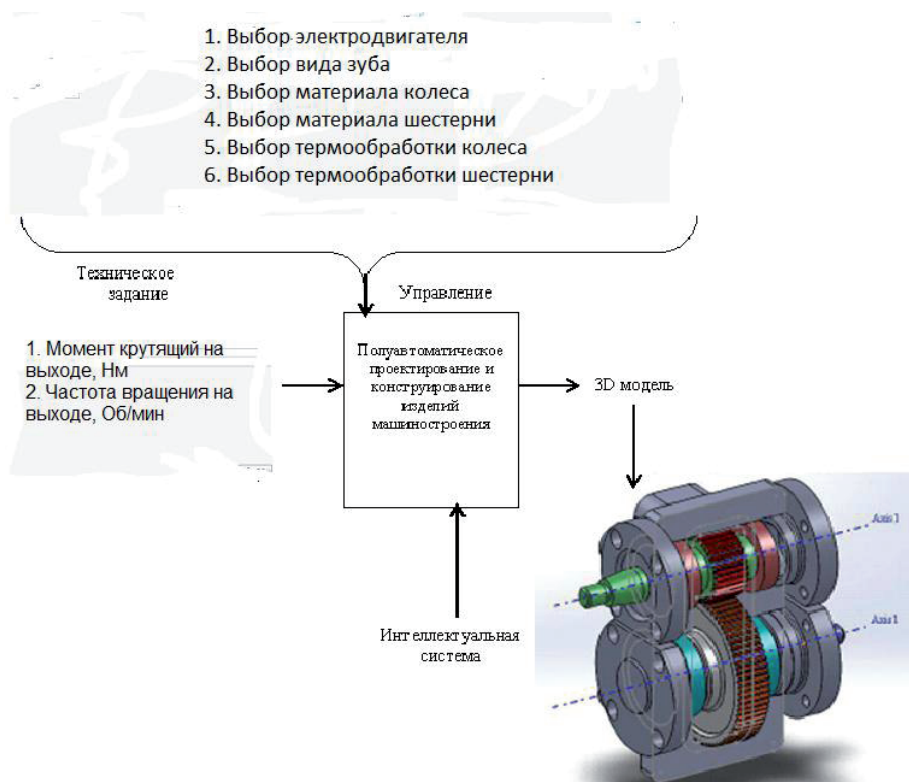


Рисунок 9 - Схема функционирования полуавтоматической системы

Управление процессом проектирования осуществляется с помощью ответов на вопросы с заданной номенклатурой вариантов значений, выводимых БЗ. В окне первого вопроса (порядок на рисунке слева направо и сверху вниз) разработчику представляется возможность выбора электродвигателя для привода. Приведено подмножество электродвигателей из БД, способных обеспечить выполнение ТЗ по мощности и частоте вращения. Во втором вопросе предлагается выбрать вид зуба колеса (прямой, косой или шевронный). Следующие вопросы



связаны с выбором термообработки и материалов шестерни и колеса. Два вопроса позволяют уточнить значения свойств материалов шестерни и колеса.

После ответов на все вопросы ИС производит все необходимые расчёты и генерацию документации и 3D модели. Если в результате ответов на вопросы нарушаются какие-либо ограничения, встроенные в БЗ, ИС выдаёт отказ с объяснением причин.

## 5 ИС программирования обработки на оборудовании с ЧПУ

Одной из наиболее эффективных систем автоматизации программирования на оборудовании с ЧПУ является отечественная система SprutCAM [14,15], обладающая интеллектуальными возможностями. Здесь рассматриваются возможности этой системы применительно к генерации траекторий и формированию зон фрезерной обработки детали.

Чистовая построчная оптимизированная обработка состоит из двух чистовых построчных траекторий, рабочие ходы которых лежат во взаимно перпендикулярных плоскостях (см. пример на рисунке 10). При этом каждая траектория обрабатывает только те зоны, в которых фронтальный угол к поверхности в направлении перемещения инструмента меньше или равен 45 градусам. Это гарантирует, что одни и те же участки не будут обработаны дважды. Кроме того, при такой стратегии достигается равномерная высота гребешка по всей поверхности детали. В результате улучшаются условия резания и уменьшается машинное время.

В задании на обработку для чистовой построчной оптимизированной операции указывается набор твёрдых тел, поверхностей и сеточных объектов. Для каждого геометрического объекта или группы объектов может быть установлен дополнительный припуск, который при обработке будет добавлен к основному припуску операции. Если заготовка и ограничивающая модель не заданы, то производится обработка всей доступной поверхности рабочего задания. Иначе обрабатываются лишь участки поверхности, лежащие внутри заготовки и вне ограничивающей модели.

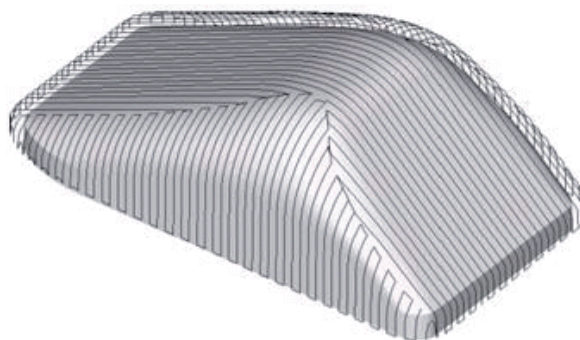


Рисунок 10 – Схема чистовой оптимизированной построчной обработки детали

Зоны обработки предназначены для выделения на детали или заготовке зон, подлежащих обработке в черновых и чистовых фрезерных переходах. Зоны обработки могут быть замкнутыми и открытыми. Задание открытых зон – новый эффективный инструмент, позволяющий технолог легко выделять на заготовке области обработки, которые без такой возможности приходилось бы строить вручную в 2D редакторе.

В системе имеется возможность создания переходов по уже выбранной геометрии. Суть этого метода заключается в том, что технолог сначала выбирает геометрию, которую планирует обработать, в графической области, и затем создаёт операцию. При этом выбранная геометрия автоматически попадает в рабочее задание созданной операции. Кроме того, при создании самой операции система анализирует рабочее задание, и пытается подобрать параметры операции исходя из параметров рабочего задания.



## 6 ИС проектирования и нормирования ТП

Основой для построения ИС являются онтологические концептуальные метамодели. Если для создания таких систем в области конструирования в качестве модели целесообразно использовать модифицированные диаграммы классов объектов UML, то применительно к системам проектирования ТП наиболее удобны модели стандарта IDEF3 [9].

Применение методики IDEF3 показано на примере формирования метамодели ТП обработки цилиндрических зубчатых колес (рисунок 11). Первым функциональным элементом метамодели ТП является блок заготовительных операций, который целесообразно декомпозировать отдельно в виде дочерней модели (рисунок 12).

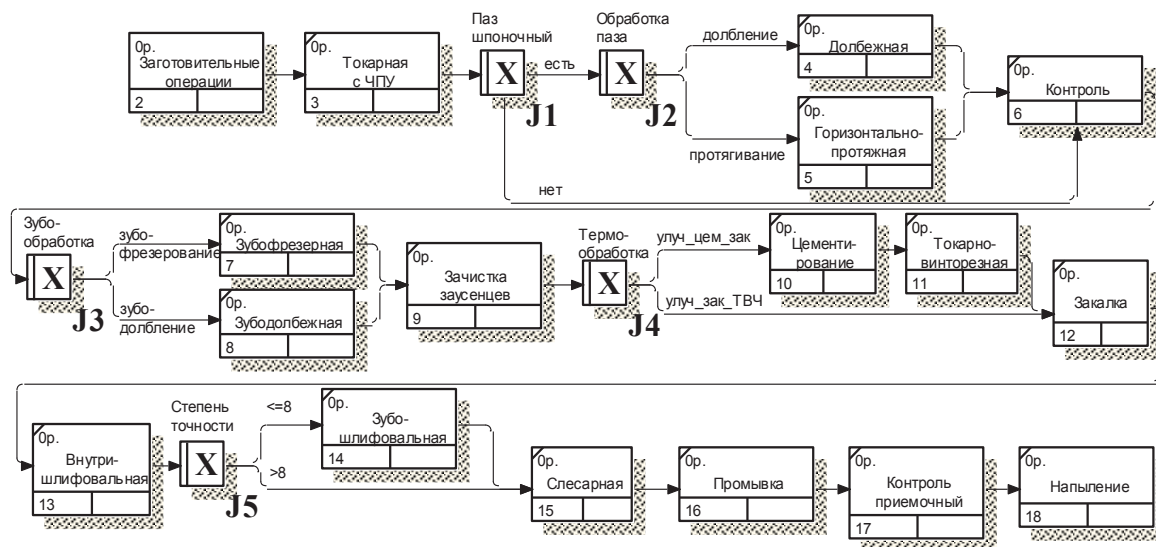


Рисунок 11 – Метамодель ТП обработки цилиндрических зубчатых колес

Эта модель начинается с перекрёстка типа исключающего *ИЛИ*. Такие перекрёстки наиболее часто используются при формировании метамодели ТП. Этот перекрёсток помимо формального идентификатора J6 имеет имя «Заготовка» и относится к типу разветвления стрелок. Каждая из отходящих от него стрелок имеет собственное имя: «литьё», «штамповка» и «круг». Таким образом, можно считать, что «Заготовка» является символьной переменной, принимающей одно из трёх перечисленных значений. В зависимости от этих значений выполняются различные заготовительные операции. Если переменная принимает значение «литьё», то выполняется соответствующая литейная операция. В случае значения «штамповка» выполняются операции «Пило-отрезная», посредством которой подготавливается заготовка для штамповки, и собственно операция «Штамповка».

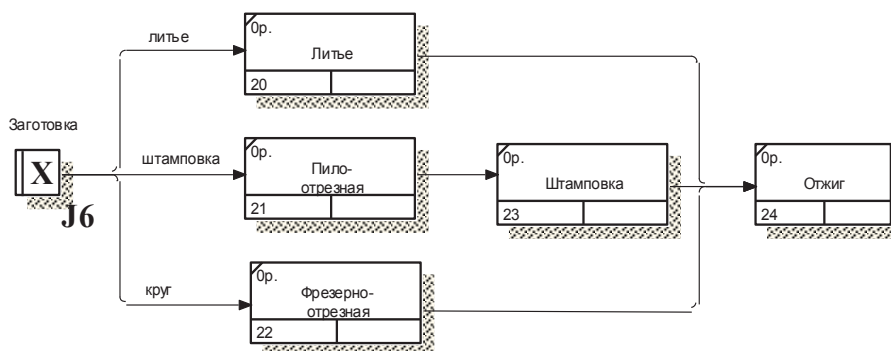


Рисунок 12 - Дочерняя модель заготовительных операций для цилиндрических зубчатых колес



## 7 ИС оперативного управления производством

На рисунке 14 представлена общая функциональная модель работы машиностроительного предприятия при позаказном производстве. На входе - заказы, а на выходе – готовая продукция с документами для отгрузки, а также отчётные бухгалтерские документы.

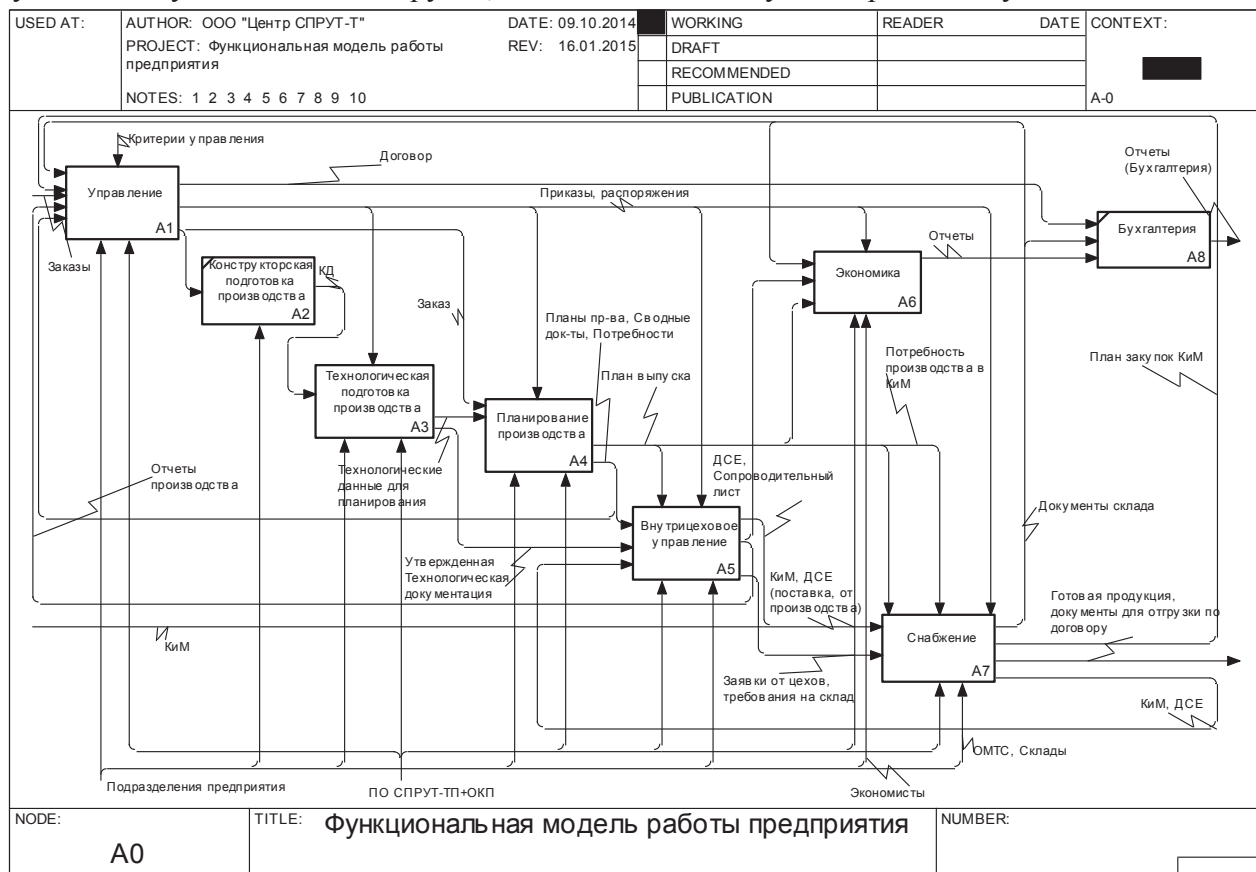


Рисунок 14 - Функциональная модель работы предприятия

Первым функциональным блоком модели является блок «Управление». Механизмами реализации всех функциональных блоков являются соответствующие подразделения предприятия, а также программные комплексы, в качестве которых здесь рассматриваются отечественные системы проектирования СПРУТ-ТП и оперативно-календарного планирования СПРУТ-ОКП [16]. Результатами работы блока «Управление» являются договоры, поступающие в блок «Бухгалтерия», приказы и распоряжения, управляющие работой последующих функциональных блоков, а также заказы, определяющие конструкторско-технологическую подготовку производства и его планирование. Помимо спецификации заказа на вход блока «Управление» поступают многочисленные обратные связи от последующих блоков, необходимые для принятия управленческих решений.

Вторым функциональным блоком на рисунке 14 является конструкторская подготовка производства.

Система СПРУТ-ТП обеспечивает автоматизированное нормирование на базе отечественных нормативов и работает совместно с системой оперативно-календарного планирования СПРУТ-ОКП. В результате работы СПРУТ-ТП формируются ТП и спецификации, на основе которых производится формирование сводных технологических ведомостей (трудоемкости, оборудования, оснастки и материалов) на данный заказ. Спроектированные и отнормированные в СПРУТ-ТП ТП передаются в систему оперативно-календарного планиро-

вания СПРУТ-ОКП. При формировании данных для передачи в СПРУТ-ОКП проверяется их корректность, и протокол ошибок возвращается на вход блока разработки и нормирования ТП. Планирование производства осуществляется на двух уровнях: предприятия и цеха. На уровне предприятия формируется план выпуска на год, квартал, месяц. На этой основе формируются планы-графики цехов с учётом межцеховых переходов. Описанные функциональные блоки входят в состав подсистемы СПРУТ-ОКП. Одним из функциональных блоков модели работы предприятия является блок «Снабжение».

Система СПРУТ-ОКП включает также блок «Экономика», в котором производится калькуляция по статьям затрат. На основе отчётов, поступающих с производства, формируются данные о браке деталей и сборочных единиц с указанием их количества, стоимости и причин брака. Эти отчёты используются для формирования нарядов на оплату. Экономическая информация позволяет системе СПРУТ-ОКП производить план-фактный анализ производства.

## Заключение

Фундаментом для 4ПР является Интернет вещей. Однако 4ПР должна охватывать автоматизацию всех этапов и процессов, когда изделия ещё не являются вещами, а существуют в виртуальном мире в виде информационных моделей.

В преддверии грядущей Индустрии 5.0 необходимо рассматривать в совокупности два мира: мир виртуальный, реализуемый Интернетом знаний, и мир реальный, реализуемый Интернетом вещей. Интернет знаний целесообразно строить на онтологической методологии, в основе которой лежит мир виртуальных агентов, представляющих объект-функции, способные генерировать новые данные на основе имеющейся информации. Попытка такой интеграции, охватывающей все этапы ЖЦ изделий, представлена в настоящей статье.

## Список источников

- [1] Industry 4.0: the fourth industrial revolution – guide to Industrie 4.0. - <https://www.i-scoop.eu/industry-4-0>.
- [2] Industry 4.0: the Future of Smart Manufacturing - Praim - <https://www.praim.com/Article>.
- [3] Siemens | Industrie 4.0 - <https://www.siemens.com/digital/enterprise>.
- [4] Digital Factory 4.0 | Industry 4.0 solution | antsolutions.eu - <http://www.antsolutions.eu>.
- [5] Industry 4.0 - the Nine Technologies Transforming Industrial Production. - <https://www.bcg.com/.../embracing-industry-4.0-rediscoveri>.
- [6] **Норицугу, У.** Общество 5.0: взгляд Mitsubishi Electric / У. Норицугу // Экономические стратегии. №4/2017. - С.2-11.
- [7] **Ржевский, Г.А.** Самоорганизация в социальных системах / Г.А. Ржевский // Онтология проектирования. – 2014. - №4(14). – С.8-17.
- [8] Концепция Индустрии 5.0 - <http://industry5.ru/koncept>.
- [9] **Евгениев, Г.Б.** Основы автоматизации технологических процессов и производств. Т.1: Информационные модели. Т.2: Методы проектирования и управления - Москва, Изд-во МГТУ им. Н.Э. Баумана, 2015. - Т.1: - 441 с. - Т.2: - 479 с.
- [10] **Гришин, М.В.** Онтологии проектирования шаблонной оснастки в авиационном производстве / М.В. Гришин, С.Н. Ларин, П.И. Соснин // Онтология проектирования. – 2016. – Т.6, №1(19). – С. 7-28. – DOI: 10.18287/2223-9537-2016-6-1-7-28.
- [11] **Андриченко, А.Н.** Тенденции и состояние в области управления справочными данными в машиностроении / А.Н. Андриченко // Онтология проектирования. 2012, 2(4). – С. 25-35.
- [12] **Ризванов, Д.А.** Интеллектуальная поддержка принятия решений при управлении ресурсами сложных систем на основе многоагентного подхода / Д.А. Ризванов, Н.И. Юсупова // Онтология проектирования. – 2015. Т.5, №3(17). - С.297-312. - DOI: 10.18287/2223-9537-2015-5-3-297-312.
- [13] **Thilmany, Jean.** Putting Artificial Intelligence to Work in CAD Design – Engineers Rule. December 21, 2017. - <https://www.engineersrule.com/solidworks-puts-artificial-intelligence>.
- [14] СПРУТ-Технология. Автоматизация проектирования. - <https://sprut.ru/>.

- [15] Программирование обработки на оборудовании с ЧПУ: учебник: в 2 т./ под ред. Г.Б. Евгенева и А.Х. Харджиева - Издательство МГТУ им. Н.Э. Баумана, 2018. - Т.1. – 374 с. - Т.2. – 356 с.
- [16] Центр СПРУТ. Высокопрофессиональные российские программные решения для предприятия - <https://csprut.ru/>.
- 

## INDUSTRY 5.0 AS INTEGRATION OF THE INTERNET OF KNOWLEDGE AND THE INTERNET OF THINGS

G.B. Evgenev

Moscow State Technical University named after N.E. Bauman, Moscow, Russia  
[g.evgenev@mail.ru](mailto:g.evgenev@mail.ru)

### Abstract

The methodology of creating systems of "Industry 5.0" class with the use of artificial intelligence technologies is proposed. The methodology is based on multi-agent methods of creating knowledge bases and is suitable for the development of design and control systems for digital intelligent industries. An Integrated structure of the Internet of Knowledge and the Internet of Things has been developed. The life cycle of mechanical engineering products is analyzed and the methods of application of the Internet of Knowledge and the Internet of Things at various stages of this cycle are proposed. The functional decomposition of the main stages of the life cycle is given. The conceptual foundations of the Internet of Knowledge are given. Multi-agent methods of knowledge base creation have been developed. The meta-ontology of engineering agents is proposed. The principles of construction of multi-agent systems for semi-automatic design of products are described. The description of the capabilities of intelligent programming of processing systems for CNC machining in terms of the formation of machining trajectory and transition areas is given. The possibilities of intelligent systems of design and regulation of technological processes are described. It is proposed to use the IDEF3 standard to create metamodels of technological processes and modified route maps for the formation of knowledge bases. The description of the intellectual system of operative management of industrial production is given. The General functional model of operation of an industrial enterprise operation is considered. The functional blocks of the system are described.

**Key words:** Industry 4.0, Industry 5.0, digital manufacturing, Internet of knowledge, Internet of things, intelligent systems.

**Citation:** Evgenev GB. Industry 5.0 as integration of the Internet of Knowledge and the Internet of Things [In Russian]. *Ontology of designing*. 2019; 9(1): 7-23. – DOI: 10.18287/2223-9537-2019-9-1-7-23.

### References

- [1] Industry 4.0: the fourth industrial revolution – guide to Industrie 4.0. - <https://www.i-scoop.eu/industry-4-0>.
- [2] Industry 4.0: the Future of Smart Manufacturing - Praim - <https://www.praim.com> › Article.
- [3] Siemens | Industrie 4.0 - <https://www.siemens.com/digital/enterprise>.
- [4] Digital Factory 4.0 | Industry 4.0 solution | antsolutions.eu - [//www.antsolutions.eu](http://www.antsolutions.eu).
- [5] Industry 4.0 - the Nine Technologies Transforming Industrial Production. - <https://www.bcg.com/.../embracing-industry-4-0-rediscoveri>.
- [6] *Noritsugu U.* Society 5.0: Mitsubishi Electric view [In Russian]. *Economic strategies*. 2017; 4: 2-11.
- [7] *Rzevski G.* Self-organization in social systems. *Ontology of Designing*. 2014; 4(14): 8-17.
- [8] Industry 5.0 [In Russian]. - <http://industry5.ru/koncept>.
- [9] *Evgenev GB.* The Basics of automation of technological processes and production. V.1: Information models. V.2: methods of design and management [In Russian]. Moscow, Izd-vo BMSTU, 2015. - V.1 - 441 p.; V.2 - 479 p.
- [10] *Grishin MV, Larin SN, Sosnin PI.* Ontology of designing industrial equipment within the aviation production [In Russian]. *Ontology of designing*. 2016; 6(1): 7-28. - DOI: 10.18287/2223-9537-2016-6-1-7-28.
- [11] *Andrichenko AN.* Tendencies and condition in the field of reference data management in the engineering industry [In Russian]. *Ontology of designing*. 2012; 2(4): 25-35.



- [12] **Rizvanov DA, Yusupova NI.** Intelligent decision support for resource management of complex systems based on multi-agent approach [In Russian]. *Ontology of designing*. 2015; 3(17): 297-312. - DOI: 10.18287/2223-9537-2015-5-3-297-312.
- [13] **Thilmany J.** Putting Artificial Intelligence to Work in CAD Design – Engineers Rule. December 21, 2017. - <https://www.engineersrule.com/solidworks-puts-artificial-intelligence>.
- [14] SPRUT-Technology. Automation design [In Russian]. - <https://sprut.ru/>.
- [15] Programming of machining on CNC equipment [In Russian]; Edited by G.B. Evgenev and A.H. Harajiev. Publishing house of the MSTU N.Uh. Bauman, 2018. - V.1 – 374 p.; V.2 – 356 p.
- [16] Center SPRUT. Highly professional Russian software solutions for companies [In Russian]. - <https://csprut.ru/>.

## Сведения об авторе



**Евгениев Георгий Борисович**, 1938 г. рождения. Окончил Московский авиационный технологический институт в 1960 г., к.т.н. (1964), д.т.н. (1975). Лауреат премии Совета Министров СССР (1991). Профессор кафедры «Компьютерные системы автоматизации производства» МГТУ им. Н.Э. Баумана. Член Российской ассоциации искусственного интеллекта. В списке научных трудов более 150 работ в области САПР и ИИ.

**Georgy Borisovich Evgenev** (b. 1938) graduated from the Moscow aviation technological Institute in 1960, PhD (1964), Doctor of engineering (1975). He is Council of Ministers of the USSR prize winner (1991). Professor at Moscow State Technical University named after N.E. Bauman. Member of the International Association for Ontology and its Applications. Co-author of about 150 scientific articles and abstracts in the field of CAD and AI.

## Примечание редакции

Для читателя, проявившего интерес к поставленным в статье вопросам, может быть полезна краткая справка о грядущих социальных преобразованиях в связи с эволюционно-технологическим переходом к Обществу 5.0.

### От Индустрии 4.0 к Обществу 5.0: план социальных преобразований в Японии

(Источник: <https://www.i-scoop.eu/industry-4-0-society-5-0/>)

Общество 5.0. Дорога к супер-умному обществу вымощена важной ролью технологий в таких областях, как Интернет вещей, искусственный интеллект, кибер-физические системы, Big Data и др.

Ожидается, что во всем мире более 20 процентов населения будут старше 60 лет к 2050 году. Это означает, что все страны должны следить за тем, что делает Япония, и как это Общество 5.0 работает в реальности. Среди проблем, с которыми Япония сталкивается больше, чем многие другие страны, - стихийные бедствия и загрязнение. Супер-умное общество должно предложить путь решения этих и других проблем.

Эволюция общества представлена в пяти социальных этапах: 1) охотничье общество, 2) аграрное общество, 3) индустриальное общество, 4) информационное общество и 5) супер-умное общество.

*Пять стен, которые необходимо преодолеть при переходе к Обществу 5.0.*

- 1) стена министерств и ведомств (разработка национальных стратегий и интеграции государственной системы продвижения, включая разработку «удобной системы Интернет вещей» и функции мозгового центра).
- 2) стена правовой системы (реформы в области регулирования административной «оцифровки»).
- 3) стена технологий (формирование, использование и защита «фонда знаний»: от кибербезопасности до робототехники, нано, био и системных технологий).
- 4) стена человеческих ресурсов (образовательная реформа, ИТ-грамотность, расширение доступных человеческих ресурсов со специализацией в продвинутых цифровых навыках).
- 5) стена общественного признания (необходимость социального консенсуса, тщательное рассмотрение социальных последствий и этических проблем, в том числе в отношении человек-машина, и таких философских вопросов, как определение того, что означает индивидуальное счастье и человеческое богатство).

(См. также: Keidanren (Japan Business Federation). Toward realization of the new economy and society. April 19, 2016. - [http://www.keidanren.or.jp/en/policy/2016/029\\_outline.pdf](http://www.keidanren.or.jp/en/policy/2016/029_outline.pdf)).

УДК 629.78

## ПРОЕКТИРОВАНИЕ СЛОЖНЫХ ИЗДЕЛИЙ С МАЛОЙ ВЕРОЯТНОСТЬЮ ОТКАЗОВ В УСЛОВИЯХ ИНДУСТРИИ 4.0

**Ю.П. Похабов**

АО «НПО ПМ - Малое Конструкторское Бюро», Железногорск, Россия  
pokhabov\_yury@mail.ru

### Аннотация

Успешность перехода на цифровые технологии проектирования и инжиниринга в условиях Индустрии 4.0 определяется не только совершенствованием технических и программно-вычислительных средств проектирования, но и эффективностью человеческой деятельности при их использовании. Одним из вызовов новой промышленной революции является необходимость повышения профессионального уровня проектанта и конструктора в области подготовки и верификации исходных данных на входе в компьютерную модель и валидации результатов его работы на выходе. Главным образом это касается наиболее трудоёмких и ответственных работ по проектированию сложных изделий с малой вероятностью отказов. Результат может быть достигнут за счёт применения новых методов конструкторско-технологического анализа для смягчения или исключения человеческих ошибок при принятии технических решений. Использование методов конструкторско-технологического анализа надёжности не нарушает естественный ход процессов проектирования и конструирования, при этом обеспечивает обоснование параметров функционирования, которое необходимо для принятия конструкторских решений. Для этого используются методы анализа, позволяющие представить конструкторскую документацию в виде параметрической модели функционирования. Данная модель позволяет эмулировать аварийные ситуации при худших сочетаниях режимов и условий эксплуатации, что невозможно достичь при экспериментальной отработке. На основании такого моделирования принимаются обоснованные конструкторские решения, исключающие потенциальные отказы. Для исключения маловероятных отказов предусмотрена оригинальная опция, позволяющая устанавливать необходимые и достаточные требования в конструкторской документации для достижения бездефектного производства.

**Ключевые слова:** Индустрия 4.0, цифровой двойник, цифровая тень, малая вероятность отказов, конструкторско-технологический анализ надёжности.

**Цитирование:** Похабов, Ю.П. Проектирование сложных изделий с малой вероятностью отказов в условиях Индустрии 4.0 / Ю.П. Похабов // Онтология проектирования. – 2019. – Т.9, №1(31). – С.24-35. – DOI: 10.18287/2223-9537-2019-9-1-24-35.

### Введение

В 2015 г. была определена концепция четвёртой промышленной революции – Индустрии 4.0 [1], в результате которой из привычного цикла промышленного изготовления в прошлое должны уйти длительные процессы конструкторских разработок, отработочных испытаний и постановки продукции на производство. Всё это будет заменено цифровыми технологиями на основе квазифизических моделей, которые позволят снизить стоимость разработок и создавать промышленную продукцию с уникальными потребительскими свойствами «с первого предъявления» (фактически без материальных затрат на экспериментальную отработку и доработку конструкций по результатам испытаний). Нет сомнений, что возможно осуществить основополагающие идеи Индустрии 4.0, но не лишне будет обратить внимание на то, с чем мы рискуем столкнуться на практике в техническом аспекте.



Более 40 лет в нашей стране уже развиваются отдельные элементы Индустрии 4.0 – вначале САПР, затем CAD/CAM/CAE-системы и CALS-технологии. За это время усовершенствовались технические и программно-вычислительные средства проектирования, но человек по-прежнему остаётся в автоматизированной среде проектирования слабым звеном: качество его мышления не улучшилось, аналитические способности не повысились, образовательная подготовка заметно упала. При этом значительно возросли навыки владения компьютерной техникой, но это слабо влияет на улучшение качества конструкторских разработок. Чтобы сделать шаг вперёд, необходимо подтянуть профессионализм проектантов (конструкторов) к вызовам Индустрии 4.0, и именно это должно стать одной из основных задач новой промышленной революции.

## 1 Предпосылки для реализации Индустрии 4.0 в области проектирования

Допустим, что можно построить «безупречную» *цифровую модель*<sup>1</sup> некоего объекта (включающую геометрические размеры, кинематические связи элементов конструкции, физико-механические характеристики конструкционных материалов, покрытий, параметры конструктивных соединений и пр.), которая является его тождественной копией в физическом мире (фактически это 3D-модель с детализацией физико-механических связей элементов конструкции). Предполагается, что количественное описание такой модели эквивалентно параметрам реального объекта: *геометрическим величинам* (конфигурации, формам, размерам, поперечным сечениям, допускам, зазорам, посадкам, люфтам и т.п.), *функциональным характеристикам* (силовым, жесткостным, кинематическим, электрическим, электромагнитным, термодинамическим и т.д.), *критериям* (показателям надёжности, живучести, безопасности, гидро- и газодинамическим критериям подобия и др.).

С одной стороны, такая цифровая модель представляет собой симулякр – репрезентативный образ некоего (будущего или реального) объекта в виртуальной реальности, который можно увидеть лишь с помощью цифровых средств визуализации (причём с любого ракурса – снаружи или изнутри, что в реальном мире невозможно); с другой стороны, цифровая модель существует в виртуальной среде в виде *числовой информации* об объекте и способна изменяться по законам, отражающим эквивалентные изменения данного объекта в условиях физического мира; наконец, с третьей стороны, при помощи конструкторских, технологических и производственных процедур цифровая модель может быть преобразована в физический объект, который можно не только видеть, но и пользоваться им по назначению в реальной действительности.

В отличие от цифровой модели в виртуальной реальности, материальный объект в реальном мире всегда выполняет предписанные ему функции (даже если он бездействует) и испытывает при этом различного рода нагрузки и воздействия (силовые, гравитационные, барические, атмосферные, тепловые, климатические, бактериологические, радиационные и т.п.). Такие нагрузки и воздействия могут быть внешними и внутренними (независимыми от внешних). В совокупности их принято называть *условиями и режимами* функционирования (эксплуатации) и именно они ограничивают возможности реального объекта выполнять заданные функции. Согласно тектологии А.А. Богданова [2], всё, что человек способен создать в техносфере, он может сделать с определённой (понятной ему) функциональностью, которая должна быть осуществлена в заданных условиях и режимах (т.е. известных наперёд). Другое дело, что не все условия и режимы функционирования человеку могут быть изначально известны, и вследствие этого человек рискует на практике создать неработоспособные, ненадёжные или неэффективные технические объекты. Предположим, что известны все

<sup>1</sup> Модель – сущность, воспроизводящая явление, объект или свойство объекта реального мира [ГОСТ Р 57188–2016, 2.1.1].

условия и режимы функционирования объекта и описывающая их *математическая модель*<sup>2</sup>, определяющая изменения состояния объекта (например, величины нагрузок и условия их приложения к объекту). В этом случае, совмещая в виртуальной реальности *цифровую* и *математическую модели*, можно получить *динамическую модель* поведения объекта во времени (например, в виде анимации реакций на внешние нагрузки и онлайн-протоколов параметров состояния), эквивалентную его физическому поведению в реальном мире<sup>3</sup>.

Поскольку на реальный объект и его поведение существенное влияние оказывают условия и технологии изготовления, то *цифровая модель* должна отражать технологические возможности реального производства, обусловленные, например, типом имеющегося оборудования для изготовления конструктивных соединений деталей (сварки, пайки, клёпки, склеивания, скрепления болтами, винтами и т.п.), а *математическая модель* – включать наследственные факторы, порождённые технологическими режимами изготовления, которые способны изменять поведение изделия при функционировании, например из-за остаточных и монтажных деформаций, ошибок позиционирования при изготовлении и проведении сборочных операций и т.д. При изготовлении изделия на той или иной производственной площадке его *цифровые* и *математические модели* должны быть приведены в соответствие с возможностями конкретного производства, чтобы адекватно отражать его технологию и не навредить общему замыслу конструктора.

*Цифровая модель* объекта и *математическая модель* условий и режимов его функционирования являются субъективным отражением мозговой деятельности человека по восприятию объективной реальности. Чем больше человек осведомлён, тем больше он способен видеть причинно-следственных связей в физическом мире, соответственно, более точно может описать их логико-математическими зависимостями. Однако познания человека не могут быть абсолютными, поэтому в любом случае *динамическая модель* будет отличаться от поведения реального объекта «пропорционально числу переменных, фигурирующих в его описании» [3].

Для достижения целей Индустрии 4.0, *динамическая модель* поведения объекта должна с допустимой погрешностью совпадать с поведением физического объекта в реальном мире. Понятно, что сравнить напрямую поведение реального объекта с его *динамической моделью* невозможно. Для этого используют т.н. *цифровую тень* – виртуальный образ физической реальности в режиме реального времени. Построить *цифровую тень* можно путём получения необходимой и достаточной информации о поведении реального объекта (*безошибочно изготовленного* согласно цифровой рабочей конструкторской документации и находящегося в заданных условиях и режимах функционирования). Формально, соответствие с заданной погрешностью *цифровой тени* и *динамической модели* служит критерием качества выполненной конструкторской разработки.

Чтобы осуществить *безошибочное изготовление* объекта (одно из основных условий для получения достоверной *цифровой тени*), необходимо чтобы требования технологической документации строго соответствовали требованиям рабочей конструкторской документации, элементы конструкции и изделие в целом были изготовлены без отступлений от этих требований, а ожидаемое качество изготовления было подтверждено объективными средствами технического контроля. Предположим, что перечисленные условия полностью обеспечены

<sup>2</sup> Математическая модель – модель, в которой сведения об объекте моделирования представлены в виде математических символов и выражений [ГОСТ Р 57188–2016, 2.1.2]. В данном контексте под математической моделью подразумевается пакет математических моделей для проведения мультидисциплинарных CAE- и HPC-расчётов (расчёты прочности, оптимизации, теплообмена, динамики, колебаний, вибрации, ударов, электромагнетизма, CFD-расчёты и т.п.).

<sup>3</sup> В простейшем виде динамическую модель могут характеризовать реакции математической модели, визуализированные в виде отчётов и спецификаций, – внутренние силовые нагрузки, напряжения, деформации, перемещения, формы и частоты собственных колебаний и т.п.

системой менеджмента качества на производстве (например, ИСО 9001) и потому не влияют на процесс проектирования/конструирования (по результатам изготовления не нужно принимать конструкторские решения по доработкам).

Если критерии качества разработки выполнены, то *динамическая модель* по факту является *цифровым двойником* реального объекта (*«единой моделью, достоверно описывающей все характеристики, процессы и взаимосвязи как для отдельного объекта, так и для всего производства»* [4]). Такой *цифровой двойник* способен служить информационно-справочной основой для прогнозирования, диагностики, коррекции режимов работы и проведения ремонтов материального объекта в реальном времени, а также отработки чрезвычайных действий на случай угроз или возникновения нештатных ситуаций (поломок).

Теоретически *цифровой двойник* можно получить напрямую (минуя составление *динамической модели* и *цифровой тени*), но для этого необходимо иметь генеральную совокупность данных, включающую вариации нагрузок/воздействий, характеристик материалов и параметров технологических процессов, что неизбежно потребует огромных организационно-технических ресурсов по сбору, хранению, обработке и передаче информации, не говоря уже о привлечении значительных финансовых средств.

Таковы теоретические предпосылки до соприкосновения с реальностью (если на практике будут соблюдены все вышеизложенные условия и допущения реализации задуманного).

## 2 Реалии проектно-конструкторских разработок

Создание технологий автоматизированного проектирования позволили перевести процессы рисования и расчётов в область компьютерных вычислений на основе обработки больших массивов данных. Функции человека и компьютера резко дифференцировались. Компьютер выполняет рутинную часть работы по графическому отображению и компилированию геометрических фигур, копированию изображений, редактированию, форматированию, генерированию 3D-моделей и вычислениям параметров на основе конечно-элементных или иных моделей. Скорость проведения разработок выросла, а вот с качеством всё оказалось сложнее. Если у конструкторов, заставших «докомпьютерные времена» и переобучившихся на новые методы конструирования, качество разработок повысилось, то качество разработок молодых конструкторов заметно упало. Конструкторы нового поколения стали делать больше ошибок, связанных с принятием инженерно-технических решений. Примерно то же самое произошло и с расчётчиками. Старые расчётчики ещё помнят, как проверить правильность результатов автоматизированного расчёта с помощью аналитических формул, а молодёжь этого уже не знает и воспринимает результаты компьютерных вычислений как «истину в последней инстанции». Не менее значимым является тот факт, что процессы проектирования и конструирования переместились в виртуальную область, и это для сторонних лиц вуалирует ценность занятия конструктора. Видимая лёгкость появления конструкторской документации «из-под компьютера» приводит к нивелированию роли конструктора (с позиций стороннего наблюдателя чертежи разрабатывает не конструктор, а компьютер).

Таким образом реалии компьютеризации конструкторских работ привели к возникновению двух групп причин падения качества конструкторских разработок. С одной стороны, менеджеры совершенно искренне стали полагать, что компьютеры думают за человека. Они воспринимают конструкторов и расчётчиков как обслуживающий персонал и ставят их в условия директивного сокращения сроков разработки, не оставляя им времени на осмысление результатов своих действий (всерьёз обсуждается тезис о бессмысленности обучения современных инженеров сопротивлению материалов, поскольку на практике им необходимы только умения пользоваться готовыми программами [5]).

С другой стороны, современное поколение конструкторов/расчётчиков не всегда осознаёт, что основная их задача – это не виртуозное владение компьютером, а способность (по Роберту Шекли) *правильно задать компьютеру вопрос, для чего прежде нужно знать бóльшую часть ответа*. Именно эта часть работы конструктора/расчётчика не изменилась и не изменится по мнению автора даже в случае применения подходов «проектирования за гранью интуиции генерального конструктора» [4]. Сегодня функции человека и компьютера максимально разделены. Роль и задача компьютера – это проведение вычислений по заданию человека. Роль и задача конструктора/расчётчика сводится к *подготовке и верификации исходных данных на входе в компьютер и к валидации результатов его работы на выходе*. В противном случае конструктор неизбежно превращается в «рисовальщика», а расчётчик – в оператора по загрузке компьютера малопродуктивной работой.

### 3 Пути повышения качества цифровых технологий проектирования

Возможности по достижению требуемой точности *динамической модели* достаточно очевидны. Прежде всего необходимо иметь методики проектирования и конструирования изделий, которые необходимы для повышения эффективности принимаемых человеком решений и выполнения им чётких действий. На практике добиться этого достаточно просто: для каждого типа изделия разрабатываются формализованные принципы проектирования и правила конструирования [6], что позволяет избежать неоправданных ошибок, которые по тем или иным причинам невозможно обнаружить при разработке цифровой модели и нельзя учесть в математической модели.

При проектно-конструкторских разработках сложно избежать ошибок (пропуск графических элементов чертежа, неустановление требований, упущение смысловых нюансов, задание не тех допусков, не задание твёрдости или шероховатости рабочих поверхностей и т.п.). Количество таких ошибок резко возрастает при разработках в условиях ограниченного времени. Для их обнаружения необходим беспристрастный критически настроенный проверяющий, который думает не о том, как конструкция должна работать без сбоев, а о том, по какой причине она непременно не будет работать так, как это задумал конструктор. Данный аспект связан с тем, что разработчик всегда склонен находить оправдания своим решениям.

Наконец самое главное – это использование новых методик обеспечения работоспособности и надёжности, позволяющих выявлять причины отказов на ранних стадиях жизненного цикла (ЖЦ). Такие методики разработаны для высоконадёжных изделий единичного изготовления [6], в частности для механических устройств раскрытия конструкций космических аппаратов. Их концепция целиком вписывается в условия реализации Индустрии 4.0.

### 4 Конструкторско-технологический анализ надёжности

Традиционные методы классической теории надёжности основаны на сборе и обработке данных об отказах изделий при эксплуатации и выявлении логико-математических связей между показателями надёжности и параметрами конструкции для прогнозирования частоты отказов. Новые методы достижения надёжности связаны с обеспечением стабильности параметров конструкции (во времени). Это осуществляется с помощью процедур, применяемых при проектировании, конструировании и технологической подготовке производства, которые базируются на соблюдении физических законов и исключении ошибок при разработке конструкторской и технологической документации. Базовым методом новой методики обеспечения надёжности является конструкторско-технологический анализ надёжности (КТАН) [6], который подразумевает проведение оценок надёжности. Однако смысл таких оценок заклю-



чается не в получении частотных характеристик отказов реальных объектов во времени, а в своевременном обнаружении и смягчении (предупреждении) последствий отказов на основе анализа моделей: *цифровой* и *математической*. Если использовать терминологию цифровых разработок, то в классическом понимании показатели надёжности можно получить только на основе изучения вероятностно-статистических моделей надёжности *цифровой тени* при реализации генеральной совокупности. На практике, если результаты *цифровой квалификации* изделия и оценки надёжности методами КТАН удовлетворяют установленным критериям, например, расчётная вероятность безотказной работы выше заданного уровня и точность *динамической модели* не ниже допустимой, то этого может быть вполне достаточно для подтверждения заданной работоспособности и надёжности. В некоторых случаях исследование надёжности можно проводить на *цифровых двойниках*, например, при оценке ресурса работы изделия.

КТАН может использоваться на каждой из стадий проектно-конструкторско-технологических работ ЖЦ, проведение такого анализа является целенаправленной итерационной процедурой. Всякое новое конструкторско-технологическое решение в процессе разработки приводит к обновлению исходных данных и повторению итерационного цикла анализов.

#### 4.1 Анализ-инициализация

*Анализ-инициализация* – это перевод чертёжно-технической документации в параметрическую модель функционирования. Если рассматривать данную процедуру в терминах цифровых технологий, то в качестве чертёжно-технической документации могут быть использованы цифровые модели, заданные условия и режимы эксплуатации, требуемые условия функционирования.

Анализ-инициализацию проводят последовательно:

1) декомпозируют конструкцию на конструктивные узлы, простейшие механизмы, функциональные группы и интерфейсы до уровня элементов, в которых может быть реализован режим точки единичного отказа.

*Например, для однозвенной поворотной штанги космического аппарата одной из точек единичного отказа является шарнирный узел поворота.*

2) выявляют полную функциональность (набор функций), способную удовлетворить заданные и подразумеваемые потребности, и делают словесное описание отказов как гипотетических ситуаций, препятствующих выполнению рассматриваемых функций.

*Например, для шарнирного узла штанги отказами являются нарушение условий прочности (разрушения и деформации) и невозможность поворота в шарнире (без разрушений и деформаций);*

3) фиксируют конкретные причины отказов при худших сочетаниях факторов режимов (включая состояние изделия после изготовления) и условий эксплуатации. Поскольку отказ может быть вызван одновременно несколькими причинами, то каждая из них становится привязанной к конкретному конструктивному элементу или узлу изделия, что автоматически указывает на его критичность.

*Например, для поворотной штанги разрушение шарнира определяется сочетанием максимальной нагрузки и минимальной возможности конструктивного элемента сопротивляться ей (неудачный выбор конструкционного материала или назначение геометрических размеров и допусков поперечного сечения), а невозможность совершения относительных поворотов в шарнире может быть вызвана:*

- *его запрессовкой из-за наихудшего сочетания радиальных зазоров после изготовления, толщины твёрдой смазки, механических нарушений трибосопряжения и воздействия теплового поля;*
- *заклиниванием из-за наихудшего сочетания осевого зазора и теплового перемещения элементов конструкции шарнира;*
- *торможением из-за худшего сочетания движущего момента на приводе и моментов сил сопротивления в шарнире при аномально низкой температуре;*

- неключением привода в шарнире из-за отказа, например вследствие производственного брака;
- зацеплением поворотной конструкции из-за ошибок проектирования, в результате которых на пути движения может возникнуть препятствие.

4) определяют свойства критичных элементов конструкции, наличие которых делает невозможной каждую из причин отказов (процедура реализуется методом парирования).

Например, для парирования отказов в шарнире поворотной конструкции необходимо:

- обеспечить необходимую несущую способность шарнира, исключающую отказы по условиям прочности;
- установить значения радиальных зазоров в шарнире, которые при худших сочетаниях толщины смазки, возможных механических нарушений трибосопряжений и тепловых воздействий не вызывают запрессовки шарнира;
- установить осевые зазоры в шарнире больше, чем относительные тепловые деформации в шарнире  $\Delta l$ ;
- обеспечить требуемый запас движущего момента на приводе, который приводит к устойчивому движению с заданными параметрами раскрытия;
- выбрать привод с требуемой вероятностью включения  $R_d \geq R_{lim}$ ;
- доказать, что поворотная конструкция имеет траектории движения свободные от зацеплений за смежные конструкции, в том числе из-за возможных неконтролируемых положений нежёстких конструктивных элементов в невесомости.

Выполнение процедуры парирования возможных причин отказов приводит к составлению вектор-столбцов параметров, которые должны быть присущи конструкции, чтобы исключить причины возможных отказов. Для шарнира поворотной конструкции, используемой в качестве примера, вектор-столбец параметров, характеризующих способность выполнять заданные функции, выглядит следующим образом:

$$(1) \quad X = (T, \Delta_0, \Delta_{sh}, M_{дв}, R_d, Q_{st})^T,$$

где

$T$  – несущая способность шарнира;

$\Delta_0$  – радиальный зазор в шарнире;

$\Delta_{sh}$  – осевой зазор в шарнире;

$M_{дв}$  – движущий момент в шарнире;

$R_d$  – надёжность включения привода;

$Q_{st}$  – вероятность зацепления поворотной конструкции.

5) каждое из свойств критичных элементов определяют с помощью параметров, которые характеризуют отношения элементов конструкции как взаиморасположение, взаимосвязь и взаимодействие внутри конструкции и вне её.

Параметры определяются:

- $T$  взаимодействием объекта с внешней средой посредством нагрузки  $N$ ;
- $\Delta_0$  – взаиморасположением оси и проушины шарнира в радиальном направлении, их взаимосвязью через твёрдую смазку, взаимодействием с внешней средой при тепловом воздействии среды или путём попадания абразивных частиц из окружающей среды в радиальный зазор;
- $\Delta_{sh}$  – взаиморасположением оси и проушины шарнира в осевом направлении и взаимодействием с внешней средой при тепловом воздействии;
- $M_{дв}$  – взаимосвязью с сопротивлением повороту в шарнире  $M_c$  и взаимодействием с температурой внешней среды, приводящей к увеличению сопротивления;
- $R_d$  – взаимодействием с приводом;
- $Q_{st}$  – взаиморасположением поворотной конструкции с внешними объектами.

6) каждому параметру определяют диапазон допустимых значений исходя из требований технического задания на разработку (представления изделия с точки зрения заказчика) и конструктивного исполнения изделия (представления конструкции изделия с точки зрения разработчика).

В результате вектор-столбец (1) может быть преобразован следующим образом

$$(2) \quad D_x = \left\{ \begin{array}{l} T > N \\ \Delta_0 > 0 \\ \Delta_{sh} > \Delta l \\ M_{дв} > M_c \\ R_d \geq R_{lim} \\ Q_{st} \rightarrow 0 \end{array} \right\}.$$

Множество  $D_x$  (2) является представлением будущего изделия в виде параметров, характеризующих способность выполнять требуемые функции в заданных режимах и условиях эксплуатации.

## 4.2 Оценка теоретической надёжности

Оценка теоретической надёжности по проектным параметрам объекта – это расчёты надёжности в том смысле, как это понимается в современной инженерии. Отличие в том, что в новой методике расчёты надёжности проводят по каждому параметру в отдельности (определяется условная вероятность) и затем производится интегрирование условных вероятностей для получения полной надёжности [6–8]. Предусмотрены два метода оценки теоретической надёжности: детерминированный (путём задания конструктивных запасов по каждому из параметров таким образом, чтобы гарантировать нахождение их значений в допустимой области) [6] и стохастический (путём оценки конструкционной индивидуальной надёжности [7], суть которой заключается в вычислении вероятностей нахождения параметров в допустимых областях исходя из индивидуальных характеристик материалов, процессов нагружения/воздействий и технологий изготовления изделий).

В качестве примера приведём оценку надёжности поворотной штанги детерминированным методом с учётом (2). Запишем параметрическую надёжность штанги в виде

$$(3) \quad \begin{aligned} R_1 &= P(\underline{T} > \bar{N}); \\ R_2 &= P(\underline{\Delta}_0 > 0); \\ R_3 &= P(\underline{\Delta}_{sh} > \bar{\Delta l}); \\ R_4 &= P(\underline{M}_{дв} > \bar{M}_c); \\ R_5 &= R_d; \\ R_6 &= 1 - Q_{st}. \end{aligned}$$

Каждая из составляющих параметрической надёжности (3) определяется с учётом обеспечения конструктивных запасов:

- резервирования путём  $m$ -ой кратности структурного резервирования для определения  $R_5 = 1 - (1 - R_d)^{m+1}$ ;
- нормированных коэффициентов безопасности и запасов прочности для определения  $R_1 = P(T/\eta > fN)$ ;
- запасов движущих моментов для определения  $R_4 = P(\underline{M}_{дв} > kM_c)$ ;
- параметрической избыточности в виде силовых и температурных развязок, например для определения  $R_3 = P(\underline{\Delta}_{sh} > \bar{\Delta l})$ , когда конструктивными методами достигается условие  $\underline{\Delta}_{sh} \gg \bar{\Delta l}$  (если из двух шарниров, образующих единую ось вращения, один не имеет ограничений по перемещению вдоль оси);
- процедур получения гарантированных результатов с использованием, например, минимаксных критериев для определения  $R_2 = P(\underline{\Delta}_0 > 0)$ , когда могут быть применены геометрические построения (3D-моделирование), либо факторов инженерной психологии для определения  $R_6 = 1 - Q_{st}$  с учётом критического анализа ошибок проектирования, допускающих несанкционированные действия человека при работе с изделием.

Суммарная надёжность с учётом  $R_1 \dots R_6$  определяется по формуле

$$(4) \quad R = \prod_{i=1}^n R_i.$$

При соответствующем выборе конструктивных запасов при определении проектно-конструкторских параметров теоретически может быть достигнута надёжность  $R \equiv 1$ . На основании формулы (4) производится подтверждение того, что выбранные значения параметров соответствуют требованиям технического задания, если  $R > R_{зад}$ .

## 4.3 Анализ соответствия

Суть анализа соответствия проектных параметров требованиям конструкторской и технологической документации заключается в том, чтобы убедиться, что выбранные проектные параметры функционирования соответствуют требованиям, установленным в документации согласно формуле



(5) 1 параметр : $\Leftrightarrow$  минимум 1 требование.

Исходя из (5) каждому параметру должно быть предусмотрено соответствующее требование для изготовления и/или эксплуатации, выполнение которого может быть проверено средствами технического контроля.

#### 4.4 Оценка рисков

*Оценка рисков* возникновения отказов вследствие неустановления требований в документации, т.е. по состоянию «как есть». Нерелевантность параметров и требований конструкторской и/или технологической документации, риски невыполнения или ненадлежащего выполнения требований при изготовлении рассматриваются как события  $C_i$ , где индекс  $i$  соответствует  $i$ -му элементу рассматриваемой системы. Вероятность совершения каждого из таких событий в формуле (3) может быть определена в виде

$$(6) \quad P(C_i) = \alpha_i R_i,$$

где  $\alpha_i$  – корректирующие коэффициенты, которые могут быть получены экспертным путём, например, с применением шкалы балльных оценок критичности отказов согласно [8].

Для вычисления окончательной вероятности выполнения функции с учётом (3) и состояния конструкторской и технологической документации (6) используют формулу

$$(7) \quad P(C) = \prod_{i=1}^n P(C_i).$$

#### 4.5 Окончательной оценка надёжности

*Окончательная оценка надёжности* с учётом теоретической надёжности и рисков возникновения отказов производится с учётом (7) по формуле  $P(C) > R_{\text{зад}}$ .

Последние три процедуры отражают новый подход к анализу и оценке надёжности, целесообразность которых возникает при требованиях к безотказности выше 0,999, т. е. для высокоответственных изделий. Смысл проведения анализа и оценки рисков возникновения отказов из-за нерелевантности проектных параметров требованиям документации вытекает из простой логики. Исходя из оценки теоретической надёжности можно получить расчётно-экспериментальный результат, согласно которому следует, что все параметры выбраны верно. Но такая оценка не учитывает, что не отражение или нечёткость любого из параметров в требованиях документации способна привести к непредвиденному отказу (как правило, с малой вероятностью). Например, отказ может оказаться следствием неопределённости требований по умолчанию<sup>4</sup>. Документация должна содержать все необходимые и достаточные требования, тогда их строгое выполнение приведёт к реализации проектных параметров.

Оценка теоретической надёжности является процедурой верификации параметров функционирования, а оценка рисков – это валидации соответствия реального объекта технической документации.

### 5 Об обязательности проведения анализов

Методика КТАН основана на парадигме А.И. Уёмова о триединстве вещей, их свойств и отношений [9], который, в свою очередь, опирался на модель Аристотеля. При этом отношения рассматриваются как взаиморасположение, взаимосвязь и взаимодействие вещей при определении свойств [6]. Такое понимание позволяет определять любые свойства через от-

<sup>4</sup> Умолчание – соглашение о характеристике языкового объекта или выполняемом действии при отсутствии их явного описания [10].

ношения физических объектов, а сами свойства выражать количественно через параметры и показатели. Допустимые значения параметров определяются заданными требованиями технического задания на разработку и конструктивным исполнением изделия. В методике КТАН *цифровая модель* строится на основе определения взаиморасположения и взаимосвязей элементов конструкций, а *математическая модель* – с учётом взаимодействий объектов реального мира. Таким образом реальные объекты представляются в виде набора свойств, которые могут быть выражены через параметры с допустимыми границами изменения их значений. Такой подход универсален и может быть реализован при рассмотрении свойств как гипотетических отношений вещей, проявление которых в реальном мире возможно и вероятно. Сочетая в моделях предельные отношения свойств вещей, можно исследовать *динамическую модель* в критических ситуациях, приближая её к *цифровому двойнику*.

Указанный подход в процессе цифровых разработок позволяет:

- выявлять и принимать меры по устранению отказов на ранних этапах ЖЦ путём парирования причин возможных отказов за счёт придания будущему изделию предписанных свойств соответствующих критичных элементов;
- дополнять цифровые модели информацией, которую в виртуальной реальности человеку невозможно представить без применения органолептических методов;
- уйти от стереотипов при проведении расчётов, которые характерны для длительной практики конструкторских разработок однотипной продукции, и не всегда отражают целесообразность, необходимость и достаточность их применения с точки зрения обеспечения заданной работоспособности и надёжности;
- дополнить математические модели вероятностными моделями поведения элементов конструкций с сингулярностями, нелинейностями, логико-вероятностными зависимостями, статистическими закономерностями и пр.;
- определять целесообразность и критерии проведения анализов, расчётов<sup>5</sup>, экспериментов<sup>6</sup>, опытов и испытаний<sup>7</sup> с целью получения необходимой информации для построения моделей.

## Заключение

Производя компьютерные вычисления, необходимо заботиться о максимальном снижении доли человеческих ошибок в исходных данных и повышении знаний о физической природе мира. Анализы, подобные КТАН, позволяют балансировать на грани виртуальной реальности цифровых моделей и реальности физического мира, и на этом балансе производить валидацию конструкторских разработок до момента вывода их на стадию производства. Это позволит повысить качество разработок сложных изделий с малой вероятностью отказов.

## Благодарности

Работа выполнена при частичной финансовой поддержке НИР «Гарантия-2021-НПО ПМ МКБ» согласно Федеральной космической программе России на 2016-2025 годы.

<sup>5</sup> Расчёт – документ, содержащий расчёты параметров и величин, например, расчёт на прочность и др. [11].

<sup>6</sup> Эксперимент – система операций, воздействий и/или наблюдений, направленных на получение информации об объекте при исследовательских испытаниях [12].

<sup>7</sup> Испытание – экспериментальное определение количественных и/или качественных характеристик свойств объекта испытаний как результата воздействия на него, при его функционировании, при моделировании объекта и/или воздействий [13].

### СПИСОК ИСТОЧНИКОВ

- [1] **Шваб, К.** Четвёртая промышленная революция. – М.: Эксмо, 2016. – 138 с. – ISBN 978-5-699-90556-0.
- [2] **Богданов, А.А.** Тектология: (Всеобщая организационная наука). – М.: Экономика, 1989. – Кн. 1. – 304 с. – ISBN 5-282-00538-7.
- [3] **Акоф, Р.** Основы исследования операций / Р. Акоф, М. Сасиени. Пер. с англ. – М.: Мир, 1971. – 533 с.
- [4] **Боровков, А.И.** Цифровые двойники и цифровая трансформация предприятий ОПК / А.И. Боровков, Ю.А. Рябов, К.В. Кукушкин и др. // Оборонная техника. – 2018. – № 1. – С.6–33.
- [5] **Кулешов, А.П.** Преодолеть сопротивление материалов [интервью с ректором Сколтеха, акад. РАН / записал А. Механик] // Стимул: журнал об инновациях в России. – [https://stimul.online/articles/interview/preodolet-soprotivlenie-materialov/?sphrase\\_id=1295](https://stimul.online/articles/interview/preodolet-soprotivlenie-materialov/?sphrase_id=1295).
- [6] **Похабов, Ю.П.** Теория и практика обеспечения надёжности механических устройств одноразового срабатывания. – Красноярск: Сиб. федер. ун-т, 2018. – 338 с. – ISBN 978-5-7638-3812-1.
- [7] **Тимашев, С.А.** Инфраструктуры. Т. 1: Надёжность и долговечность. – Екатеринбург, 2016. – 530 с. – ISBN 978-5-7691-2434-1.
- [8] ГОСТ 27.310-95. Надёжность в технике. Анализ видов, последствий и критичности отказов. Основные положения. Dependability in technics. Failure mode, effects and criticality analysis. Basic principles. Актуализация 12.09.2018.
- [9] **Уёмов, А.И.** Вещи, свойства и отношения. – М.: Изд-во АН СССР, 1963. – 184 с.
- [10] ГОСТ 33707-2016 (ISO/IEC 2382:2015) Информационные технологии (ИТ). Словарь. Information technologies. Vocabulary. Дата введения 2017-09-01.
- [11] ГОСТ 2.102-2013 Единая система конструкторской документации (ЕСКД). Виды и комплектность конструкторских документов. Unified system for design documentation. Types and sets of design documentation. Дата введения 2014-06-01.
- [12] ГОСТ 24026-80 Исследовательские испытания. Планирование эксперимента. Термины и определения. Research tests. Experiment planning. Terms and definitions. Дата введения 1981-01-01.
- [13] ГОСТ 16504-81. Система государственных испытаний продукции. Испытания и контроль качества продукции. Основные термины и определения. The state system of testing products. Product test and quality inspection. General terms and definitions. Дата введения 1982-01-01

---

## DESIGNING COMPLEX PRODUCTS WITH SMALL PROBABILITY OF FAILURE IN THE CONTEXT OF INDUSTRY 4.0

**Yu.P. Pokhabov**

*JSC "NPO PM SDB", Zheleznogorsk, Russia  
pokhabov\_yury@mail.ru*

### Abstract

The success of the transition to digital design and engineering in Industry 4.0 is determined not only by the improvement of technical and computer-aided design tools, but also by the efficiency of human activity in their use. One of the challenges of the new industrial revolution is to increase the professional level of designers and developers in the field of preparing and verifying initial computer input data validating the results of its work at the output. This mainly concerns the most time-consuming and responsible work on the design of complex products with a low probability of failure. The result can be achieved through the use of new methods of design and technological analysis to mitigate or eliminate human errors when making technical decisions. Using the methods of design & technology reliability analysis does not violate the natural course of the design and development processes, while providing a justification of the parameters of operation, which is necessary for making faultless design decisions. For this purpose, analysis methods are used, which allow to submit design documentation in the form of a parametric model of operation. This model allows simulating emergency situations with the worst combinations of modes and operating conditions, which is impossible to achieve with experimental testing. Based on such modeling, reasonable design decisions are made that exclude potential failures. To eliminate low probability failures, an original option is provided, which allows establishing the necessary and sufficient requirements in the design documentation to achieve defect-free production.

**Keywords:** Industry 4.0, digital twin, digital shadow, small probability of failures, design & technology reliability analysis (DTRA).

**Citation:** Pokhabov YuP. Designing complex products with small probability of failure in the context of Industry 4.0 [In Russian]. *Ontology of designing*. 2019; 9(1): 24-35. – DOI: 10.18287/2223-9537-2019-9-1-24-35.

## Acknowledgment

The work was partially supported from the research project “Garantiya-2021- NPO PM SDB” according to the Federal Space Program of Russia for 2016-2025.

## References

- [1] **Schwab K.** Chetvortaya promyshlennaya revolyutsiya [The Fourth Industrial Revolution]. – Moscow: Eksmo, 2016. – 138 p. – ISBN 978-5-699-90556-0. [In Russian].
- [2] **Bogdanov AA.** Tektologiya: (Vseobshchaya organizatsionnaya nauka) [Tectology: (Universal organizational science)]. – Moscow: Ekonomika, 1989. – B. 1. – 304 p. – ISBN 5-282-00538-7. [In Russian].
- [3] **Akof RL, Sasieni MW.** Fundamentals of operations research. – John Wiley and sons, NY, 1968.
- [4] **Borovkov AI, Ryabov YuA, Kukushkin KV.** et al. [Digital counterparts and digital transformation of enterprises DIE] / *Oboronnaya tekhnika*. – 2018; 1: 6–33. [In Russian].
- [5] **Kuleshov AP.** [Overcome the resistance of materials] [interview with the Skoltech rector, acad. RAS / recorded A. Mechanic] // Stimul: zhurnal ob innovatsiyakh v Rossii [In Russian] – [https://stimul.online/articles/interview/preodolet-soprotivlenie-materialov/?sphrase\\_id=1295](https://stimul.online/articles/interview/preodolet-soprotivlenie-materialov/?sphrase_id=1295).
- [6] **Pokhabov YuP.** Teoriya i praktika obespecheniya nadezhnosti mekhanicheskikh ustroystv odnorazovogo srabatyvaniya [Theory and practice of ensuring the reliability of mechanical devices of one-time operation]. – Krasnoyarsk: SFU, 2018. – 338 p. – ISBN 978-5-7638-3812-1. [In Russian].
- [7] **Timashev SA.** Infrastruktury. T. 1: Nadezhnost i dolgovechnost [Infrastructure. Vol. 1: Reliability and durability]. – Yekaterinburg, 2016. – 530 p. – ISBN 978-5-7691-2434-1. [In Russian].
- [8] GOST 27.310-95. Dependability in technics. Failure mode, effects and criticality analysis. Basic principles. [In Russian]. Update: 12.09.2018.
- [9] **Uyomov AI.** Veshchi, svoystva i otnosheniya [Things, properties and relationships]. – Moscow: USSR Academy of Sciences, 1963. – 184 p. [In Russian].
- [10] GOST 33707-2016 (ISO/IEC 2382:2015) Information technologies. Vocabulary. [In Russian]. Publication date: 2017-09-01.
- [11] GOST 2.102-2013 Unified system for design documentation. Types and sets of design documentation. [In Russian]. Publication date: 2014-06-01.
- [12] GOST 24026-80 Research tests. Experiment planning. Terms and definitions. [In Russian]. Publication date: 1981-01-01.
- [13] GOST 16504-81. The state system of testing products. Product test and quality inspection. General terms and definitions. [In Russian]. Publication date: 1982-01-01.

## Сведения об авторе



**Похабов Юрий Павлович**, 1959 г. рождения. Окончил филиал Красноярского политехнического института (завод-ВТУЗ, ныне – СибГУ им. М.Ф. Решетнёва) в 1982 г. Область интересов: проектирование, конструирование и обеспечение надёжности конструкций космических аппаратов. Канд. техн. наук (2013). Автор более 50 научных работ, в том числе 2 монографий.

**Yuri P. Pokhabov** (b. 1959). Graduated from the branch of the Krasnoyarsk Polytechnic Institute (currently – Reshetnev Siberian State University of Science and Technology) in 1982. Area of interests: design, develop and maintenance of reliability of structures of spacecraft. Ph.D (2013). Author of more than 50 scientific works, including 2 monographs.

УДК 004.822

## ОНТОЛОГИЯ КОНЕЧНО-АВТОМАТНОЙ КРИПТОГРАФИИ

А.А. Шарипбай<sup>1</sup>, Ж.С. Сауханова<sup>2</sup>, Г.Б. Шахметова<sup>3</sup>, М.С. Сауханова<sup>4</sup>

Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан

<sup>1</sup> sharalt@mail.ru, <sup>2</sup> saukhanova@mail.ru, <sup>3</sup> shakhmetova.gb@gmail.com, <sup>4</sup> m.saukhanova@mail.ru

### Аннотация

На сегодняшний день совершенствование методов защиты информации является актуальной задачей. Статья посвящена применению альтернативных методов разработки более стойких и эффективных криптосистем с открытыми ключами. В качестве модели приняты конечные автоматы. Для систематизации знаний в области конечно-автоматной криптографии использована онтология. В работе рассматриваются вопросы построения онтологической модели конечно-автоматной криптографии. Предлагаемая онтологическая модель имеет четыре основных уровня: онтология представления знаний, онтология верхнего уровня, онтология предметной области и прикладные онтологии. В качестве онтологии представления знаний использована методология концептуальных карт. Онтология верхнего уровня содержит основную информацию о криптографии и теории автоматов, онтология предметной области описывает непосредственно конечно-автоматную криптографию. В качестве примера прикладной онтологии был рассмотрен алгоритм криптосистемы с открытым ключом на основе конечных автоматов. Представленная онтология построена впервые и даёт чёткое понимание применения конечных автоматов в криптографии, систематизирует полученные в ходе исследования сведения о данной предметной области, является предпосылкой дальнейшей разработки криптосистем, основанных на теории автоматов.

**Ключевые слова:** онтология, концептуальная карта, конечный автомат, криптография.

**Цитирование:** Шарипбай, А.А. Онтология конечно-автоматной криптографии / А.А. Шарипбай, Ж.С. Сауханова, Г.Б. Шахметова, М.С. Сауханова // Онтология проектирования. – 2019. – Т.9, №1(31). – С.36-49. – DOI: 10.18287/2223-9537-2019-9-1-36-49.

### Введение

В условиях быстрого развития информационных технологий и внедрения их во все сферы человеческой деятельности обеспечение надёжной защиты информации, передаваемой по незащищённым каналам связи, становится актуальной проблемой современного общества. Известно, что криптография занимается исследованием методов преобразования (шифрования и дешифрования) информации с целью скрытия её содержания [1, 2], а одним из перспективных направлений в криптографии является применение теории конечных автоматов (КА) для шифрования и дешифрования информации.

Для систематизации знаний в области конечно-автоматной криптографии (КАКГ) было принято решение использовать *онтологию*. Известно, что онтология позволяет концептуализировать предметную область (ПрО), формализовать накопленные знания: определить ключевые понятия, задать семантические отношения между понятиями, необходимые для постановки задач и описания процессов их решения в данной ПрО. Кроме того, преимуществом использования онтологии является возможность анализа, накопления и повторного применения знаний о ПрО, полученной из разных источников [3]. Поэтому основной целью настоящей статьи является построение онтологической модели КАКГ.

В качестве инструмента был выбран общедоступный и хорошо себя зарекомендовавший на практике программный продукт SmartTools [4], который используется для построения концептуальных карт (К-карт).



## 1 Особенности создания онтологий с помощью концептуальных карт

При исследовании ПрО необходима систематизация полученных знаний. Область исследования можно представить в виде *концептуальной модели* ПрО, которая содержит в себе множества понятий (концептов), классификацию, свойства и характеристики этих понятий. В настоящее время применяют *онтологию* для наглядного представления концептуальной модели ПрО. Под онтологией понимается «формальная спецификация концептуализации, которая имеет место в некотором контексте ПрО» [5].

Формально онтологию можно определить следующим образом [6]:

$O = \langle C, R, A \rangle$ , где:

- $C$  – конечное множество концептов (понятий, терминов) ПрО;
- $R$  – конечное множество отношений между концептами (понятиями, терминами) ПрО;
- $A$  – конечное множество аксиом или функций интерпретации, заданных на концептах и (или) отношениях.

Для построения онтологической модели КАКГ выбрана методология концептуальных карт (concept maps, К-карт). К-карта позволяет графически представить знания изучаемой ПрО. Она представляет собой ациклический граф, вершины которого есть основные понятия (концепты) рассматриваемой ПрО, а ребра – связи между понятиями (отношения) [7]. «Концепты и связи между ними имеют универсальный характер для некоторого класса понятий ПрО. Поэтому любая разработка К-карты подразумевает анализ структурных взаимодействий между отдельными понятиями ПрО» [8]. На рисунке 1 показан пример построения онтологии К-карты с помощью инструмента SmartTools.

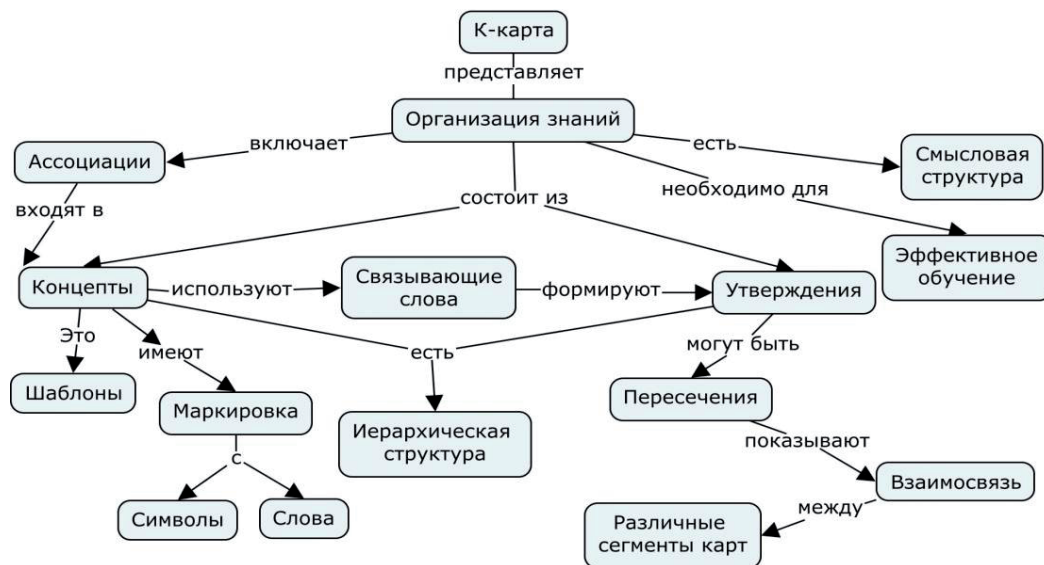


Рисунок 1 – Онтология К-карты

К-карты можно отнести к «лёгким» онтологиям, которые не содержат аксиом [9] и имеют вид:

$O_k = (C, R)$ , где:  $C$  – конечное множество концептов ПрО;  $R$  – множество отношений между концептами.

Выбор К-карт для создания онтологической модели КАКГ обусловлен их следующими преимуществами [6]:

- *системность* – К-карты позволяют представить целостный взгляд на изучаемую ПрО;
- *единообразие* – материал воспроизводится и воспринимается эффективней, если представлен в единой форме;



- *научность* – формирование К-карты ПрО позволяет выявить недостающие логические связи во всей их полноте;
- *когнитивность* – в процессе построения К-карт используются все виды памяти человека, что позволяет быстро запоминать представленные картами сведения об изучаемой ПрО.

Рассматриваемая ПрО представляет собой сложно-структурированную область и включает в себя понятия криптографии, теории автоматов и прикладную терминологию, содержащую понятия по конкретным реализациям криптосистем, основанных на КА. Онтологию КАКГ можно разделить на четыре основных уровня, которые показаны на рисунке 2.

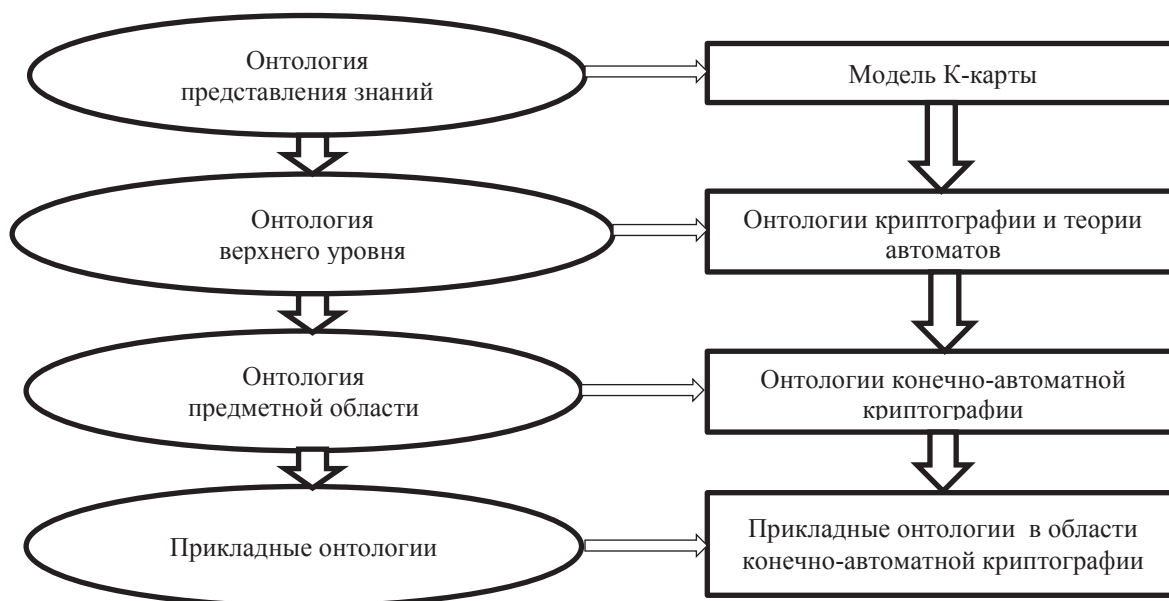


Рисунок 2 – Уровни онтологической модели области конечно-автоматной криптографии

Первый уровень – онтология представления знаний – описывает область представления знаний. Целью первого уровня является создание языка для спецификаций других онтологий более низкого уровня. В качестве онтологии первого уровня использована модель К-карты, показанной на рисунке 1.

Онтология верхнего уровня – это онтология криптографии и теории автоматов, поскольку ПрО объединяет криптографию и теорию КА.

Третий уровень онтологии – это онтология КАКГ, в которой даются описания основных концептов конечно-автоматной модели. По своей структуре онтология КАКГ является логическим продолжением онтологии верхнего уровня.

Четвёртый уровень онтологии – это прикладные технологии КАКГ. Данный уровень содержит специфичную информацию – концепты и отношения, которые раскрывают особенности определённых криптосистем, основанных на КА.

## 2 Онтология конечно-автоматной криптографии

КАКГ основана на использовании теории автоматов и криптографии. Каждая из этих наук включает в себя огромный понятийный аппарат. В статье выбраны основные концепты, которые дают ясное представление об изучаемой ПрО. На рисунке 3 показано объединение данных направлений. В верхней части рисунка 3 (а) концептуальной модели расположены основные концепты и отношения, относящиеся к криптографии, нижняя часть рисунка 3 (с)

содержит важные концепты и связи теории автоматов, на стыке двух онтологий (рисунок 3 (b)) представлены существующие криптосистемы, построенные на основе КА.

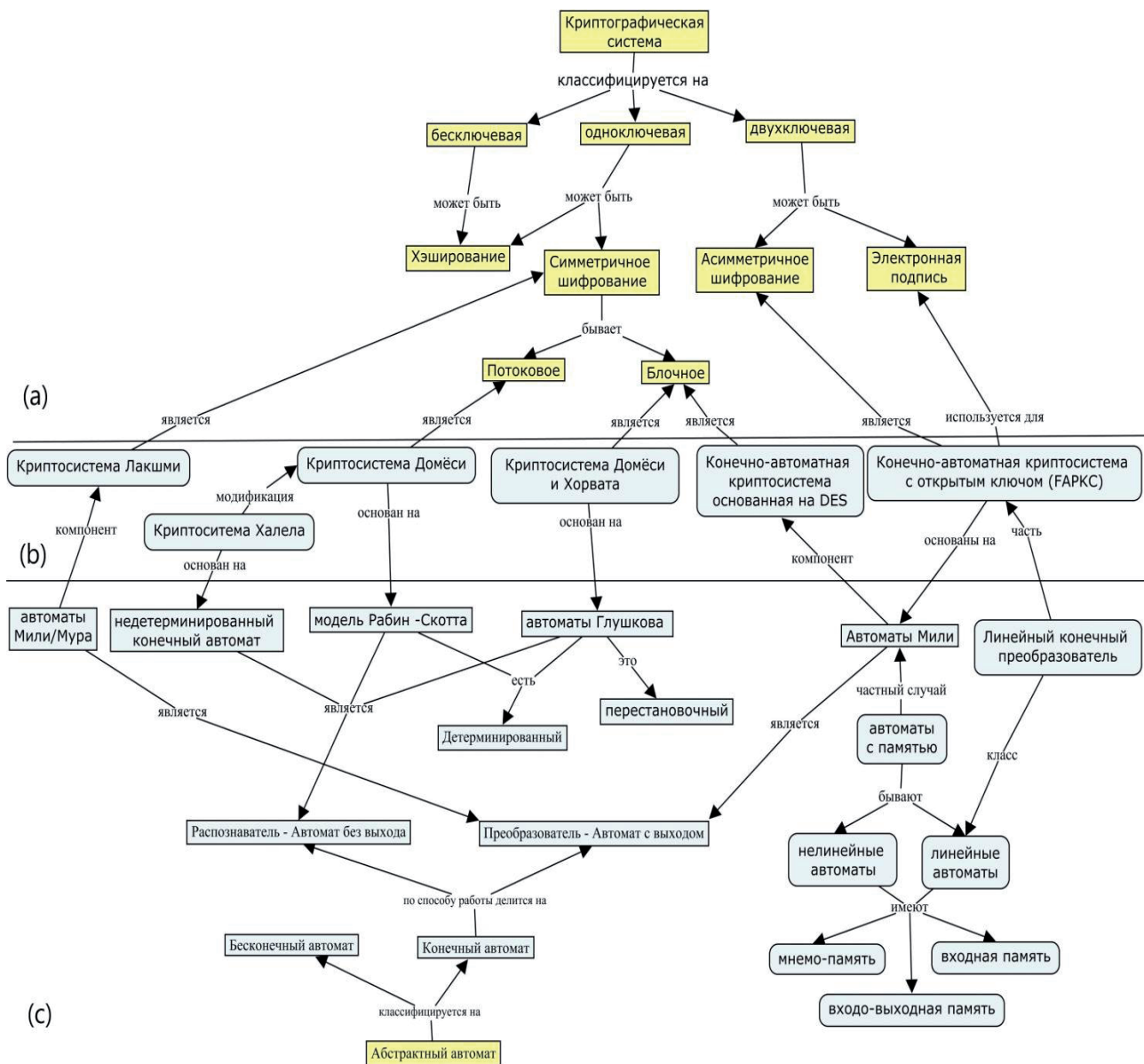


Рисунок 3 – Онтология конечно-автоматной криптографии

В онтологической модели криптографии вершиной является криптографическая система. Криптографическая система представляет собой набор из криптографического алгоритма, открытого и закрытого текстов и ключей. Криптографический алгоритм (шифр) – это математическая функция, которая используется для преобразования открытого текста в зашифрованный (шифрование) и обратно (дешифрование) [10]. Ключ является основным компонентом шифра, который непосредственно отвечает за выбор преобразований криптотекста и хранится в тайном месте. В зависимости от числа ключей криптографические системы разделяют на бесключевые, одноключевые и двухключевые.

Бесключевая криптографическая система реализует метод хэширования (*hashing*) или контрольное преобразование информации. Одноключевая криптосистема – это способ симметричного шифрования, в котором используется один секретный ключ, как для шифрова-

ния, так и для дешифрования. Симметричная криптосистема подразделяется на блочное и потоковое шифрование. Двухключевая криптосистема применяется в асимметричном шифровании и в электронной цифровой подписи. В асимметричном шифровании используют два типа ключей: открытые – для шифрования исходного текста, и секретные – для дешифрования зашифрованного текста. Оба эти ключа связаны между собой сложным соотношением. Электронная цифровая подпись необходима для подтверждения целостности и авторства данных [11].

Криптографические системы, основанные на КА – это алгоритмы, в качестве ключей которых используются КА. В вершине онтологии теории автоматов лежит концепт «Абстрактный автомат» (АА). Согласно [12] АА – это модель дискретного устройства, которое описано пятиместным кортежем:

$A = \langle X, Y, S, \delta, \lambda \rangle$ , где:

$X$  – множество входных символов;

$Y$  – множество выходных символов;

$S$  – множество внутренних состояний;

$\delta: S \times X \rightarrow S$  – функция переходов;

$\lambda: S \times X \rightarrow Y$  – функция выходов.

В случае, когда множества  $X, Y, S$  – конечны, АА является КА. Если хотя бы одно из перечисленных множеств бесконечно, то АА называется *бесконечным*.

По принципу однозначности функции перехода КА можно классифицировать как *детерминированный* или *недетерминированный КА*.

Детерминированность автомата заключается в выполнении условия однозначности переходов, т.е., если автомат находится в некотором состоянии и под воздействием произвольного входного символа переходит в одно и только одно состояние. При недетерминированности КА, он под воздействием одного и того же входного символа может перейти в различные состояния из множества состояний  $S$  [13].

По способу работы КА разделяют на два вида [14]:

- *автомат – преобразователь (автомат с выходом)*: данный вид автомата преобразовывает поступившую на входе информацию в выходную последовательность, т.е. реализует автоматные отображения;
- *автомат – распознаватель (автомат без выхода)*: данный вид автомата распознаёт поступившую на вход последовательность, т.е. отвечает на вопрос: принадлежит ли входная последовательность к данному множеству.

В класс преобразователей входят автоматы Мили/Мура. Согласно [15] данные автоматы имеют следующую формальную запись:

$$\text{Автомат Мили: } \begin{cases} s(t+1) = \delta(s(t), x(t)) \\ y(t) = \lambda(s(t), x(t)) \end{cases}, t=0,1,2,\dots$$

$$\text{Автомат Мура: } \begin{cases} s(t+1) = \delta(s(t), x(t)) \\ y(t) = \lambda(s(t)) \end{cases}, t=0,1,2,\dots$$

КА Мили, находясь в начальном состоянии  $s(0)$ , под действием входной последовательности  $x(0)x(1)\dots$  проходит последовательность состояний  $s(0)s(1)\dots$  и вырабатывает выходную последовательность  $y(0)y(1)\dots$ . Зависимость между входными символами, состояниями автомата и выходными символами в дискретном времени  $t$  показана с помощью данных систем канонических уравнений.

В автоматах Мура, в отличие от автоматов Мили, выходная последовательность определяется только состоянием автомата в какой-то момент времени  $t$  и не зависит от входной последовательности в этот же момент времени.

Как видно из рисунка 3 автомат с памятью является частным случаем автомата Мили. Данные автоматы применяются в реализации конечно-автоматной криптосистемы с открытым ключом (ФАРКС). Согласно [16], если функция  $\varphi: Y^k \times X^{h+1} \rightarrow Y$  для некоторых целых  $k, h \geq 0$ , и если КА  $M = \langle X, Y, Y^k \times X^{h+1}, \delta, \lambda \rangle$  может быть определён  $y(i) = \varphi(y(i-1), \dots, y(i-k), x(i), \dots, x(i-h))$ ,  $i = 0, 1, \dots$ , а именно

$$\delta(\langle y_{-1}, \dots, y_{-k}, x_{-1}, \dots, x_{-h} \rangle, x_0) = \langle y_0, \dots, y_{-k+1}, x_0, \dots, x_{-h+1} \rangle,$$

$$\lambda(\langle y_{-1}, \dots, y_{-k}, x_{-1}, \dots, x_{-h} \rangle, x_0) = \langle y_0, \dots, y_{-k+1}, x_0, \dots, x_{-h+1} \rangle,$$

$$y_0 = \varphi(y_{-1}, \dots, y_{-k}, x_0, \dots, x_{-h}),$$

тогда  $M$  называется КА с памятью порядка  $(h, k)$  и обозначается через  $M_\varphi$ . Тогда  $h$  и  $k$  называются входной и выходной памятью автомата  $M$  соответственно. В случае, когда  $k=0$ , автомат  $M_\varphi$  называется КА с входной памятью порядка  $h$ .

Пусть функция  $f: Y^k \times U^{p+1} \times X^{h+1} \rightarrow Y$ , а функция  $g: Y^k \times U^{p+1} \times X^{h+1} \rightarrow U$  для некоторых целых  $k, h \geq 0, p \geq -1$ , и если КА  $M_{f,g} = \langle X, Y, Y^k \times U^{p+1} \times X^h, \delta, \lambda \rangle$  может быть определён  $y(i) = f(y(i-1), \dots, y(i-k), u(i), \dots, u(i-p), x(i), \dots, x(i-h))$ ,

$$u(i+1) = g(y(i-1), \dots, y(i-k), u(i), \dots, u(i-p), x(i), \dots, x(i-h)), i = 0, 1, \dots$$

а именно

$$\delta(\langle y_{-1}, \dots, y_{-k}, u_0, \dots, u_{-p}, x_{-1}, \dots, x_{-h} \rangle, x_0) = \langle y_0, \dots, y_{-k+1}, u_1, \dots, u_{-p+1}, x_0, \dots, x_{-h+1} \rangle,$$

$$\lambda(\langle y_{-1}, \dots, y_{-k}, u_0, \dots, u_{-p}, x_{-1}, \dots, x_{-h} \rangle, x_0) = \langle y_0, \dots, y_{-k+1}, u_1, \dots, u_{-p+1}, x_0, \dots, x_{-h+1} \rangle,$$

$$y_0 = f(y_{-1}, \dots, y_{-k}, u_0, \dots, u_{-p}, x_0, \dots, x_{-h}),$$

$$u_1 = g(y_{-1}, \dots, y_{-k}, u_0, \dots, u_{-p}, x_0, \dots, x_{-h}),$$

тогда  $M$  называется КА с псевдо-памятью порядка  $(h, k, p)$  и обозначается через  $M_{f,g}$ .

Автомат с памятью подразделяют на линейные и нелинейные. Если функции, определяющие КА, линейны, то автомат является линейным (ЛКА). Добавление к ЛКА любой нелинейной функции приводит к нелинейному КА с памятью (НЛКА).

Рассмотрим КА в качестве распознавателя.

Если для КА  $A = \langle X, Y, S, \delta, \lambda \rangle$ ,  $|Y| = 1$ , то он называется автоматом без выхода и определяется тройкой  $\langle X, S, \delta \rangle$ .

Если  $\langle X, S, \delta \rangle$  КА без выхода,  $s_0 \in S$  – начальное состояние,  $F \subseteq S$  – конечное множество финальных состояний, тогда пятерка  $\langle X, S, \delta, s_0, F \rangle$  называется КА – распознаватель.

КА  $A = \langle X, S, \delta \rangle$  с  $S = X$  называется перестановочным автоматом, если для любых  $a, b \in S$  ( $a \neq b$ ) и  $x, y \in X$  ( $x \neq y$ ),  $\delta(a, x) \neq \delta(b, x)$  и  $\delta(a, x) \neq \delta(a, y)$ .

Модель Рабин-Скотта – это детерминированный КА – распознаватель.

### 3 Онтология Про

Модель включает в себя концепты из описанных в разделах 1 и 2 уровней и новые понятия, относящиеся непосредственно к КАКГ (см. рисунок 4). В КАКГ введены следующие концепты: композиция КА, обратимость КА, слабо обратимый КА с задержкой, обратный автомат. Эти характеристики стали основополагающими в разработке криптосистем на основе КА. Как показано на рисунке 3, криптосистема ФАРКС основана на КА Мили, а именно на КА с памятью. Основная идея ФАРКС заключается в использовании последовательной композиции слабо обратимых КА для генерации открытого ключа, тогда как секретный ключ состоит из их обратных КА. Данная семантическая связь видна на рисунке 4. Дадим определение концептов.

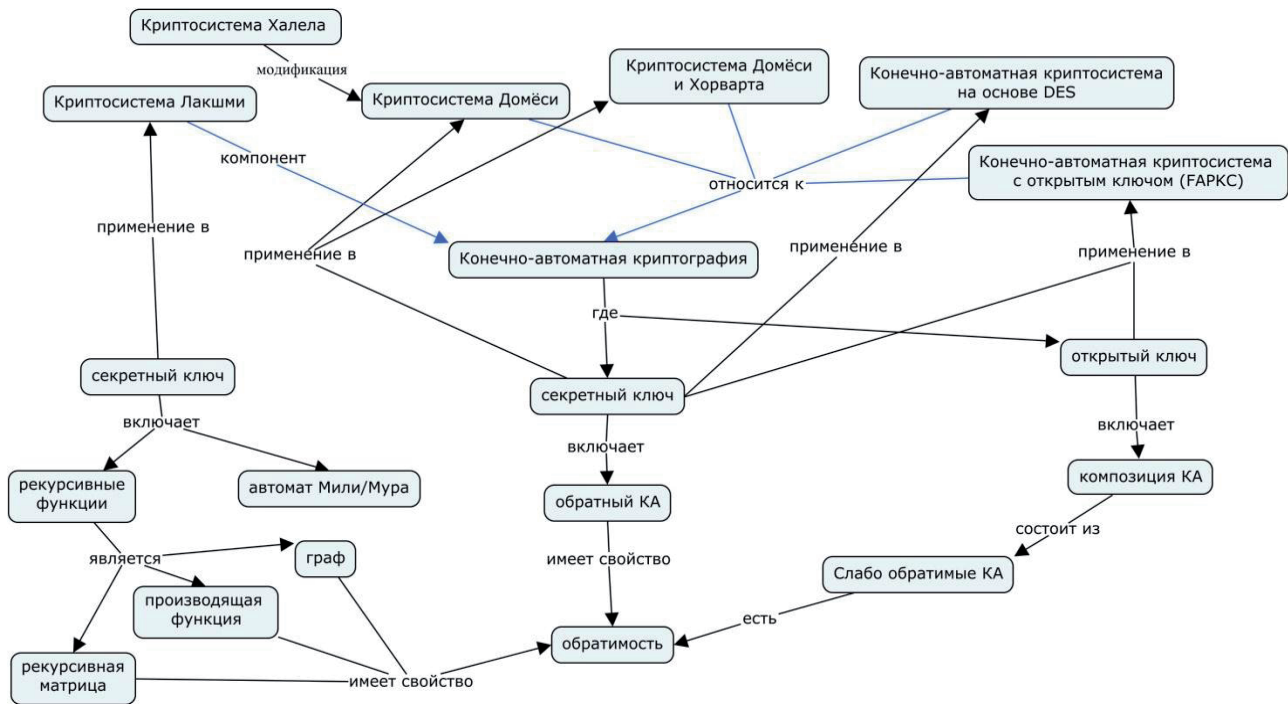


Рисунок 4 – Онтология предметной области

Пусть даны два КА  $M = \langle X, Y, S, \delta, \lambda \rangle$  и  $M' = \langle X, Y, S', \delta', \lambda' \rangle$ . КА  $M = \langle X, Y, S, \delta, \lambda \rangle$  называется *слабо обратимым с задержкой  $\tau$* , где  $\tau$  – целое неотрицательное число, если  $\forall s \in S$  и  $\forall x_i \in X, i = 0, 1, \dots, \tau$ ,  $x_0$  может быть однозначно определено состоянием  $s$  и функцией выхода  $\lambda(s, x_0 \dots x_\tau)$ .

Для  $\forall s \in S$  и  $\forall s' \in S'$ , если  $\forall a \in X^\omega, \exists a_0 \in X^k: \lambda'(s', \lambda(s, a)) = a_0 a$  и  $|a_0| = \tau$ , тогда  $(s', s)$  называется парой с задержкой  $\tau$  ( $\tau$ -парой), т.е.  $s'$  соответствует  $s$  с задержкой  $\tau$  [17].

Автомат  $M'$  называется *обратным с задержкой  $\tau$*  к автомату  $M$ , если  $\forall s \in S \exists s' \in S'$  такой что  $(s', s)$  является  $\tau$ -парой в  $M' \times M$ .

Пусть заданы два КА  $M_1 = \langle X_1, Y_1, S_1, \delta_1, \lambda_1 \rangle$  и  $M_2 = \langle X_2, Y_2, S_2, \delta_2, \lambda_2 \rangle$ , где  $X_2 = Y_1$ . Тогда композиция двух КА, определяется так:

$$C(M_1, M_2) = \langle X_1, Y_2, S_1 \times S_2, \delta, \lambda \rangle,$$

где  $\delta(\langle s_1, s_2 \rangle, x) = \langle \delta_1(s_1, x), \delta_2(s_2, \lambda_1(s_1, x)) \rangle, \lambda(\langle s_1, s_2 \rangle, x) = \lambda_2(s_2, \lambda_1(s_1, x)), x \in X_1, s_1 \in S_1, s_2 \in S_2$ .

Если функция  $g: Y_1^r \times Y_2^{p+1} \rightarrow Y_2$ , а функция  $f: X_1^{t+1} \rightarrow Y_1$ , то КА  $C'(M_f, M_g)$  с памятью порядка  $(p + t, r)$  определяется так:

$$\begin{aligned} y(i) &= g(y(i-1), \dots, y(i-r), f(x(i), \dots, x(i-t)), \dots, f(x(i-p), \dots, x(i-p-t))) \\ &= g'(y(i-1), \dots, y(i-r), x(i), \dots, x(i-p-t)), i = 0, 1, \dots \end{aligned}$$

Конечно-автоматная криптосистема на основе DES оперирует КА с такими же характеристиками, что и в криптосистеме FAPKC.

Криптосистема Домёси (P. Dömösi) схожа с криптосистемами, построенными на основе КА Мили, тем, что для шифрования и дешифрования использует ключевой автомат. Ключевой автомат представляет собой матрицу перехода перестановочного автомата без выхода. Нужно отметить, что для дешифрования используется обратный ключевой автомат. P. Dömösi спроектировал симметричную криптосистему для поточного шифрования, основанную на Рабин-Скотта модели КА. Для блочного шифрования нашли своё применение произведения автоматов Глушкова (перестановочные автоматы без выхода) [18].



Заметим, что для *обратимости* автомата необходимо и достаточно, чтобы в его табличном представлении в каждой строке таблицы переходов все состояния были различны, т.е. автомат должен быть перестановочным. Это важное свойство, которое обеспечивает однозначность зашифрованного текста для любого открытого текста. Для безопасности предполагаем, что все столбцы таблицы переходов образуют перестановку набора состояний.

Как известно, дешифрование – это обратная функция к шифрованию, соответственно необходимо определить понятие обратного автомата без выхода.

Автомат  $A^{-1} = \langle X, Q, \delta^{-1} \rangle$ , с функцией переходов  $\delta^{-1}(b, x) = a$ , где  $a, b \in Q, x \in X$  называется *обратным* к автомату  $A = \langle X, Q, \delta \rangle$  тогда и только тогда, когда  $\delta(a, x) = b$ .

Тогда для  $\forall a, b \in Q (a \neq b)$  и для  $\forall x \in X^*$  выполняется равенство  $A^{-1}(A(x)) = x$  [19].

Криптосистема, предложенная Лакшми (Lakshmi), использует автоматы Мили/Мура и рекурсивную функцию. КА является компонентом данной системы и входит в состав секретного ключа. Он предлагает использовать такие рекурсивные функции как рекуррентная матрица, производящая функция и граф, для которых определяется обратимость [20].

*Рекуррентная матрица* - это квадратная матрица порядка  $n, n > 0$ , элементы которой взяты из рекуррентного соотношения, например, это могут быть числа Фибоначчи. Данная матрица должна быть обязательно *невыврожденной*.

Производящая функция  $f(x)$  для рекуррентного отношения есть многочлен порядка  $(n-1)$  и имеет вид:  $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ .

Обратный многочлен  $F(x)$  к многочлену  $f(x)$  является многочленом степени  $(n-1)$ , если удовлетворяет следующему свойству:  $f * F = 1$ .

Пусть граф есть структура  $G = (V, E, \varphi)$ , где  $V$  – это непустое множество, элементы которого называются вершинами или узлами,  $E$  представляет собой набор из двух элементных подмножеств  $V$ , называемых ребрами, а  $\varphi$  – это функция с областью  $E$  и совместной областью  $P_2(V)$ .

Пусть функция  $f(G)$  преобразует граф в число  $n$ , а функция  $f(n)$  является обратной к функции  $f(G)$ . Матрица смежности  $A$  определяет граф  $G$ . Обратное неверно. Путём перестановки вершин группы  $G$  можно получить множество матриц смежности. Следовательно, должна быть предоставлена дополнительная информация для обеспечения инъективного свойства отображения.

#### 4 Прикладные онтологии

Прикладные онтологии КАКГ охватывает широкую область знаний и предназначены в первую очередь для того, чтобы описать концептуальную модель конкретной задачи или приложения. Данный уровень описывает концепты, зависящие от верхних уровней онтологии. В статье приведён иллюстративный пример прикладной онтологии асимметричного криптографического алгоритма на основе КА – FAPКС. Данный алгоритм включает в себя большую часть описанных концептов. FAPКС с момента создания подвергся модификациям, которые представлены на рисунке 5. Нужно отметить, что общий алгоритм FAPКС остался неизменным, менялись виды используемых в криптосистеме автоматов.

Концепция конечно-автоматной криптосистемы с открытыми ключами заключается в том, что для шифрования открытого текста и верификации подписи используется открытый ключ, который состоит из последовательной композиции обратимых автоматов, в то время как обратные им автоматы входят в состав закрытого ключа, который используется для расшифровки и подписи сообщения. Считается, что без знания секретного ключа трудно инвертировать последовательность композиции автоматов [16].

Общая схема работы алгоритма FAPКС представлена на рисунке 6.

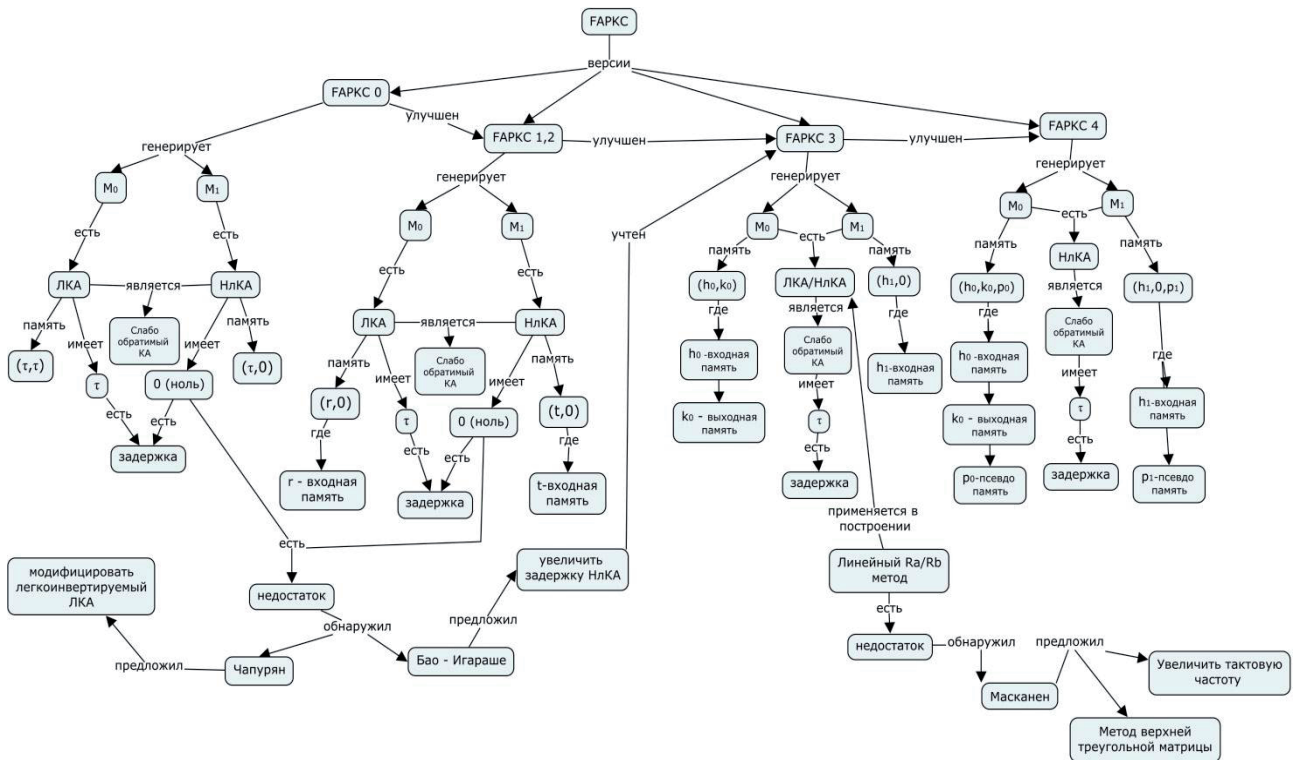


Рисунок 5 – Прикладная онтология FAPKC

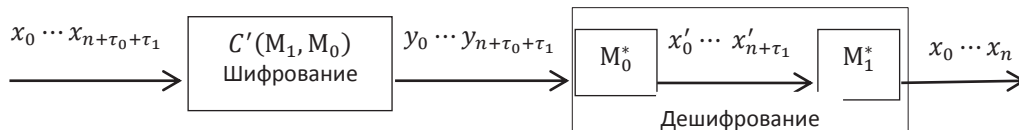


Рисунок 6 – Схема работы FAPKC

В отличие от теории чисел, где большое число можно всегда разложить на простые сомножители, для которых порядок их взаимного расположения в произведении не важен, в теории КА в композиции примитивных автоматов имеет значение как их набор, так и порядок взаимного расположения примитивных автоматов в композиции. Другими словами, композиция КА из примитивных автоматов не обладает свойством коммутативности. Поэтому разложение композиции КА на примитивные автоматы позволяет создавать сверхнадёжные системы защиты информации [21].

#### Общий алгоритм FAPKC.

В конечно-автоматной криптосистеме с открытыми ключами пользователь выбирает открытый и закрытый ключи по следующему алгоритму.

- 1) генерируются обратимый автомат  $M_0 = \langle X, Y, S_0, \delta_0, \lambda_0 \rangle$  с памятью порядка  $(r_0, t_0)$  и обратимый к нему автомат  $M_0^* = \langle Y, X, S_0^*, \delta_0^*, \lambda_0^* \rangle$  с памятью порядка  $(r_0^*, t_0^*)$  и с задержкой  $\tau_0$ .
- 2) генерируется обратимый автомат  $M_1 = \langle X, X, S_1, \delta_1, \lambda_1 \rangle$  с входной памятью порядка  $-h_1$  и обратный к нему автомат  $M_1^* = \langle X, X, S_1^*, \delta_1^*, \lambda_1^* \rangle$  памятью порядка  $(\tau_1, h_1)$  с задержкой  $\tau_1$ .
- 3) строится композиция автоматов  $M_1$  и  $M_0$   $C'(M_1, M_0) = \langle X, Y, S, \delta, \lambda \rangle$ .
- 4) определяется  $\tau$ . Выбирается произвольное состояние  $s_e$  автомата  $C'(M_1, M_0)$ , которое и будет началом шифрования. Определяются части, необходимые для дешифрования:  $s_{1,d}^{out}$  и  $s_{0,d}^{out}$ , для подписи и проверки:  $s_v^{out}, s_v^{in}, s_{0,s}, s_{1,s}$ .
- 5) открытый ключ пользователя состоит из  $\{C'(M_1, M_0), s_v^{out}, s_v^{in}, s_e, \tau_0 + \tau_1\}$ . Закрытый ключ пользователя состоит из  $\{M_0^*, M_1^*, s_{0,s}, s_{1,s}, s_{1,d}^{out}, s_{0,d}^{out}, \tau_0, \tau_1\}$ .

**Шифрование:** К концу заданного открытого текста  $x_0 \dots x_n$  добавляются произвольные символы длины  $x_{n+1} \dots x_{n+\tau_0+\tau_1}$  и вычисляется шифртекст по открытому ключу,  $y_0 \dots y_{n+\tau_0+\tau_1} = \lambda(s_e, x_0 \dots x_{n+\tau_0+\tau_1})$ .

**Дешифрование:** Открытый текст получают в два этапа. Вначале вычисляется  $x'_0 \dots x'_{n+\tau_1} = \lambda_0^*(\langle x_{-1,e}, \dots, x_{-h_0,e}, y_{\tau_0-1}, \dots, y_0, y_{-1,e}, \dots, y_{-k_0,e} \rangle, y_{\tau_0} \dots y_{n+\tau_0+\tau_1})$ .

Затем находят  $x_0 \dots x_n = \lambda_1^*(\langle x_{-1,e}, \dots, x_{-h_1,e}, x'_{\tau_1-1}, \dots, x'_0 \rangle, x'_{\tau_1}, \dots, x'_{n+\tau_1})$ .

**Подпись:** К концу сообщения  $y_0 \dots y_n$  добавляются произвольные символы длины  $y_{n+1} \dots y_{n+\tau_0+\tau_1}$ . Затем вычисляется подпись  $x_0 \dots x_{n+\tau_0+\tau_1} = \lambda_1(s_{1,s}, \lambda_0^*(s_{0,s}, x_0 \dots x_{n+\tau_0+\tau_1}))$ , используя часть секретного ключа  $M_0^*, M_1^*, s_{0,s}, s_{1,s}$ .

**Проверка:** Проверка подлинности подписи сообщения  $x_0 \dots x_{n+\tau_0+\tau_1}$  проводится использованием части открытого ключа  $C'(M_1, M_0)$ ,  $s_v^{out}$  и  $s_v^{in}$ . Вычисляется

$\lambda(\langle y_{-1,s}, \dots, y_{-t_0}, x_{\tau_0+\tau_1-1}, \dots, x_0, x_{-1,s}, \dots, x_{-r_0-r_1+\tau_0+\tau_1,s} \rangle, x_{\tau_0+\tau_1} \dots x_{n+\tau_0+\tau_1})$ ,

которое должно совпадать с сообщением  $y_0 \dots y_n$ .

В таблице 1 показаны характеристики версий криптосистемы FAPKC.

Таблица 1 – Характеристики криптосистемы FAPKC

FAPKC версии	Вид автомата	$(r_0, t_0)$	$(r_0^*, t_0^*)$	За-держка	Формальное представление
FAPKC0	М <sub>0</sub> -линейный КА	$(\tau, \tau)$	$(\tau, 0)$	$\tau$	$y(i) = \sum_{j=1}^{\tau} A_j y(i-j) + \sum_{j=0}^{\tau} B_j x'(i-j), i = 0, 1, 2, \dots$
	Нелинейная функция $f$	$(r+1)$		0	$f(v_0, \dots, v_r)   \forall v_1, \dots, v_r, f(v_0, \dots, v_r)$ – обратима от аргумента $v_0$
FAPKC1	М <sub>0</sub> -линейный КА	$(r, 0)$	$(\tau, r)$	$\tau$	$x'(i) = g(y'(i-r), \dots, y(i)), i = 0, 1, \dots$
	М <sub>1</sub> -нелинейный КА	$(t, 0)$	$(0, t)$	0	$y'(i) = f(y'(i-t), \dots, y(i)), i = 0, 1, \dots$
FAPKC2	М <sub>0</sub> -линейный КА	$(r, 0)$	$(r, r)$	$r$	$x'(i) = g(y'(i-r), \dots, y(i)), i = 0, 1, \dots$
	М <sub>1</sub> -нелинейный КА	$(t, 0)$	$(\tau, r)$	$\tau$	$y'(i) = f(y'(i-t), \dots, y(i)), i = 0, 1, \dots$
FAPKC1 (Bao, Igarashe)	М <sub>0</sub> -линейный КА	$(r, t)$	$(t+\tau, r)$	$\tau$	$y(i) = \sum_{j=0}^r A_j x(i-j) + \sum_{j=1}^{\tau} B_j y(i-j), i = 0, 1, 2, \dots$
	М <sub>1</sub> -нелинейный КА	$(r, 0)$	$(\tau, r)$	$\tau$	$y(i) = \sum_{j=0}^{\tau} A_j x(i-j) + \sum_{j=1}^{t-1} B_j y(i-j), i = 0, 1, 2, \dots$
FAPKC3	М <sub>0</sub> -линейный/нелинейный КА	$(h_0, k_0)$	$(k_0 + \tau_0, h_0)$	$\tau_0$	$y(i) = \sum_{j=1}^{k_0} A_j y(i-j) + \sum_{j=0}^{h_0} B_j x(i-j), i = 0, 1, 2, \dots$ $y(i) = \sum_{j=1}^{k_0} A_j y(i-j) + \sum_{j=0}^{h_0} B_j x(i-j) + \sum_{j=0}^{h_0-\epsilon} B'_j s(x(i-j), \dots, x(i-j-\epsilon)),$
	М <sub>1</sub> -нелинейный КА	$(h_1, 0)$	$(\tau_1, h_0)$	$\tau_1$	$x'(i) = \sum_{j=0}^{h_1} F_j x(i-j) + \sum_{j=0}^{h_1-\epsilon} F'_j s(x(i-j), \dots, x(i-j-\epsilon)), i = 0, 1, 2, \dots$ где $s(x(i-j), \dots, x(i-j-\epsilon))$ - нелинейная функция, а $\epsilon$ – маленькое положительное целое число.
FAPKC3 (Meskanen)	М <sub>0</sub> -линейный	$(h_0, k_0)$	$(k_0 + \tau_0, h_0)$	$\tau_0$	$y(i) = \sum_{j=1}^{k_0} A_j y(i-j) + \sum_{j=0}^{h_0} B_j x(i-j) + \sum_{j=h_0+1}^{2h_0} B_{2h_0-j} x(i-j), i=0, 1, 2, \dots$
	М <sub>1</sub> -нелинейный КА	$(h_1, 0)$	$(\tau_1, h_0)$	$\tau_1$	$x'(i) = \sum_{j=0}^{h_1} F_j x(i-j) + \sum_{j=0}^{h_1-\epsilon} F'_j s(x(i-j), \dots, x(i-j-\epsilon)), i = 0, 1, 2, \dots$
FAPKC4	М <sub>0</sub> -линейный/нелинейный КА	$(h_0, k_0, p_0)$	$(\tau_0 + k_0, h_0, p_0)$	$\tau_0$	$y(i) = f(y(i-1), \dots, y(-k_0), u(i), \dots, u(i-p_0), x(i), \dots, x(-h_0)),$ $u(i+1) = g(y(i-1), \dots, y(-k_0), u(i), \dots, u(i-p_0), x(i), \dots, x(i-h_0)),$ $i = 0, 1, \dots$
	М <sub>1</sub> -нелинейный КА	$(h_1, 0, p_1)$	$(\tau_1, h_1, p_1)$	$\tau_1$	$y(i) = f(u(i), \dots, u(i-p_1), x(i), \dots, x(i-h_1)),$ $u(i+1) = g(u(i), \dots, u(i-p_1), x(i), \dots, x(i-h_1)), i = 0, 1, \dots$

В версии FAPKC0, приведённой в [22], открытый ключ содержит составной КА из обратимого линейного автомата с памятью порядка  $(\tau, \tau)$  и с задержкой  $\tau$  и слабо обратимого нелинейного КА с входной памятью и с задержкой 0. Две другие схемы FAPKC1 и FAPKC2 приведены в [23], где открытый ключ для FAPKC1 содержит композицию двух КА: обратимый линейный автомат с входной памятью порядка  $\tau$  и с задержкой  $\tau$ , слабо обратимый нелинейный КА с входной памятью и с задержкой 0. В работе [24] доказано, что FAPKC1 небезопасен в шифровании и предлагается модификация с использованием квазилинейных КА. В [23] разработан метод генерации своего рода нелинейных слабо обратимых КА; затем две схемы, названные FAPKC3 и FAPKC4, были предложены в [25, 26]. Для версий FAPKC3 были предложены иные модификации в работе [17].

## Заключение

Актуальность создания онтологии КАКГ обусловлена необходимостью в систематизации полученных знаний в криптографии. Цель предложенной онтологической модели - формирование наглядной когнитивной модели, которая отражает все основные концепты криптосистем, основанных на теории автоматов, и характер их связей. Данная модель может улучшить понимание такого рода криптосистем при планировании их реализации.

Для формирования онтологии КАКГ использована методология К-карт, которая обладает такими преимуществами как системность, единообразие, научность и когнитивность.

Структура области КАКГ рассматривается как сложно-структурированный объект, который получен путём интеграции двух самостоятельных областей - криптографии и теории автоматов. Использована четырёхуровневая онтологическая модель, включающая представление знаний – методологию К-карт, онтологии криптографии и теории автоматов, онтологию КАКГ, прикладные онтологии.

Предложенную онтологию можно использовать для решения таких задач:

- обеспечение общим понятийным и терминологическим аппаратом специалистов, интересующихся данной ПрО;
- создание интеллектуальных систем принятия решения в криптографической защите информации;
- организация эффективного поиска информации о КАКГ.

## Список источников

- [1] *Гатченко, Н.А.* Криптографическая защита информации / Н.А. Гатченко, А.С. Исаев, А.Д. Яковлев - СПб: НИУ ИТМО, 2012. -142 с.
- [2] *Суцевский, Д.Г.* Современные криптосистемы и их особенности / Д.Г. Суцевский, О.В. Панченко, В.Н. Кугураков // Вестник технического университета. – 2015. – №11(18). – С.194-197.
- [3] *Мирзагитов, А.А.* Методы разработки онтологии по информационной безопасности, основанные на прецедентном подходе / А.А. Мирзагитов, Д.Е. Пальчунов // Вестник Новосиб. гос. ун-та. Серия: Информационные технологии. – 2013. – №3(11). – С. 37-46.
- [4] *Cañas, A.J.* SMARTTOOLS: a knowledge modeling and sharing environment / A.J. Cañas, G. Hill, R. Carff, N. Suri, J. Lott, G. Gómez, Th. C. Eskridge, M. Arroyo, R. Carvajal. – <http://cmc.ihmc.us/papers/cmc2004-283.pdf>.
- [5] *Gruber, T.R.* A Translation Approach to portable ontology specifications/ T.R. Gruber // Knowledge acquisition. – 1993. - №5 (2). – P. 199-220.
- [6] *Ивлеев, А.А.* Онтология военных технологий: основы, структура, визуализация и применение (1 часть) / А.А. Ивлеев, В.Б. Артеменко // Вооружение и экономика. – 2011. - №4 (16). – С.35-52 - <http://www.viek.ru/16/35-52.pdf>
- [7] *Novak, J.D.* The Theory Underlying Concept Maps and How to Construct and Use Them / J.D. Novak, A.J. Cañas. - <http://cmap.ihmc.us/docs/theory-of-concept-maps>.



- [8] **Гаврилова, Т.А.** Анализ ошибок студентов при визуальном структурировании знаний / Т.А. Гаврилова, В.А. Онуфриев // Компьютерные инструменты в образовании. – 2016. – № 6. – С.42-54 – <http://ipo.spb.ru/journal>.
- [9] **Бачурина, Е.П.** Стратегия онтологического инжиниринга при управлении знаниями в области ЧС / Е.П. Бачурина // Тез. VI Всерос. науч.-техн. конф. студентов, аспирантов и молодых ученых «Молодежь и наука». – Красноярск, 2010. – С. 304-308.
- [10] **Муромский, А.А.** Использование онтологического подхода для защиты данных при их пересылке и архивации / А.А. Муромский, Н.П. Тучкова // Онтология проектирования. – 2016. – Т. 6, №2(20). – С. 136-148. – DOI: 10.18287/2223-9537-2016-6-2-136-148.
- [11] **Саломая, А.** Криптография с открытым ключом / А. Саломая. Перевод с английского И.А. Вихлянцева под редакцией А.Е. Андреева и А.А. Болотова – М.: Мир, 1995. – 318 с.
- [12] **Гуренко, В.В.** Введение в теорию автоматов / В.В. Гуренко. – М.: МГТУ им. Н.Э. Баумана, 2013. – 62 с.
- [13] **Шарипбай, А.А.** Теория языков и автоматов / А.А. Шарипбай. – Алматы, Эверо, 2015. – 207 с.
- [14] **Трахтенброт, Б.А.** Конечные автоматы (поведение и синтез) / Б.А. Трахтенброт, Я.М. Барздинь. – Москва, Наука, 1970. – 400 с.
- [15] **Шарипбай, А.А.** Автоматные модели в криптографии / А.А. Шарипбай // Вестник КазНУ. Серия: математика, механика, информатика. – 2016. – №3/1 (90). – С.96-104.
- [16] **Tao, R.J.** Finite Automata and Application to Cryptography/ R.J. Tao. – Tsinghua University Press, 2009. – 406 p.
- [17] **Meskanen, T.** On finite automaton public key cryptosystems / T. Meskanen // TUCS Technical Report. – Turku, 2001. – No.408. – 46 p.
- [18] **Dömösi, P.A.** Novel Cryptosystem Based on Gluskov Product of Automata / P. Dömösi, G. Horváth // Acta Cybernetica. – 2015. – P.359–371.
- [19] **Dömösi, P.A.** Novel Cryptosystem Based on Finite Automata Without Outputs / P. Dömösi // Automata, Formal Languages and Algebraic Systems. – 2010. – P.23-32.
- [20] **Gandhi, B.K.** Cryptographic Scheme for Digital Signals using Finite State Machines / B.K. Gandhi, A. Ch. Sekhar, S.S. Lakshmi // International Journal of Computer Applications. – 2011. – №6 (29). – P. 61-63.
- [21] **Шахметова, Г.Б.** Применение конечных автоматов для разработки асимметричных шифров / Г.Б. Шахметова, А.А. Шарипбай, Ж.С. Сауханова, Г.Ж. Исабаева // Труды V Международной научно-практической конференции «Интеллектуальные информационные и коммуникационные технологии – средство осуществления третьей индустриальной революции в свете стратегии «КАЗАХСТАН-2050». – 2018. – С.286-289.
- [22] **Tao, R.** A finite automaton public key cryptosystem and digital signatures [китайский]/ R. Tao, Sh. Chen // Chinese Journal of Computers. – 1985. Vol.8(6). – P. 401-409.
- [23] **Tao, R.** Two varieties of finite automaton public key cryptosystem and digital signatures / R. Tao, Sh. Chen // Journal of computer science and technology. – 1986. Vol. 1(1). – P.9-18.
- [24] **Bao, F.**, Break finite automata public key cryptosystem / F. Bao, Y. Igarashi. // In International Colloquium on Automata, Languages, and Programming, Springer Berlin Heidelberg. – Berlin, 1995. – P.147-158.
- [25] **Tao, R.** FAPKC3: a new finite automaton public key cryptosystem / R. Tao, Sh. Chen, X. Chen // Journal of Computer science and Technology. – 1997. – Vol.12(4). – P.289-305.
- [26] **Tao, R.** The generalization of public key cryptosystem FAPKC4 / R. Tao, Sh. Chen // Chinese science bulletin. – Chinese, 1999. – Vol.44(9). – P.784-790.

## ONTOLOGY OF FINITE-AUTOMATION CRYPTOGRAPHY

**A.A. Sharipbay<sup>1</sup>, Zh.S. Saukhanova<sup>2</sup>, G.B. Shakhmetova<sup>3</sup>, M.S. Saukhanova<sup>4</sup>**

*Eurasian national university named after L.N. Gumilyov, Astana, Kazakhstan*

<sup>1</sup> [sharalt@mail.ru](mailto:sharalt@mail.ru), <sup>2</sup> [saukhanova@mail.ru](mailto:saukhanova@mail.ru), <sup>3</sup> [shakhmetova.gb@gmail.com](mailto:shakhmetova.gb@gmail.com), <sup>4</sup> [m.saukhanova@mail.ru](mailto:m.saukhanova@mail.ru)

### Abstract

Today, the improvement of methods for protecting information is an urgent task, therefore, the authors of the article are interested in using alternative methods of developing more robust and efficient public-key cryptosystems. Finite automata are taken as such a model. To systematize knowledge in the field of finite-state cryptography, it was decided to use ontology. This paper discusses the construction of an ontological model of finite-automaton cryptography. The proposed ontological model will have four main levels: the ontology of knowledge representation, the ontology of the top



level, the domain ontology and practical ontologies. The methodology of conceptual maps was taken as the ontology of knowledge representation. The ontology of the top level contains basic information about cryptography and automata theory, the domain ontology describes directly the finite automaton cryptography. As an example of practical ontology, a public-key cryptosystem based on finite automata was considered. The presented ontology is modeled for the first time and gives a clear understanding of the use of finite automata in cryptography, systematizes the information obtained in the course of the research on this subject area, is a prerequisite in the further development of cryptosystems based on the theory of automata.

**Key words:** *ontology, conceptual map, finite automata, cryptography.*

**Citation:** Sharipbay AA, Saukhanova ZhS, Shakhmetova GB, Saukhanova MS. Ontology of finite-automation cryptography [In Russian]. *Ontology of designing*. 2019. 9(1): 36-49. – DOI: 10.18287/2223-9537-2019-9-1-36-49.

## References

- [1] **Gatchenko NA, Isaev AS, Yakovlev AD.** Cryptographic protection of information: textbook [In Russian]. Petersburg: NRU ITMO; 2012.
- [2] **Sushevsky DG, Panchenko OV, Kugurakov VN.** Modern cryptosystems and their features [In Russian]. Bulletin of the Technical University 2015; 11(18): 194-197.
- [3] **Mirzagitov AA, Palchunov DE.** Methods for developing ontologies for information security based on a case-law approach [In Russian] Bulletin of Novosib. state un-that. Series: Information Technology 2013; 3(11): 37-46.
- [4] **Cañas AJ, Hill G, Carff R, Suri N, Lott J, Gómez G, Eskridge ThC, Arroyo M, Carvajal R.** CMAPTOOLS: a knowledge modeling and sharing environment. - <http://cmc.ihmc.us/papers/cmc2004-283.pdf>.
- [5] **Gruber TR.** A Translation Approach to portable ontology specifications: Proc. of Knowledge acquisition 1993; 5(2): 199-220.
- [6] **Ivlev AA, Artemenko VB.** Ontology of military technologies: fundamentals, structure, visualization and application (1 part) [In Russian] Armament and economy 2011; 4(16): 35-52. - <http://www.viek.ru/16/35-52.pdf>.
- [7] **Novak JD, Cañas AJ.** The Theory Underlying Concept Maps and How to Construct and Use Them - <http://cmap.ihmc.us/docs/theory-of-concept-maps>.
- [8] **Gavrilova TA, Onufriev VA.** Analysis of students' mistakes in the visual structuring of knowledge [In Russian] Computer tools in education 2016; 6: 42-54. - <http://ipo.spb.ru/journal>.
- [9] **Bachurina EP.** Strategy of ontological engineering in knowledge management in the field of emergency situations [In Russian]. Thesis of VI All-Russian scientific and technical conf. students, graduate students and young scientists "Youth and Science." – Krasnoyarsk; 2010: 304-308.
- [10] **Muromskii AA, Tuchkova NP.** Ontological approach to the data protection for their transfer and archiving [In Russian] *Ontology of designing*. 2016; 2(20): 136-148. DOI: 10.18287/2223-9537-2016-6-2-136-148.
- [11] **Salomaa A.** Public-key Cryptography. - Springer-Verlag; 1996.
- [12] **Gurenko VV.** Introduction to Automata Theory [In Russian]. - Moscow: Moscow State Technical University named after NE Bauman; 2013.
- [13] **Sharipbay AA.** Theory of languages and automata [In Russian]. - Almaty, Evero; 2015.
- [14] **Trakhtenbrot BA., Barzdin YaM.** State machine (behavior and synthesis) [In Russian]. – Moscow, Nauka; 1970.
- [15] **Sharipbay AA.** Automata models in cryptography [In Russian] KazNU Bulletin. Mathematics, Mechanics, Computer Science Series. 2016; 3/1(90): 94-104.
- [16] **Tao RJ.** Finite Automata and Application to Cryptography. – Tsinghua University Press; 2009.
- [17] **Meskanen T.** On finite automaton public key cryptosystems. *TUCS Technical Report*. – Turku; 2001.
- [18] **Dömösi P, Horváth G.** A Novel Cryptosystem Based on Gluskov Product of Automata. *Acta Cybernetica*; 2015: 359–371.
- [19] **Dömösi P.** A Novel Cryptosystem Based on Finite Automata Without Outputs. *Automata, Formal Languages and Algebraic Systems*; 2010: 23-32.
- [20] **Gandhi BK, Sekhar ACh, Lakshmi SS.** Cryptographic Scheme for Digital Signals using Finite State Machines. *International Journal of Computer Applications*. 2011; 6 (29): 61-63.
- [21] **Shakhmetova GB, Sharipby AA, Saukhanova ZhS, Isabaeva GZh.** The use of finite automata for the development of asymmetric ciphers [In Russian]. Proceedings of the V International Scientific and Practical Conference "Intelligent Information and Communication Technologies - the medium of the third industrial revolution in the light of the strategy" KAZAKHSTAN-2050 "; 2018: 286-289.
- [22] **Tao R, Chen Sh.** A finite automaton public key cryptosystem and digital signatures [In Chinese]. Chinese Journal of Computers 1985; 8(6): 401-409.

- [23] **Tao R, Chen Sh.** Two varieties of finite automaton public key cryptosystem and digital signatures. *Journal of computer science and technology* 1986; 1(1): 9-18.
- [24] **Bao F, Igarashi Y.** Break finite automata public key cryptosystem. In *International Colloquium on Automata, Languages, and Programming*, Springer Berlin Heidelberg (Berlin, 1995); 1995: 147-158.
- [25] **Tao R, Chen Sh, Chen X.** FAPKC3: a new finite automaton public key cryptosystem. *Journal of Computer science and Technology*. 1997; 12(4): 289-305.
- [26] **Tao R, Chen Sh.** The generalization of public key cryptosystem FAPKC4. *Chinese science bulletin*. 1999; 44(9): 784-790.

## Сведения об авторах

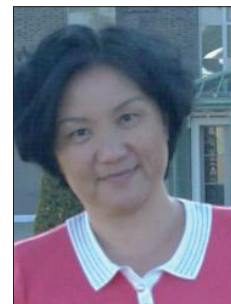


**Шарипбай Алтынбек Амирович**, 1952 г. рождения. К.ф.-м.н., д.т.н., профессор по группе специальностей «Информатика, вычислительная техника и управление», академик «Международной Академии информатизации», академик «Академии педагогических наук Республики Казахстан». Профессор кафедры «Информатика и информационная безопасность», директор НИИ «Искусственный интеллект» Евразийского национального университета им. Л.Н. Гумилева (ЕНУ). В списке научных трудов более 200 работ в области искусственного интеллекта, компьютерной лингвистики, информационной безопасности.

**Altynbek Amirovich Sharipbay**, born in 1952. Candidate of Physical and Mathematical Sciences, Doctor of Technical Sciences, Professor in the group of specialties "Informatics, computer engineering and control", academician of the "International Academy of Informatization", academician of the "Academy of Pedagogical Sciences of the Republic of Kazakhstan". Professor of the Department "Informatics and Information Security", Director of the Research Institute "Artificial Intelligence" of the Eurasian National University. L.N. Gumilyov. In the list of scientific works more than 200 works in the field of artificial intelligence, computational linguistics, information security.

**Сауханова Жанат Сергазиевна**, 1964 г. рождения. Окончила Санкт-Петербургский государственный университет, факультет прикладной математики - процессов управления. К.ф.-м.н., доцент кафедры «Информатика и информационная безопасность» ЕНУ им. Л.Н. Гумилева. В списке научных трудов более 50 работ в области программирования, больших данных, информационной безопасности.

**Zhanat Sergazievna Saukhanova**, born in 1964. She graduated from St. Petersburg State University, faculty of applied mathematics - management processes. Candidate of Physical and Mathematical Sciences, Associate Professor of the Department "Informatics and Information Security" of the Eurasian National University. L.N. Gumilyov. In the list of scientific works more than 50 works in the field of programming, big data, information security.



**Шахметова Гульмира Балтабаевна**, 1982 г. рождения. Окончила магистратуру ЕНУ им. Л.Н. Гумилева по специальности «Информатика». Является докторантом 2 курса специальности «Информатика» ЕНУ им. Л.Н. Гумилева. В списке научных трудов более 15 статей в области теории автоматов и формальных языков, информационной безопасности, ИКТ в образовании.

**Gulmira Baltabaevna Shakhmetova**, born in 1982. She graduated from the magistracy of the Eurasian National University. L.N. Gumilyov specialty "Computer Science". She is a PhD student of the 2 course of the specialty "Computer Science" of the ENU L.N. Gumilyov. The list of scientific works includes more than 15 articles in the field of automata theory and formal languages, information security, ICT in education.



**Сауханова Магрипа Сергазиевна**, 1962 г. рождения. Окончила магистратуру ЕНУ им. Л.Н. Гумилева по специальности «Информатика». Старший преподаватель кафедры «Вычислительная техника и программное обеспечение» Евразийского национального университета им. Л.Н. Гумилева. В списке научных трудов более 20 статей в области ИКТ в образовании, автоматного программирования, информационной безопасности.

**Magripa Sergazievna Saukhanova**, born in 1962. She graduated from the magistracy of the Eurasian National University L.N. Gumilyov specialty "Informatics". She is a senior lecturer in the department of "Computing equipment and software" of the Eurasian National University L.N. Gumilyov. In the list of scientific works more than 20 articles in the field of ICT in education, automated programming, information security.



УДК 004.421.2

## ПРЕДСТАВЛЕНИЕ МАТЕМАТИЧЕСКИХ ПОНЯТИЙ В ОНТОЛОГИИ НАУЧНЫХ ЗНАНИЙ

А.А. Муромский<sup>1</sup>, Н.П. Тучкова<sup>2</sup>

Вычислительный центр им. А.А. Дородницына ФИЦ ИУ РАН, Москва, Россия

<sup>1</sup>mirom@ccas.ru, <sup>2</sup>tuchkova@ccas.ru

### Аннотация

Рассматриваются особенности математических предметных областей в контексте построения онтологии научных предметных областей. Приводятся примеры из разрабатываемого тезауруса по уравнениям смешанного типа из раздела уравнений с частными производными математической физики. Особое внимание уделяется вопросам представления и актуальной поддержки тезауруса адресата. Предлагается использовать специализированные тезаурусы, где наряду с определениями на естественном языке присутствуют символьные выражения. Этот подход позволяет уточнить запрос, используя математическую запись в TeX-нотации независимо от языка первоисточника, что приводит к уменьшению поискового шума и сокращению времени поиска. Обсуждается вопрос о том, почему всё ещё недостаточно создано математических тезаурусов и что надо сделать, чтобы они использовались в цифровых библиотеках наряду с другими информационными ресурсами. Приводится пример создания тезауруса для отдельной математической предметной области. Особая трудность состоит в описании родо-видовых отношений в математических предметных областях, предлагается один из вариантов описания этих связей для уравнения смешанного типа.

**Ключевые слова:** поисковый запрос, сравнение математических текстов, семантический поиск, тезаурус математических нотаций, TeX-нотации.

**Цитирование:** Муромский, А.А. Представление математических понятий в онтологии научных знаний / А.А. Муромский, Н.П. Тучкова // Онтология проектирования. – 2019. – Т.9, №1(31). – С. 50-69. – DOI: 10.18287/2223-9537-2019-9-1-50-69.

### Введение

Почему всё ещё недостаточно математических тезаурусов и что надо сделать, чтобы они использовались в цифровых библиотеках наряду с другими информационными ресурсами (ИР)? От классиков известно, что «математика - царица наук» (Фридрих Гаусс) и «язык, на котором Бог написал вселенную» (Галилео Галилей). Воспользоваться этим языком могут лишь профессионалы, поскольку уникальные математические знания излагаются на языке формул. В то же время математические знания востребованы всеми естественными науками и существует необходимость их систематизации и поиска.

Кеннету Мэю (Kenneth O. May) принадлежит авторство уникального цифрового проекта по созданию словарей и тезаурусов по математике совместно с издательством McGraw-Hill. Ранее тезаурусы рассматривались, в основном, как инструменты для языковых словарей, например, тезаурус английских слов Роже [1], который переиздаётся уже больше столетия (впервые опубликован в 1852 году). Основной и очевидный для многих тезис К.О. Мэя, прозвучавший на одной из конференций, гласит: «Вы не можете сделать научную работу и систематически работать ни в какой-либо области без словарей в этой области» [2]. Этот тезис сохранил актуальность и приобрёл новое практическое значение в эпоху интернет-технологий и междисциплинарных исследований. В цифровой век появилась возможность

обработки символьной информации для презентации в Интернете (MathML, Coristo WordXML и др.). Ставятся задачи поддержки пользователей в процессе обучения, поиска и взаимодействия с компьютером в области математических наук, машинного обучения и формализации математических знаний. Последней тематике посвящены, например, разработки научных групп Университета Кэмбриджа и Технологического университета Мюнхена, которые ведутся с середины 80-х годов XX в. [3]. Аналогичные исследования ведутся в отечественных проектах [4-5]. Значительное число работ посвящено проблемами представления знаний в Интернете и связанных с ними развитием онтологий [6-7] и предметных тезаурусов [8]. Обзор глобальных цифровых математических библиотек даётся в работе [9].

Для обеспечения востребованности ресурса информационные системы должны также отражать реальное место математических наук в системе научных знаний и предлагать настраиваемый на пользователя интерфейс. Опираясь на тезис, сформулированный ещё Муэрсом (Moore) [10], необходимо отметить, что использование информационно-поисковой системы должно приносить пользу авторам.

Современные средства представления предметных областей (ПрО) на основе построения онтологий и технологий социальных систем обеспечивают такие возможности. Трудности заключаются в описании внутренних взаимосвязей в самих ПрО, поскольку это подразумевает совместные междисциплинарные исследования экспертов ПрО и информационных технологий (ИТ). Каждый такой ИР уникален в своём роде, поскольку отражает индивидуальные взгляды на математическую область, традиции математических школ и подходов к рассмотрению задач и их приложений. Известно, также, что тренд на экспоненциальный рост информации в Интернете распространяется и на математические публикации, что формирует проблему информационного шума при поиске. Это особенно негативно отражается на эффективности использования ИР в научной работе. Составление обзоров при подготовке публикаций стало практически такой же частью научных исследований, как и сами научные исследования. Хотя обзоры всегда необходимы для подтверждения актуальности предлагаемых исследований, сегодня, во многом, они опираются на анализ интернет-изданий. Процесс поиска научной литературы занимает значительное количество времени в отличие от ситуации, когда было достаточно ознакомиться с реферативными журналами, чтобы составить представительный список цитирований по изучаемой проблеме. Поиск и анализ данных в интернет-ресурсах становится самостоятельным видом деятельности [11].

ИТ не прекращают двигаться по пути предоставления поисковых средств для научных исследований. Это оправданно, поскольку только сам специалист, исследователь может адекватно сформулировать свой информационный запрос, и необходимо предоставить ему такую возможность. Исследования не теряют своей актуальности и складываются из нескольких направлений:

- использование символьной информации для передачи данных;
- представление формул в системах компьютерной алгебры;
- системы доказательства теорем и поиск доказательных (reasoning) данных;
- разработка методов и технологий представления математических знаний в цифровых библиотеках;
- создание платформ для поддержки пользователей в изучении математических наук;
- поиск математических публикаций и взаимодействие учёных в процессе исследований;
- создание систем поддержки актуальных математических знаний в научном интернет-пространстве и управления этими данными на основе семантических связей ПрО;
- создание и поддержка цифровых библиотек математических публикаций.

Для формирования информационных образов разделов математики необходимо учитывать перечисленные аспекты. А именно, использовать представительные словари с толкова-



нием терминов и формул и связей между ними, как основы информационно-поискового тезауруса данной ПрО. В работе [12] был представлен тезаурус для ПрО «обыкновенные дифференциальные уравнения» (ОДУ), и теперь разрабатывается его расширение на область «уравнения с частными производными» (Partial Differential Equations, PDE) как части общего математического ресурса по «уравнениям математической физики и смежных областей» (Equations of Mathematical Physics, EMPh) [13]. Тестовая реализация разрабатывается на базе цифровой библиотеки LibMeta [14]. Для накопления персональных научных знаний о ПрО используется понятие тезауруса адресата (см. работу [15]).

## **1 Особенности информационных образов математических ПрО**

Интерес к проблеме использования символьной записи в виде формул в поисковом запросе сформировался достаточно давно. В работе [9] даётся обзор источников, посвящённых основным тенденциям исследований в этой области: использованию индексирования публикаций на основе таксономии и включению формул в поисковый запрос. В целом, разделяют проблемы представления формул для хранения, визуализации и поиска публикаций с формулами. Изучаются методы кодирования математических знаний, анализа символьных выражений, индексации математических статей ключевыми словами и формулами. Разрабатываются системы контекстного поиска с учётом таксономии. Это направление, которое сочетает семантические связи и рекурсивный поиск индексированных документов с формулами, оценивается как наиболее развиваемое и актуальное. Такой подход предполагает создание предварительной структуры данных и разметку документов для быстрого поиска и дальнейшего уточнения запроса по предварительной выборке.

Можно отметить отличительные особенности математических публикаций, такие как наличие формул и структуру изложения (утверждения, доказательства). Элементы предварительной аналитико-синтетической обработки математических текстов кроме создания словарей, тезаурусов и наборов ключевых слов для индексации, дополняются списками формул и связей между ними, а также утверждениями, где они встречаются. Связи утверждений и формул диктуются логикой самой математической ПрО. Некоторые технические аспекты математических цифровых библиотек отражены в работе [16].

### **1.1 О формулах и поиске по формулам**

Формула - понятие, выражающее разные значения. Трудно назвать область, где это понятие не употребляется или заменяется другим термином. Формула, как некая запись, совокупность знаков, имеющих определённый смысл (значение), имеет довольно широкое употребление. Например, запись «формула поведения» отношения к математике не имеет, если поведение, как некое действие, не является объектом математического описания. Формула - понятие, которое образуется в результате процесса деятельности в науке (математике, химии и др.). Математическая формула имеет входные данные и выходные, получаемые в результате математических операций над входными данными. Поэтому каждая формула может быть представлением определённого процесса, если имеется соответствующая интерпретация данных. Для интерпретации математических знаний часто прибегают к теории графов [17].

Всякая математическая запись, выражающая определённый смысл, может восприниматься как формула. Приведём некоторые определения, чтобы избежать излишних рассуждений на тему термина «формула» и его трактовок. Термин «формула» иностранного происхождения (от лат. formula), определения из [18-20] основаны на определениях из ряда энциклопедий, изданных ранее.

Формула - это:



- точное определение какого-либо правила, отношения, закона, приложимое в определённых условиях ко всем частным случаям; всякое определение, выраженное в краткой форме; совокупность величин, выраженных числами и буквами и соединённых посредством математических знаков; сокращённое обозначение состава какого-либо соединения с помощью букв, заменяющих названия элементов, и чисел, указывающих на количественное соотношение этих элементов в данном соединении [18];
- общее краткое и точное выражение (мысли, закона), определение [19];
- норма, масштаб, схема, образец, правило, по которому что-либо делают (вначале термин имел геометрическое содержание, его корень – forma) [20].

В англоязычном варианте некоторые из определений лексической базы данных Princeton's WordNet [21] отражают термин «формула» как математическое выражение:

*Formula*, expression (noun) a group of symbols that make a mathematical statement; a standard procedure for solving a class of mathematical problems.

Синонимы: *formulation, aspect, rule, chemical formula, saying, look, verbal expression, expression, manifestation, verbalism, grammatical construction, pattern, face, convention, locution, normal, reflexion*.

Приведённые определения позволяют использовать при составлении поискового запроса *словесный контекст формулы*. Синонимы составляют множество выражений для этого контекста наряду с другими содержательными текстовыми фрагментами.

Для поиска на естественном языке можно употреблять то, что можно по аналогии назвать *словесной формулой*, выражающей *описание объекта* поиска. *Набор ключевых слов* с высказанной выше точки зрения – *словесная формула* в данной тематике.

Структура математической формулы, т.е. её символьная запись, для разных случаев может быть одна и та же, различны только интерпретации входных и выходных данных, кроме, естественного временного параметра или входного данного. Рассмотрим примеры.

*Пример 1:*  $u_{tt} = a^2 u_{xx}$ , или в TeX-нотации [22]:  $\$ u_{tt} = a^2 u_{xx} \$$  – уравнение свободных поперечных колебаний струны.

*Пример 2:* уравнение Лаврентьева-Бицадзе;  $u_{xx} + \text{sign } y \cdot u_{yy} = 0$ ;

$\$ u_{xx} + \text{sign } y \cdot u_{yy} = 0 \$$ .

В поисковых системах, поддерживающих формулы в TeX-нотации, как правило, символьные выражения сравниваются с точностью до переменных.

Если формула имеет общепризнанное название, например, «формула бинома», «Даламбера», то тем самым её *входным и выходным данным назначена определённая интерпретация в зависимости от её тематического применения*.

Именные формулы также могут быть записаны по-разному.

*Пример 3:* «уравнение Трикоми»; «уравнение Т»;  $y Z_{xx} + Z_{tt} = 0$ ,  $y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ ;

$\$ y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \$$ .

Если формула не имеет общепризнанного названия (обозначения), то для её идентификации более значимым становится контекст. Например, есть «неизвестная» формула в контексте, где встречаются термины из списка некоторых ключевых слов для термина «колебание струны» из примеров 1 и 4. Значит, эта формула может быть помещена как «синоним» или примечание в соответствующую статью указателя и будет полезным комментарием при поиске.

*Пример 4:*

- малые плоские поперечные колебания струны;
- вертикальное перемещение струны;

- закон свободного колебания струны, Даламбера решение, Даламбера формула;
- общее решение однородного волнового уравнения,
- однородное волновое уравнение.

Предлагаемый в статье вариант работы с математическими текстами рассчитан на публикации в TeX-нотации. При этом, поскольку используется идея сравнения «по аналогии» и контекстный поиск, то в определённой степени предложенный механизм может быть использован и для работы с текстовыми документами и MSword-документами. Работа с pdf-файлами - это отдельная задача. Эти файлы, если они размещены в сети для специалистов, как правило, имеют стандартные наборы сопроводительных описаний в виде библиографических записей и других полей для поиска, включая тематическую принадлежность и указания на классификаторы.

При анализе текста необходимо выявить его структуру, определить список ключевых слов и связей. Найти «главный» объект, воздействие на него и действие главного объекта, установить «отклонение» от «основного» смысла. Приведём простой пример на основе использования ключевых слов текста, не касаясь его структуры: есть «k» текстов, в них содержится какое-то количество слов, совпадающих с терминами тезауруса ПрО. Главными объектами текста считаются ключевые слова, совпадающие с дескрипторами тезауруса ПрО. Вводится интегральный показатель ( $N_i$ ,  $i=1,2,...,k$ ) этих совпадений для оценки тематической близости текстов к тематике ПрО. Если  $N_1 < N_2 < ... < N_k$ , то чем больше частотный показатель совпадений, тем ближе исследуемый текст к тематике ПрО.

Любой процесс в математической физике можно описать по-разному. Если есть совпадения, то можно говорить об аналогии на основе тезауруса ПрО. Например, в [23] приводится множество аналогий математического и физического описания волновых процессов. Эти примеры характерны для случаев различных традиционных представлений (математической записи) волновых уравнений математической физики, используемых в математике и в физике. При сравнении текстов с математическими записями возможны варианты: а) математические представления различные, но описывается одно и то же явление (процесс), б) математические записи совпадают, а словесное описание отличается. И в случае а), и в случае б) сравнение текстов выявит аналогию в смысловом содержании.

Несмотря на тот факт, что ключевые слова могут относиться к терминам различных тезаурусов, совпадающие математические записи позволяют сравнивать тексты и можно сделать вывод, что термины двух тезаурусов необходимо дополнить ассоциативными связями, так как они ссылаются на одинаковые математические записи.

Таким образом, можно использовать сравнение текстов для автоматического расширения тезауруса некоторой ПрО за счёт выявленных новых связей между терминами, если это не противоречит логике самих текстов. В частности, для больших текстов можно реализовать механизм создания тезаурусов на основе частотного анализа словаря ключевых слов и выявления главных терминов в тексте (в качестве дескрипторов) и связанных с ним терминов (в качестве ассоциативных). При реализации этого механизма сжатия смысл текста будет в какой-то мере определяться списком дескрипторов и связями, что позволит ускорить процесс сравнения при поиске. В цифровой библиотеке LibMeta множество признаков сравнения задаётся значениями полей атрибутов (см. пример 5 на рисунке 1).

Наличие механизма сравнения формул в нотации MathML или TeX-нотации весьма расширяет возможности поиска математических текстов, хотя и может привести к известным дополнительным сложностям из-за неоднозначности записи в этих нотациях. Известно достаточное количество разработок, посвящённых сравнению математических текстов в формате TeX, однако на практике применяется контекстное сравнение текстов. Например, в реа-

лизации проекта EgoMath2 [24], используется процедура конвертации математических формул в некоторые текстовые аналоги для последующего сравнения.

Естественный механизм «разметки» документов задаётся при наличии ключевых слов в пристатейных вторичных документах. В противном случае, используя частотный анализ, составляется локальный тезаурус или словарь для статьи. Т.е. выделяются главные термины и работы, связанные с ними реферативным материалом. Таким образом статья «приписывается» к определённой ПрО, где выбранные главные термины входят в состав дескрипторов.

Основная работа, безусловно, - это составление самого предметного тезауруса и сопровождение его формулами в TeX-нотации. Формулы при поиске сравниваются с точностью до переменных, а наличие формул в виде «картинок» даёт дополнительную визуальную информацию - это полезная информация, которая служит уточнению запроса.

В цифровой версии словаря всевозможные данные (знания) о специальных функциях помещаются в поля определений и используются в дальнейшем поиске (реализовано в цифровой библиотеке LibMeta [14]). Это названия на русском и английском, запись в TeX-нотации, изображения в jpg-формате, индексы классификаторов УДК (универсальный десятичный классификатор) и соответствующие в MSC (Mathematics Subject Classification). Возможна другая дополнительная информация, ссылки на литературу и на другие термины (ассоциативные связи). Фрагмент из двуязычного словаря специальных функций математической физики, составленный с использованием справочника [25], приведён в примере 6.

*Пример 6:*

- Термин, Ru: Бернулли многочлен обобщённый (многочлен Нёрлунда)
- Термин, En: Generalized Bernoulli polynomial (Noerlund polynomial)
- Термин, TeX:  $B_n^{(\alpha)}(z)$
- Термин, jpg:  $b_{m,n}(z)$
- Определение, TeX:  $\left( \frac{1-e^{-t}}{t} \right)^{\alpha} e^{zt} = \sum_{k=0}^{\infty} B_k^{(\alpha)}(z) \frac{t^k}{k!}$
- Определение, jpg:  $b_{m,n}(z) = \binom{n}{m} z^m (1-z)^{n-m}$
- УДК: 517.589
- MSC: 11B68

Статья словаря в таком виде содержит все поля для идентификации соответствующих формул в тексте. Источником информационного шума становится «поиск по аналогии», который часто применяется в информационно-поисковых системах. Уточнение с помощью формулы позволяет квалифицированному пользователю сократить время поиска и принятия решения при обращении к первоисточнику.

LibMeta - Список наборов атрибутов

## LibMeta

СЕМАНТИЧЕСКАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА

[На главную](#)
[Новый набор атрибутов](#)

### Список наборов атрибутов

Название (на русск.)	Обозначение (на англ.)	Описание
<a href="#">Набор атрибутов персоны</a>	FOAF	Набор атрибутов для персоны (FOAF)
<a href="#">Набор атрибутов публикации</a>	DC	Набор атрибутов для публикации (DC)
<a href="#">Набор атрибутов для лит-ры ОДУ</a>	ODU attribute set	Набор атрибутов для описания экземпляров литературы по ОДУ
<a href="#">Набор атрибутов авторов ОДУ</a>	auth_ODU	Набор атрибутов авторов из литературы, использовавшейся при составлении ОДУ
<a href="#">Набор для описания формул</a>	Set for formulas	Набор для описания формул

© LibMeta 2019

Рисунок 1 – Пример 5: описание формулы через набор атрибутов в цифровой библиотеке LibMeta

## 1.2 Использование семантических структур для описания связей в ПрО уравнений математической физики

Одно из первых появление формул в тезаурусе относится к ПрО химии, что вполне оправдано и стало возможным после появления средств визуализации в Интернете. В русскоязычном сегменте в 1972 г. появился «Тезаурус научно-технических наук» [26], в 1978 г. «Тезаурус по теории формальных и алгебраических групп» [27], в 1991 г. «Русско-немецкий информационно-поисковый тезаурус по робототехнике» [28]. В 2005 г. был опубликован «Тезаурус информационно-поисковый по ПрО: обыкновенные дифференциальные уравнения» [12], где были даны основные дескрипторы и связи терминов в этой области. В работе [13] был дан обзор этих и других ИР, но теперь через 10 лет многое изменилось. В основном это англоязычные указатели, тезаурусы и классификаторы. Ниже приведены наиболее распространённые открытые источники с указанием в таблицах иерархических структур и родовых отношений в тех случаях, когда это явно представлено в самих ресурсах.

### 1.2.1 Проект «Математика в Web-представлении»

Проект «Математика в Web-представлении» (Mathematics on the Web) американского математического общества (AMS), <http://www.ams.org/mathweb/mi-classifications.html>. Он включает предметные классификации MSC и Zentralblatt/Math Reviews Classification Statistics, представленные в виде списка гиперссылок, реализующих иерархические (род-вид) и горизонтальные (синонимы, ассоциативные) связи. Для поиска нужного раздела можно перемещаться по цепочке гиперссылок или использовать одно или несколько ключевых слов ПрО. Добавлены новые, «нестандартные» классификаторы издателей, но многие из них перестали существовать или вошли в состав более крупных ресурсов и также стали недоступны.

Ранее раздел «уравнения с частными производными» входил в подраздел «дифференциальные уравнения, интегральные уравнения и операторы», в который включаются также разделы «обыкновенные дифференциальные уравнения», «динамические системы», «интегральные уравнения», «вариационное исчисление и оптимизация» и «глобальный анализ». В качестве дополнительных гиперссылок данного подраздела были приведены «численные методы», «численный анализ» и «приложения» (см. таблицу 1). Теперь этот ресурс расширился и охватывает практически всю ПрО «уравнения с частными производными» на английском языке. Появилась расшифровка позиции кода MSC:35 из предыдущей версии (см. таблицу 2).

### 1.2.2 Указатель библиотеки Конгресса США

Указатель, использующий систему библиотеки Конгресса США, (Library of Congress Classification Outline), <https://www.loc.gov/catdir/cpsolcco/>. В нём нет детализации в разделе математики и тема «уравнения математической физики» не выделена в самостоятельный раздел (см. таблицу 3).

Таблица 1 – Дифференциальные уравнения, интегральные уравнения и операторы (Differential Equations and integral equations and Operators) в системе MSC 2008 г.

Коды родовых терминов	Родовые термины
MSC:34	Ordinary differential equations
MSC:35	Partial differential equations
MSC:45	Integral equations
MSC:48	Calculus of variations and optimization
MSC:49	Global analysis
	▪ numerical methods, numerical analysis, applications

Таблица 2 – Уравнения с частными производными (Partial differential equations) в системе MSC 2018 г.

Коды	Коды	Видовые термин
	35-00	General reference works (handbooks, dictionaries, bibliographies, etc.)
	35-01	Instructional exposition (textbooks, tutorial papers, etc.)
	35-02	Research exposition (monographs, survey articles)
	35-03	Historical (must also be assigned at least one classification number from Section 01)
	35-04	Explicit machine computation and programs (not the theory of computation or programming)
	35-06	Proceedings, conferences, collections, etc.
35Axx		General theory
35Bxx		Qualitative properties of solutions
35Cxx		Representations of solutions
35Dxx		Generalized solutions of partial differential equations
35Exx		Equations and systems with constant coefficients [See also 35N05]
35Fxx		General first-order equations and systems
35Gxx		General higher-order equations and systems
35Hxx		Close-to-elliptic equations
35Jxx		Partial differential equations of elliptic type [See also 58J10, 58J20]
35Kxx		Parabolic equations and systems [See also 35Bxx, 35Dxx, 35R30, 35R35, 58J35]
35Lxx		Partial differential equations of hyperbolic type [See also 58J45]
35Mxx		Partial differential equations of special type (mixed, composite, etc.) {For degenerate types, see 35J70, 35K65, 35L80}
35Nxx		Overdetermined systems [See also 58Hxx, 58J10, 58J15]
35Pxx		Spectral theory and eigenvalue problems for partial differential operators [See also 47Axx, 47Bxx, 47F05]
35Qxx		Equations of mathematical physics and other areas of application [See also 35J05, 35J10, 35K05, 35L05]
35Rxx		Miscellaneous topics involving partial differential equations {For equations on manifolds, see 58Jxx; for manifolds of solutions, see 58Bxx; for stochastic PDEs, see also 60H15}
35Sxx		Pseudodifferential operators and other generalizations of partial differential operators [See also 47G30, 58J40]

Таблица 3 – Раздел «математика» (Mathematics) указателя библиотеки Конгресса США

Родовой подкласс	Коды подклассов (4 уровня) подкласса QA и (видовые) термины
Subclass QA	<ul style="list-style-type: none"> <li>▪ QA1-939 Mathematics <ul style="list-style-type: none"> <li>○ QA1- 43 General</li> <li>○ QA1- 47- 59 Tables</li> <li>○ QA71-90 Instruments and machines <ul style="list-style-type: none"> <li>▪ QA75-76.95 Calculating machines</li> <li>▪ QA75.5-76.95 Electronic computers. Computer science <ul style="list-style-type: none"> <li>➤ QA76.75-76.765 Computer software</li> </ul> </li> </ul> </li> </ul> </li> <li>▪ QA101-(145) Elementary mathematics. Arithmetic</li> <li>▪ QA150-272.5 Algebra</li> <li>▪ QA273-280 Probabilities. Mathematical statistics</li> <li>▪ QA299.6-433 Analysis</li> <li>▪ QA440-699 Geometry. Trigonometry. Topology</li> <li>▪ QA801-939</li> </ul>

### 1.2.3 Указатель на основе десятичной системы Дьюи

Классификация на основе десятичной системы Дьюи (Dewey Decimal Classification, DDC) существует с 1873 года. Сейчас редакция 2011 г. используется в основном в библиотечном деле. Раздел «математическая физика» представлен в науке «физика», подчинённом в разделе «естественные науки и математика»: DDC - 500 - Natural sciences & mathematics, [http://bpeck.com/references/DDC/ddc\\_mine500.htm](http://bpeck.com/references/DDC/ddc_mine500.htm) (см. таблицу 4).



Таблица 4 – Раздел «физика» (Physics) указателя Дьюи

Код и родовой термин	Коды и вложенные (видовые) термины
530 Physics	<ul style="list-style-type: none"> <li>▪ 530 Physics <ul style="list-style-type: none"> <li>○ 530.1 Theories and mathematical physics</li> <li>○ 530.4 States of matter</li> <li>○ 530.8 Testing and measurement</li> </ul> </li> <li>▪ 531 Classical mechanics; Solid mechanics</li> <li>▪ 532 Fluid mechanics; Liquid mechanics</li> <li>▪ 533 Gas mechanics</li> <li>▪ 534 Sound &amp; related vibrations <ul style="list-style-type: none"> <li>○ 534.5 Subsonic and ultrasonic vibrations</li> </ul> </li> <li>▪ 535 Light &amp; related radiations <ul style="list-style-type: none"> <li>○ 535.5 Beams</li> <li>○ 535.6 Color</li> <li>○ 535.8 Spectroscopy</li> </ul> </li> <li>▪ 536 Heat</li> <li>▪ 537 Electricity &amp; electronics <ul style="list-style-type: none"> <li>○ 537.5 Electronics</li> <li>○ 537.6 Electric currents (Electrodynamics) and thermoelectricity</li> </ul> </li> <li>▪ 538 Magnetism</li> <li>▪ 539 Modern physics <ul style="list-style-type: none"> <li>○ 539.2 Radiations (Radiant energy)</li> <li>○ 539.7 Atomic and nuclear physics</li> </ul> </li> </ul>

Перечисленные наиболее известные англоязычные ресурсы используются в различных информационно-поисковых и библиографических системах. Многие из цитированных ранее в [13] ресурсов за истекший период стали закрытыми, например: классификация препринтов на Лос Аламос (Index using Los Alamos Preprint-Server Classifications), проект американского математического общества «Математический атлас» (The Mathematical Atlas), проект Массачусетского технологического университета «Открытые знания» (Unlocking Knowledge). Это связано во многом с тем, что системы классификации стали частью информационных, в основном библиографических, баз данных с соответствующими авторскими правами. Сюда можно отнести классификаторы, которые используются в известных программных продуктах таких, как MathWorks пакета MATLAB ([http://www.mathworks.com/products/product\\_listing/index.ht](http://www.mathworks.com/products/product_listing/index.ht)), Wolfram Research (<http://mathworld.wolfram.com/PartialDifferentialEquation.html>), тематические классификаторы WoS (<http://apps.webofknowledge.com>) и Scopus (<https://www.scopus.com>), которые закрыты для внешнего пользователя.

Наиболее полно ПрО «уравнения с частными производными» отражена в системе MSC. Формулы и язык для записи формул, разработанные в системах Mathematica Wolfram Research и MathWorks, являются внутренними средствами этих систем. Формулы используются активно в «Encyclopedia of Mathematics» ([https://www.encyclopediaofmath.org/index.php/Main\\_Page](https://www.encyclopediaofmath.org/index.php/Main_Page)).

Из русскоязычных ресурсов в электронном виде продолжают развиваться известные классификаторы ГРНТИ (Государственный рубрикатор научно-технической информации) и УДК (Универсальная десятичная классификация) Всероссийского института научной и технической информации (ВИНИТИ). Наряду с ними можно отметить ресурсы <http://eqworld.ipmnet.ru/>, <http://ontomathpro.org/ontology/> как самостоятельные версии указателей с иерархическими и горизонтальными связями, а также «Общероссийский математический портал» Math-Net.Ru (<http://www.mathnet.ru/>).

### 1.2.4 Классификатор ГРНТИ

ГРНТИ представляет собой универсальную иерархическую классификацию областей знания на русском языке, принятую для систематизации всего потока научно-технической информации (<http://www.viniti.ru/products/classification-systems/rubricator-grnti>). На основе рубрикатора построена система локальных (отраслевых, проблемных) тематических указателей научно-технической информации и рубрикатор отраслей знаний (<http://www.viniti.ru/products/classification-systems/rubricator-viniti>). Последняя редакция ГРНТИ 2018 г. содержит, в частности, раздел «дифференциальные уравнения в частных производных» (см. таблицу 5).

Таблица 5 – Раздел «дифференциальные уравнения в частных производных» рубрикатора ГРНТИ

Код и родовой термин	Коды и видовые термины
27.31 дифференциальные уравнения в частных производных	<ul style="list-style-type: none"> <li>▪ 27.31.17 линейные и квазилинейные уравнения</li> <li>▪ 27.31.19 нелинейные уравнения</li> <li>▪ 27.31.21 общие уравнения первого порядка</li> <li>▪ 27.31.33 уравнения высших порядков</li> </ul>

### 1.2.5 Классификатор УДК

Основное назначение УДК - тематическая классификация. На сайте ВИНТИ (<http://scs.viniti.ru/udc/>) по номеру УДК можно получить значение, например: на запрос «517.956.6», получить ответ «уравнения смешанного типа».

Таблицы УДК найти можно на разных ресурсах в Интернете. Раздел УДК 517.95 «дифференциальные уравнения с частными производными» показан в таблице 6. В позициях 517.956 и др. есть более глубокая иерархическая цепочка, но полного соответствия, например классификации MSC, нет, как нет и иллюстраций в виде формул.

### 1.2.6 Интернет-проекты

Интернет-проект (<http://eqworld.ipmnet.ru/indexr.htm>) «Мир математических уравнений» (EqWorld) поддерживает шесть языков (русский, английский, немецкий, французский, итальянский, испанский). Научный редакторский коллектив под руководством А.Д. Полянина состоит из представителей различных математических коллективов России, США, Канады и др. EqWorld содержит многочисленные ссылки на вспомогательный материал, библиографию, вычислительные средства (Maple, Mathematica, MATLAB и др.). Указатель (предметный и именной) математических уравнений включает разделы: ОДУ, дифференциальные уравнения с частными производными (уравнения математической физики), интегральные уравнения, функциональные уравнения. EqWorld наиболее приближен к российской классической математической школе как учебник и справочник с формулами. Это реализуется за счёт ссылок на справочник [29], т.е. по ссылке из указателя переходим к страничке справочника (см. рисунок 2).

Ресурс «Общероссийский математический портал» (<http://www.mathnet.ru/>) Математического института им. В.А. Стеклова РАН использует классификаторы MSC, УДК и др. На странице Math-Net.Ru содержатся также ссылки на актуальные программные продукты, такие как MikTeX (пакет для работы с TeX в операционной системе Windows), Magma (компьютерно-алгебраическая система для вычислений в алгебре, теории чисел и геометрии) и др. Имеются также ссылки на известные классификаторы (УДК и MSC 2000) и ряд полезных ссылок на известные издательства, математические ресурсы в Интернете, реферативные базы и математические научные сообщества.

Таблица 6 – Раздел «дифференциальные уравнения с частными производными» УДК ВИНТИ

Код и родовой термин	Коды и видовые термины	Коды и ассоциативные термины
517.95 Дифференциальные уравнения с частными производными	517.951 Общая теория дифференциальных уравнений с частными производными: вариационные, аналитические, основанные на интегральных преобразованиях и другие методы. Геометрическая теория. Аппроксимация	<ul style="list-style-type: none"> <li>см. 517.518.22 Компактные семейства функций</li> <li>см. 517.972.5 Вариационные методы решения дифференциальных, интегральных и других функциональных уравнений</li> <li>см. 517.984.5 Изучение спектра конкретных операторов, в том числе дифференциальных и интегральных операторов</li> </ul>
	517.952 Общие уравнения первого порядка и системы: свойства, типы и т.п.	
	517.953 Общие уравнения высших порядков и систем: свойства, типы и т.п.	
	517.954 Краевые задачи: общая теория, уравнения на многообразиях	<ul style="list-style-type: none"> <li>см. 517.927 Краевые задачи для обыкновенных дифференциальных уравнений</li> <li>см. 517.929.7 Краевые задачи</li> </ul>
	517.955 Задачи Коши для дифференциальных уравнений с частными производными	
	517.956 Линейные и квазилинейные уравнения и системы	
	517.957 Нелинейные уравнения и системы	
	517.958 Дифференциальные и интегральные уравнения математической физики	

Ресурс «*OntoMathpro Ontology classes*» (<http://ontomathpro.org/ontology>) Казанского федерального университета содержит англо-русский алфавитный указатель терминов с горизонтальными связями, например термин «Численное решение сингулярных интегральных уравнений с ядром Коши» (см. рисунок 3).

Exact Solutions > Linear Partial Differential Equations >  
Second-Order Hyperbolic Partial Differential Equations > Telegraph Equation

## 2.7. Telegraph Equation $\frac{\partial^2 w}{\partial t^2} + k \frac{\partial w}{\partial t} = a^2 \frac{\partial^2 w}{\partial x^2} + bw$

For  $k > 0$  and  $b < 0$ , it is the telegraph equation.

The substitution  $w(x, t) = \exp(-\frac{1}{2}kt)u(x, t)$  leads to the Klein-Gordon equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + (b + \frac{1}{4}k^2)u,$$

which is discussed in Subsection 2.3.

### Reference

Polyanin, A. D., *Handbook of Linear Partial Differential Equations for Engineers and Scientists*, Chapman & Hall/CRC, 2002.

Telegraph Equation

Copyright © 2004 Andrei D. Polyanin

<http://eqworld.ipmnet.ru/en/solutions/lpde/lpde207.pdf>

Рисунок 2 – Представление термина «телеграфное уравнение» в указателе EqWorld

## Class: 'Численное решение сингулярных интегральных уравнений с ядром Коши'

<http://cll.niimm.ksu.ru/ontologies/mathematics#E433>

### Annotations (2)

- comment "См. Габдулхаев Б.Г. Прямые методы решения сингулярных интегральных уравнений первого рода. Численный анализ. - Казань: КГУ, 1994. Иванов В.В. Теория приближенных методов и ее применение к численному решению сингулярных интегральных уравнений. - Киев: Наукова Думка, 1968. Лифанов И.К. Метод сингулярных интегральных уравнений и численный эксперимент (в математической физике, аэродинамике, теории упругости и дифракции волн). - М.: ТОО Янус, 1995." (ru)
- label "Численное решение сингулярных интегральных уравнений с ядром Коши" (ru)

### Superclasses (1)

- 'Computational solution of integral equation'

### Usage (1)

- Class: 'Численное решение сингулярных интегральных уравнений с ядром Коши'

Рисунок 3 – «Численное решение сингулярных интегральных уравнений с ядром Коши» в словаре OntoMathpro

Среди авторов и создателей этих ресурсов нет единого мнения и/или классификатора, для ПрО «уравнения математической физики», но существует достаточно много версий и уточнений. Каждое уточнение может дополнять существующие системы отношений и знаний ПрО в целом. Для сохранения терминологической чистоты научных текстов добавление новых терминов должно сопровождаться расширением тезауруса силами экспертного научного сообщества.

## 2 Уравнения математической физики смешанного типа

### 2.1 О предметной области уравнений математической физики

Изучение ПрО уравнений математической физики затрагивает многие разделы математики. За определениями, которые помогут установить семантические связи понятий этой ПрО, следует обратиться к математической энциклопедии [30], где они представлены:

*В.С. Владимиров:* «Математической физики уравнения – уравнения, описывающие математические модели физических явлений. ...»

*А.Н. Тихонов:* «Математическая модель – приближённое описание какого-либо класса явлений внешнего мира, выраженное с помощью математической символики».

«Математическая физика (МФ) - теория математических моделей физических явлений; занимает особое положение и в математике, и в физике, находясь на стыке этих наук. М.Ф. тесно связана с физикой в той части, которая касается построения математической модели, и в то же время МФ – раздел математики, поскольку методы исследования моделей являются математическими. В понятие методов МФ включаются те математические методы, которые применяются для построения и изучения математических моделей, описывающих большие классы физических явлений...»

Введём обозначения: «математическая физика» (mathematical physics, MPh), «методы математической физики» (methods of mathematical physics, MMPh), «уравнения математической физики» (equations of mathematical physics, EQMPh), «математические методы» (mathematical methods, MM), «методы теоретической физики» (methods of theoretical physics, MTPh), «математическая символика» (mathematical symbols, MSb).

Все обозначенные уравнения и методы связаны между собой в ПрО MPh (subject domain of mathematical physics, SD MPh), как указано в [30], следующим образом: EQMPh  $\subset$  MTPh; MSb  $\subset$  MM  $\subset$  MMPh.

Уравнения смешанного типа отличаются тем, что принимают гиперболический тип (hyperbolic type, HT), параболический тип (parabolic type, PT) и эллиптический тип (elliptic type, ET), в зависимости от области определения (рассмотрения), как часть PDE второго порядка. В таблице 7 приведены иерархические отношения для PDE второго порядка.

Таблица 7 – Раздел «PDE второго порядка»

Родовой термин	Видовые термины
✓ PDE второго порядка с двумя независимыми переменными	<ul style="list-style-type: none"> <li>▪ именные PDE</li> <li>▪ нарицательные PDE               <ul style="list-style-type: none"> <li>• Линейные PDE</li> <li>• Линейные PDE относительно старших производных с коэффициентами у старших производных                   <ul style="list-style-type: none"> <li>➤ постоянные коэффициенты</li> <li>➤ переменные коэффициенты                       <ul style="list-style-type: none"> <li>❖ PDE гиперболического типа</li> <li>❖ PDE параболического типа</li> <li>❖ PDE эллиптического типа</li> </ul> </li> </ul> </li> </ul> </li> </ul>

Для эффективности работы каждый термин сопровождается идентификатором, который в дальнейшем используется в поисковом запросе. Идентификаторы терминов PDEXXXX(YY), где PDE - «partial differential equation», XXXX - десятичный номер, YY - принимает значение типа уравнения HT, PT, ET.

*Видовые термины* в таблице 8 - уравнения типа HT, PT и ET.

Таблица 8 – Терминологические обозначения PDE второго порядка с двумя независимыми переменными

Идентификатор	Термин (словесное обозначение, символьные выражения, идентификатор)
[PDE1003]	«Линейное однородное PDE второго порядка с постоянными коэффициентами и двумя независимыми переменными» $L_1 = 0$ . $L_1 \equiv Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = 0$ , где $A, B, C, D, E$ - постоянные.
Вид:	[PDE1004(HT)], [PDE1005(PT)], [PDE1006(ET)]
[PDE1007]	«Линейное неоднородное PDE второго порядка с постоянными коэффициентами и двумя независимыми переменными» $L_2 = \varphi(x, y)$ $L_2 \equiv Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = \varphi(x, y)$ , где $A, B, C, D, E$ - постоянные, $\varphi$ - достаточно гладкая функция переменных $x$ и $y$ .
Вид:	[PDE1008(HT)], [PDE1008(PT)], [PDE1010(ET)]
[PDE1011]	«Линейное относительно старших производных PDE второго порядка с постоянными коэффициентами у старших производных и двумя независимыми переменными» $L_3 = 0$ $L_3 \equiv Au_{xx} + 2Bu_{xy} + Cu_{yy} + F_1(x, y, u, u_x, u_y) = 0$ , где $A, B, C$ - постоянные
Вид:	[PDE1012(HT)], [PDE1013(PT)], [PDE1014(ET)]
[PDE1015]	«Линейное относительно старших производных PDE второго порядка с постоянными коэффициентами у старших производных, правой частью и двумя независимыми переменными» $L_4 = \varphi(x, y)$ , $L_4 \equiv Au_{xx} + 2Bu_{xy} + Cu_{yy} + F_1(x, y, u, u_x, u_y) = \varphi(x, y)$ , где $A, B, C$ - постоянные, $\varphi$ - достаточно гладкая функция переменных $x$ и $y$ .
Вид:	[PDE1016(HT)], [PDE1017(PT)], [PDE1018(ET)]
[PDE1019]	«Линейное однородное PDE второго порядка с переменными коэффициентами двумя независимыми переменными» $L_5 = 0$ , $L_5 \equiv a_{11}(x, y)u_{xx} + 2a_{12}(x, y)u_{xy} + a_{22}(x, y)u_{yy} + b(x, y)u_x + c(x, y)u_y + d(x, y)u = 0$ , где $a_{11}, a_{12}, a_{22}, b, c, d$ - достаточно гладкие функции переменных $x$ и $y$
Вид:	[PDE1020(HT)], [PDE1021(PT)], [PDE1022(ET)]
[PDE1023]	«Линейное неоднородное PDE второго порядка с переменными коэффициентами двумя независимыми переменными» $L_6 = f(x, y)$ $L_6 \equiv a_{11}(x, y)u_{xx} + 2a_{12}(x, y)u_{xy} + a_{22}(x, y)u_{yy} + b(x, y)u_x + c(x, y)u_y + d(x, y)u = f(x, y)$ $a_{11}, a_{12}, a_{22}, b, c, d, f$ - достаточно гладкие функции переменных $x$ .
Вид:	[PDE1024(HT)], [PDE1025(PT)], [PDE1026(ET)]
[PDE1027]	«Линейное относительно старших производных PDE второго порядка с переменными коэффициентами у старших производных и двумя независимыми переменными» $L_7 = 0$ $L_7 \equiv A(x, y)u_{xx} + 2B(x, y)u_{xy} + C(x, y)u_{yy} + F_2(x, y, u, u_x, u_y) = 0$ , где $A, B, C$ - достаточно гладкие функции переменных $x$ и $y$ .
Вид:	[PDE1028(HT)], [PDE1029(PT)], [PDE1030(ET)]
[PDE1031]	«Линейное относительно старших производных PDE второго порядка с переменными коэффициентами у старших производных, правой частью и двумя независимыми переменными» $L_8 = f_1(x, y)$ $L_8 \equiv A(x, y)u_{xx} + 2B(x, y)u_{xy} + C(x, y)u_{yy} + F_2(x, y, u, u_x, u_y) = f_1(x, y)$ , где $A, B, C, f_1$ - достаточно гладкие функции переменных $x$ и $y$ .
Вид:	[PDE1032(HT)], [PDE1033(PT)], [PDE1034(ET)]



Наименования рассматриваемых PDE второго порядка по возможности должны обозначать структуру уравнений и/или указывать на их использование. Наименования PDE часто сопровождаются математическими записями «уравнение колебаний», например:

$$\rho \frac{\partial^2 u}{\partial t^2} = \operatorname{div}(p \operatorname{grad} u) - qu + F(x, t); \text{ «уравнение колебаний струны», } \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 \text{ и др.}$$

Для данных *именных* PDE в любой информационной среде математическая запись обязательна.

## 2.2 Об указателе терминов и уравнений смешанного типа

В задаче поиска математических публикаций для ПрО уравнений смешанного типа, как части общего ресурса по уравнениям с частными производными, формируется указатель терминов и уравнений по тематике уравнений смешанного типа. Используется термин «указатель» (index в английской терминологии).

Предлагается следующая структура и состав статьи «указатель»:

- терминологическое обозначение – дескриптор PDE, идентификатор PDE;
- термины - синонимы (Син);
- коды классификаторов (MSC и УДК);
- символьные обозначения (Форм, MSWord, Latex);
- литература (Лит);
- родовые дескрипторы (Род);
- видовые дескрипторы (Вид);
- ассоциативные дескрипторы (Асс);
- ключевые слова (КлС);
- см. также (См);
- историческая справка (Ист);
- примечание: пояснения, ссылки, связи, дополнительная информация (Прим).

В качестве примера статьи указателя для PDE смешанного типа рассматривается *именной* термин-дескриптор «Трикоми уравнение» этой ПрО. Для этого примера используется более развернутый вариант идентификаторов, дополненный указаниями о типе смешанного уравнения и канонической форме, то есть следующие сокращения: PDE MXT - PDE смешанного типа, CF - каноническая форма(ы), первая - CF1, вторая - CF2.

В таблице 9 приведён вариант представления статьи указателя для дескриптора «Трикоми уравнение», где указаны ссылки на синонимы, ассоциативные, видовые и родовые термины в виде их идентификаторов. Формулы для видовых терминов с указанием первоисточников имеют вид:

$$|PDE1031(CF2,HT,y<0)| [31]: z_{\xi\eta} + \frac{1}{6(\eta - \xi)}(z_{\xi} - z_{\eta}) = 0, (\xi = x - \frac{3}{2}(-y)^{\frac{3}{2}}, \eta = x + \frac{3}{2}(-y)^{\frac{3}{2}}, y < 0)$$

$$|PDE1031(CF,ET,y>0)| [32]: z_{\xi\xi} + z_{\eta\eta} + \frac{1}{3\eta}z_{\eta} = 0, (\xi = x, \eta = \frac{3}{2}(-y)^{\frac{3}{2}}, y > 0)$$

$$|PDE1031(CF,ET, z = \eta^{\alpha}u, \alpha = -\frac{1}{6})| [32]: \frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} + \frac{5}{36} \frac{1}{\eta^2} u = 0, (z = \eta^{\alpha}u, \alpha = -\frac{1}{6})$$

$$|PDE1031(CF,ET, z = \eta^{\alpha}u, \alpha = -\frac{1}{3})| [32]: \frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} + \frac{1}{\eta} \frac{\partial u}{\partial \eta} - \frac{1}{9\eta^2} u = 0, (z = \eta^{\alpha}u, \alpha = -\frac{1}{3})$$

Особенности разрабатываемого указателя: структура статьи указателя содержит разнообразную информацию о термине; позволяет реализовать три варианта входа: термин, формула, автор; использовать родо-видовые отношения для формирования семантической

структуры ПрО; выбирать коды классификаторов MSC и/или УДК; добавлять термин со ссылкой на первоисточник. Указанные свойства позволили представить термин с контекстом - ключевыми словами, авторами и др. Это расширяет поле поиска при информационном запросе. С другой стороны, наличие формул, как вспомогательной иллюстративной информации, так и в качестве «ключевых слов (символов)» в TeX-нотации, может служить уточнению запроса, уменьшению поискового шума, что, в итоге, приведёт к получению пертинентной информации для адресата. В качестве синонимов приведены русские и английские эквиваленты терминов и сочетание кодов классификаторов MSC и УДК, что необходимо при работе со статьями и при поиске в базах данных.

Таблица 9 – Статья указателя для дескриптора «Трикоми уравнение»

Позиции статьи указателя	Тип данных	Значения позиций статьи указателя
PDE1031(MXT)	Ru	Трикоми уравнение
	En	Tricomi equation
Син1	Ru	Уравнение Т
Син2	En	Equation T
	MSC	35M10 Equations of mixed type
	УДК	517.956.6 Уравнения смешанного типа
Форм 1	MSWord	$y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$
Форм2	LaTex	$y \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$
Форм3	LaTex	$y u_{xx} + u_{yy} = 0$
Лит1	[1ru]	Смирнов М. М. Уравнения смешанного типа. М., 1970. 296 с.
Лит2	[2ru]	Трикоми Ф. Лекции по уравнениям в частных производных. Пер. с итал. Д.А. Райкова с предисл. Б. М. Левитана М.: Изд-во иностр. лит., 1957. 443 с.
Лит3	[1en]	Smirnov M.M. Equations of mixed type. American Translation of the Mathematical monographs. Vol. 51. Mathematical Soc., 31 Dec 1978. p: 232.
Лит4	[2en]	Lupo D., Morawetz C.S., Payne K.R. On closed boundary value problems for equations of mixed elliptic-hyperbolic type // Communications on Pure and Applied Mathematics, 2007, 60 (9), 1319-1348
Род	Ru	PDE1030(MXT)  Уравнения смешанного типа
Род	En	PDE1030(MXT)  Equations of mixed type
Вид 1		PDE1031(CF2,HT,y<0)
Вид 2		PDE1031(CF,ET,y>0)
Вид 3		PDE1031(CF,ET, $z = \eta^\alpha u, \alpha = -\frac{1}{6}$ )
Вид 4		PDE1031(CF,ET, $z = \eta^\alpha u, \alpha = -\frac{1}{3}$ )
КлС		Трикоми задача, обобщение Трикоми задачи, Трикоми метод, задача Коши для уравнения Трикоми, краевая задача для уравнения типа Чаплыгина, краевая задача для уравнения смешанного типа
Асс	Ru	PDE1037(MXT)  Геллерстедта задача
Асс	En	PDE1037(MXT)  The Gellerstedt problem
См		PDE1039(MXT)  Газовой динамики уравнение смешанного типа
Ист		Название термина по имени Трикоми Франческо Джакомо (Tricomi Francesco Giacomo), появилось в [1ru] как ссылка на работу[2ru]
Прим		PDE1036(MXT)  Геллерстедта решение задачи Трикоми для уравнения

Эти особенности позволяют перейти на язык онтологий для описания ПрО уравнения смешанного типа (SD PDE MXT), как части ПрО уравнения с частными производными (SD PDE) и, соответственно, ПрО уравнений математической физики (SD EQMPh). Отсюда:

$$SD PDE MXT \subset SD PDE \subset SD EMPh.$$

Переход к онтологии реализуется на базе LibMeta [14], где для поддержки формул было введено отдельное понятие *Формула*, которое позволяет хранить оригинальную строку формулы из источника, где она получена. Строка может быть в формате Content MathML, Presentation MathML, L<sup>A</sup>T<sub>E</sub>X. На рисунке 4 приведён пример из этой системы для ПрО ОДУ [12].

<p><a href="#">Петровский И. Лекции по теории обыкновенных дифференциальных уравнений. Изд. 6-е, исправл.</a>  <a href="#">Матвеев Н.М. Методы интегрирования обыкновенных дифференциальных уравнений.</a>  <a href="#">Арнольд В.И. Обыкновенные дифференциальные уравнения.</a></p>	
Родовые связи	<p><a href="#">ОДУ первого порядка линейное неоднородное</a> [DE0025]  <a href="#">ОДУ первого порядка линейное однородное</a> [DE0026]</p>
Видовые связи	<p><a href="#">Абеля ОДУ второго рода</a> [DE0008]  <a href="#">Абеля ОДУ первого рода</a> [DE0009]  <a href="#">ОДУ первого порядка, разрешенное относительно производной</a> [DE0005]  <a href="#">Риккати ОДУ обобщенное</a> [DE0033]  <a href="#">Риккати ОДУ общее</a> [DE0032]  <a href="#">Якоби ОДУ первого порядка</a> [DE0037]</p>
Ассоциативные связи	<p><a href="#">Бернулли ОДУ</a> [DE0010]  <a href="#">Дарбу ОДУ первого порядка</a> [DE0015]</p>
Смотри также	<p><a href="#">Бернулли подстановка</a> [DM0011]  <a href="#">Вариация произвольной постоянной (Лагранжа метод)</a> [DM0013]  <a href="#">Общее решение (общий интеграл) ОДУ первого порядка линейного неоднородного</a> [RDE0021]  <a href="#">Общее решение (общий интеграл) ОДУ первого порядка линейного однородного</a> [RDE0022]  <a href="#">ОДУ первого порядка с разделенными и разделяющимися переменными</a> [DE0028]  <a href="#">Эйлера метод множителя</a> [DM0035]</p>
Примечание	<p><b>Теорема.</b> (И. Г. Петровский, с. 28—30.)          Пусть функции <math>p(x)</math> и <math>q(x)</math> непрерывны в интервале <math>a &lt; x &lt; b</math>. Тогда через каждую точку <math>(x_0, y_0)</math> полосы, определенной неравенствами <math>a &lt; x &lt; b, -\infty &lt; y &lt; \infty</math>, проходит одна и только одна интегральная кривая этого уравнения, определенная при всех <math>x</math> на интервале <math>(a, b)</math>.          Общее решение (общий интеграл) ОДУ первого порядка линейного неоднородного (см. (1) и (3)) есть сумма общего решения соответствующего однородного ОДУ (см. (2)) и частного решения неоднородного ОДУ (см. (3)) (см. <a href="#">Методы интегрирования обыкновенных дифференциальных уравнений. Учебное пособие. Изд. 5-е, дополн. (с. 60-62)</a>, а также <a href="#">Курс высшей математики. Т. II. Изд. 14-е, исправл. (с. 20-20)</a>; <a href="#">Дифференциальные уравнения. Изд. 3-е, стереотип. (с. 66-66)</a>). ОДУ (1) сохраняет свой вид, т. е. свойство линейности (см. <a href="#">Методы интегрирования обыкновенных дифференциальных уравнений. Учебное пособие. Изд. 5-е, дополн. (с. 55-56)</a>), при преобразованиях:          а) замене независимой переменной вида <math>x = t</math>, где <math>t</math> — непрерывно дифференцируемая функция на некотором интервале <math>(t_0, t_1)</math> и <math>a = (t_0), b = (t_1), t'(t) &gt; 0, t \in (t_0, t_1)</math>;          б) линейной заменой искомой функции <math>y(x) = \alpha(x)z + \beta(x)</math>, где <math>z = z(x)</math> — новая неизвестная функция, <math>\alpha(x), \beta(x)</math> — непрерывно дифференцируемые функции от <math>x, \alpha(x) \neq 0</math>.</p>

Рисунок 4 —Фрагмент описания понятия «ОДУ первого порядка линейное неоднородное» в библиотеке LibMeta

Расширение онтологического представления перечисленных областей математики - это добавление статей в указателе (терминов со связями на основании структуры статьи указателя). Эффект семантических связей основан на использовании терминов из тезауруса математической ПрО, снабженных формулами, как основы для математического запроса.

На примере онтологического представления ПрО, внедрённого в информационно-поисковую цифровую среду, библиотеку LibMeta показаны следующие возможности:

- обогащение тезауруса адресата за счёт информации из указателя/словаря/тезауруса ПрО;
- связи терминов и публикаций упрощают поиск публикации, способствуют учёту предыдущих исследований, т.е. цитированию в рамках ПрО;
- использование терминов из тезауруса в качестве ключевых слов публикации для уточнения значимости исследований в рамках ПрО;
- применение терминов тезауруса в поисковых запросах при составлении реферативного материала и пр.

Термины из тезауруса математической ПрО, снабжённые формулами, составляют основу для математического запроса. Терминологическое описание ПрО на основе словарей и тезаурусов выходит на первый план при решении проблемы уменьшения поискового шума и уточнения запроса с помощью включения в него формул.

## Заключение

Создание словарей, тезаурусов, наборов ключевых слов, списков формул и идентификация связей понятий и терминов ПрО составляют элементы предварительной аналитико-синтетической обработки текстов, помещаемых в Интернет в информационно-поисковую среду. В статье дан краткий обзор тематических ресурсов. Единого стандарта в этой области нет, к тому же далеко не все разделы науки представлены полно. Предлагаемый в работе указатель ПрО «уравнения смешанного типа» служит частичному заполнению указанных пробелов.

## Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ, проекты № 17-07-00217-а, 18-00-00297-комфи.

## Список источников

- [1] **Roget, P.M.** Roget's thesaurus of English words and phrases / P.M. Roget. Taylor Anderson (Editor). — Create Space Independent Publishing Platform. 2017. — 660 p.
- [2] **May, K.O.** Historiography: A Perspective for Computer Scientists / K.O. May // Invited address to International Research Conference on the History of Computing, June 10, 1976, Los Alamos.
- [3] **Paulson, L.C.** Computational logic: Its origins and applications / L.C. Paulson // Proceedings of the Royal Society. 2018. Series. A 474. <https://doi.org/10.1098/rspa.2017.0872>.
- [4] **Клещёв, А.С.** Теоретические основы оболочки для интерактивных систем верификации интуитивных математических доказательств / А.С. Клещёв, В.А. Тимченко // Онтология проектирования. — 2018. — Т. 8, №2(28). — С.219-239. — DOI: 10.18287/2223-9537-2018-8-2-219-239.
- [5] **Клещёв, А.С.** Реализация оболочки и портала знаний по верификации математических доказательств на платформе IASPaas / А.С. Клещёв, В.А. Тимченко // Онтология проектирования. — 2018. — Т. 8, №3(29). — С.427-448. — DOI: 10.18287/2223-9537-2018-8-3-428-448.
- [6] **Лукашевич, Н.В.** Проектирование лингвистических онтологий для информационных систем в широких предметных областях / Н.В. Лукашевич, Б.В. Добров // Онтология проектирования. — 2015. — Т. 5. — №. 1 (15) — С.47-69.
- [7] **Загорулько, Г.Б.** Разработка онтологии для Интернет-ресурса поддержки принятия решений в слабоформализованных областях / Г.Б. Загорулько // Онтология проектирования. — 2016. — Т. 6, №4(22). — С. 485-500. — DOI: 10.18287/2223-9537-2016-6-4-485-500.
- [8] **Boyack, K.W.** Thesaurus-based methods for mapping contents of publication sets / K.W. Boyack // Scientometrics. 2017. 111: 1141. <https://doi.org/10.1007/s11192-017-2304-3>.
- [9] **Ion, P.D.F.** The Global Digital Mathematics Library and the International Mathematical Knowledge Trust. / P.D.F. Ion, S.M. Watt // In: Geuvers H., England M., Hasan O., Rabe F., Teschke O. (eds). Intelligent Computer Mathematics. CICM 2017. Lecture Notes in Computer Science. 2017. Vol. 10383. Springer, Cham - [https://link.springer.com/chapter/10.1007/978-3-319-62075-6\\_5](https://link.springer.com/chapter/10.1007/978-3-319-62075-6_5).
- [10] **Mooers, C.** Information retrieval viewed as temporal signaling // Proceedings of the International Congress of Mathematicians. — 1950. — V.1. — P.572–573.
- [11] **Ландэ, Д.В.** Поиск знаний в Internet. Профессиональная работа / Д.В. Ландэ. — М: Издательский дом «Вильямс». 2005. — 272 с.
- [12] **Моисеев, Е.И.** Тезаурус информационно-поисковый по предметной области: обыкновенные дифференциальные уравнения / Е.И. Моисеев, А.А. Муромский, Н.П. Тучкова. — М.: МАКС Пресс. 2005. — 116 с.
- [13] **Моисеев, Е.И.** Интернет и математические знания: представление уравнений математической физики в информационно-поисковой среде / Е.И. Моисеев, А.А. Муромский, Н.П. Тучкова. — М.: МАКС Пресс. 2008. — 80 с.
- [14] **Серебряков, В.А.** Информационная модель открытой персональной семантической библиотеки LibMeta / В.А. Серебряков, О.М. Атаева // Труды XVIII Всерос. научной конференции «Научный сервис в сети интернет». Новороссийск, 19 по 24 сентября 2016 г. ИПМ им. М.В. Келдыша РАН. — 2016. — С.304-313.
- [15] **Моисеев, Е.И.** Онтология научного пространства или как найти гения / Е.И. Моисеев, А.А. Муромский, Н.П. Тучкова. // Онтология проектирования. — 2014. — №4 (14) . — с.18-33.

- [16] **Miller, B.R.** Technical aspects of the digital library of mathematical functions / B.R. Miller, A. Youssef // *Ann. Math. Artif. Intell.* — 2003. — Vol. 38. — No. 1. — P.121–136.
- [17] **Kohlhase, M.** Mathematical knowledge management: Transcending the one-brain barrier with theory graphs / **M. Kohlhase** // *European Mathematical Society (EMS) Newsletter*, June 2014. — P.22–27.
- [18] **Захаренко, Е.Н.** Новый словарь иностранных слов: свыше 25 000 слов и словосочетаний / Е.Н. Захаренко, Л.Н. Комарова, И.В. Нечаева. 3-е изд., испр. и доп. — М.: ООО ИФ «Азбуковник», 2008. — 1040 с.
- [19] **Булыко, А.Н.** Большой словарь иностранных слов. — Издательство «ИДДК», 2007. — 704 с.
- [20] **Александрова, Н.В.** История математических терминов, понятий, обозначений / Н.В. Александрова // *Словарь - справочник. Издание третье, исправленное.* - М.: Изд-во ЛКИ, — 2008. — 248 с.
- [21] Princeton's WordNet. G.A. Miller, R. Beckwith, C.D. Fellbaum, D. Gross, K. Miller. 1990. WordNet: An online lexical database. *Int. J. Lexicograph.* 3, 4, - P.235–244. - <https://wordnet.princeton.edu/>.
- [22] **Кнут, Д.Э.** Всё про TeX / Д.Э. Кнут. — М.: Издательский дом «Вильямс». 2003. — 560 с.
- [23] **Тихонов, А.Н.** Уравнения математической физики: учеб. пособие / А.Н. Тихонов, А.А. Самарский. 6-е изд., испр. и доп. — М.: Изд-во МГУ, 1999. — 799 с.
- [24] **Misutka, J.** System Description: EgoMath2 As a Tool for Mathematical Searching on Wikipedia.org. / J. Misutka, L. Galambos // In: Davenport J.H., Farmer W.M., Urban J., Rabe F. (eds) *Intelligent Computer Mathematics. CICM 2011. Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg. — 2011. Vol. 6824. — P. 307-309 [https://doi.org/10.1007/978-3-642-22673-1\\_30](https://doi.org/10.1007/978-3-642-22673-1_30).
- [25] **Brychkov, Yu.A.** Handbook of Mellin Transforms. Chapman and Hall/CRC / Yu.A. Brychkov, O.I. Marichev, N.V. Savischenko. - Boca Raton - London - New York. — 2018. — 607 с.
- [26] Тезаурус научно-технических наук / Под ред. Ю.И. Шемакина. М.: Воениздат, — 1972. — 671 с.
- [27] **Глазунов, Н.М.** Тезаурус по теории формальных и алгебраических групп и его реализация на ЕС ЭВМ: Препр. / Н.М. Глазунов, В.М. Дриянский, Т.Н. Комарова, Ю.А. Рыбалко // АН УССР. Ин-т кибернетики; 78-27. — К.: Институт кибернетики НАНУ. 1978. — 39 с.
- [28] **Барт, К.** Русско-немецкий информационно-поисковый тезаурус по робототехнике / К. Барт, А.Т. Мицевич, А.М. Петрина, Ш. Риглер, Е.В. Сукачева, В. Тош, З.Н. Хадзиламбру. - М.: ВИНТИ. 1991. - 318 с.
- [29] **Polyanin, A.D.** Handbook of Linear Partial Differential Equations for Engineers and Scientists / A.D. Polyanin, Chapman & Hall/CRC, 2002.
- [30] Математическая энциклопедия. Гл. ред. И.М. Виноградов. Т.2. Д-КОО. - М.: Советская энциклопедия, 1979. - 1104 с. [https://dic.academic.ru/contents.nsf/enc\\_mathematics](https://dic.academic.ru/contents.nsf/enc_mathematics).
- [31] **Владимиров, В.С.** Уравнения математической физики. 2-е изд., стереотип. Учебник для вузов. / В.С. Владимиров, В.В. Жаринов. — М.: ФИЗМАТЛИТ. 2004. — 400 с.
- [32] **Трикоми, Ф.** Лекции по уравнениям в частных производных / Ф. Трикоми. Пер. с итал. — М.: КомКнига., 2007. — 443 с.

## REPRESENTATION OF MATHEMATICAL CONCEPTS IN THE ONTOLOGY OF SCIENTIFIC KNOWLEDGE

**A.A. Muromskiy<sup>1</sup>, N.P. Tuchkova<sup>2</sup>**

*Dorodnicyn Computing Centre FRC CSC of RAS, Moscow, Russia*

<sup>1</sup>*murom@ccas.ru*, <sup>2</sup>*tuchkova@ccas.ru*

### Abstract

Features of mathematical data domains in the context of creation of ontology of scientific data domains are considered. Examples from the developed thesaurus on the equations of the mixed type from the section of the equations with partial equations of mathematical physics are given. Special attention is paid to questions of representation and relevant support of the thesaurus of the addressee. It is offered to use specialized thesauruses where along with definitions in a natural language there are character expressions. This approach allows to specify a request, using mathematical entry in the TeX-notation irrespective of original language that leads to reduction of search noise and reduction of search time. The issue of why there is still insufficient amount of mathematical thesauruses created and their necessity for digital libraries along with other information resources is discussed. The example of creation of a thesaurus for separate mathematical data domain is given. The special difficulty consists in the description of the taxonomy (genus and species relations) in mathematical subject domains, one of the possible versions of the description of these communications for the equation of the mixed type is proposed.



**Key words:** search query, comparison of mathematical texts, semantic search, thesaurus of mathematical notations, TeX-notations.

**Citation:** Muromskiy AA, Tuchkova NP. Representation of mathematical concepts in the ontology of scientific knowledge [In Russian]. *Ontology of designing*. 2019. 9(1): 50-69. – DOI: 10.18287/2223-9537-2019-9-1-50-69.

## Acknowledgment

The work was partially supported by the Russian Foundation for Basic Research, projects No. 17-07-00217-a, 18-00-00297-comfi.

## References

- [1] **Roget PM.** Roget's thesaurus of English words and phrases. Taylor Anderson (Editor). — CreateSpace Independent Publishing Platform. 2017. — 660 p.
- [2] **May KO.** Historiography: A Perspective for Computer Scientists. Invited address to International Research Conference on the History of Computing, June 10, 1976, Los Alamos.
- [3] **Paulson LC.** Computational logic: Its origins and applications // *Proceedings of the Royal Society*. 2018. Series. A 474. - <https://doi.org/10.1098/rspa.2017.0872>.
- [4] **Kleschev AS, Timchenko VA.** Theoretical foundations of the shell for interactive systems of intuitive mathematical proofs verification [In Russian]. *Ontology of designing*. 2018; 8(2): 219-239. DOI: 10.18287/2223-9537-2018-8-2-219-239..
- [5] **Kleschev AS, Timchenko VA.** Implementation of the shell and knowledge portal for mathematical proofs verification on the IACPaaS platform [In Russian]. *Ontology of designing*. 2018; 8(3): 427-448. - DOI: 10.18287/2223-9537-2018-8-3-427-448..
- [6] **Lukashevich NV, Dobrov BV.** Developing linguistic ontologies in broad domains [In Russian]. *Ontology of designing*. — 2015; 5(1): 47-69.
- [7] **Zagorulko GB.** Development of ontology for intelligent scientific internet resource decision-making support in weakly formalized domains [In Russian]. *Ontology of designing*. 2016; 6(4): 485-500. DOI: 10.18287/2223-9537-2016-6-4-485-500.
- [8] **Boyack, K.W.** Thesaurus-based methods for mapping contents of publication sets // *Scientometrics*. 2017. 111: 1141. - <https://doi.org/10.1007/s11192-017-2304-3>.
- [9] **Ion PDF, Watt SM.** The Global Digital Mathematics Library and the International Mathematical Knowledge Trust. // In: Geuvers H., England M., Hasan O., Rabe F., Teschke O. (eds). *Intelligent Computer Mathematics. CICM 2017. Lecture Notes in Computer Science*. 2017. Vol. 10383. Springer, Cham - [https://link.springer.com/chapter/10.1007/978-3-319-62075-6\\_5](https://link.springer.com/chapter/10.1007/978-3-319-62075-6_5).
- [10] **Mooers C.** Information retrieval viewed as temporal signaling // *Proceedings of the International Congress of Mathematicians*. — 1950; 1: 572–573.
- [11] **Landje DV.** Search for knowledge on the Internet. Professional work [In Russian]. *Poisk znaniy v Internet. Professional'naja rabota*. — M.: Publishing house "Williams" [Izdatel'skij dom «Vil'jams»]. 2005. — 272 p.
- [12] **Moiseev EI, Muromskij AA, Tuchkova NP.** Thesaurus information retrieval in the subject area: ordinary differential equations [In Russian]. *Tezaurus informacionno-poiskovyj po predmetnoj oblasti: obyknovennye differentsial'nye uravnenija*. — M.: MAKSS Press. 2005. — 116 p.
- [13] **Moiseev EI, Muromskij AA, Tuchkova NP.** Internet and mathematical knowledge: representation of equations of mathematical physics in the information retrieval environment [In Russian]. — M.: MAKSS Press. 2008. — 80 p.
- [14] **Serebrjakov VA, Ataeva OM.** Information model of the open personal semantic library LibMeta [In Russian]. *Informacionnaja model' otkrytoj personal'noj semanticheskoy bib-lioteki LibMeta* // *Proceedings of the XVIII All-Russia. scientific conference "Scientific service on the Internet."* Novorossiysk, September 19 to September 24, 2016. Institute of PM named after M.V. Keldysh RAS. — 2016. — P.304-313.
- [15] **Moiseev EI, Muromskij AA, Tuchkova NP.** Ontology of scientific space or how to find the genius [In Russian]. *Ontology of designing*. — 2014; 4(14): 18-33.
- [16] **Miller BR, Youssef A.** Technical aspects of the digital library of mathematical functions / *Ann. Math. Artif. Intell.* — 2003; 38(1): 121–136.
- [17] **Kohlhase M.** Mathematical knowledge management: Transcending the one-brain barrier with theory graphs / *European Mathematical Society (EMS) Newsletter*, June 2014. — P.22—27.
- [18] **Zaharenko EN, Komarova LN, Nechaeva IV.** [Novyj slovar' inostrannyh slov: svyshe 25 000 slov i slovosochetaniy. 3-e izd., ispr. i dop.] New dictionary of foreign words: over 25,000 words and phrases [In Russian]. — M.: «Azbukovnik», 2008. — 1040 p.
- [19] **Bulyko AN.** Bol'shoj slovar' inostrannyh slov. — Izdatel'stvo «IDDK», 2007. — 704 p.

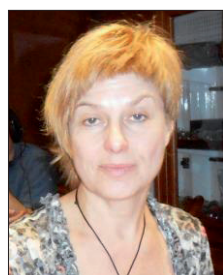
- [20] **Aleksandrova NV.** History of mathematical terms, concepts, designations [In Russian]. Istorija matematicheskikh terminov, ponjatij, oznachenij. Slovar' - spravochnik. Iz-danie tret'e, ispravlennoe. M.: Izd-vo LKI, — 2008. — 248 p.
- [21] Princeton's WordNet. G.A. Miller, R. Beckwith, C.D. Fellbaum, D. Gross, K. Miller. 1990. WordNet: An online lexical database. Int. J. Lexicograph. 3, 4. - P.235–244. - <https://wordnet.princeton.edu/>.
- [22] **Knut D.EH.** Vsyо pro TeX [In Russian]. — M.: Izdatel'skij dom «Vil'yams». 2003. — 560 p.
- [23] **Tihonov AN, Samarskij AA.** Equations of mathematical physics [In Russian]. Uravneniya matematicheskoy fiziki — M.: Izd-vo MGU, 1999. — 799 p.
- [24] **Misutka J, Galambos L.** System Description: EgoMath2 As a Tool for Mathematical Searching on Wikipedia.org. In: Davenport J.H., Farmer W.M., Urban J., Rabe F. (eds) Intelligent Computer Mathematics. CICM 2011. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. — 2011. Vol. 6824. — P. 307-309 [https://doi.org/10.1007/978-3-642-22673-1\\_30](https://doi.org/10.1007/978-3-642-22673-1_30).
- [25] **Brychkov YuA, Marichev OI, Savischenko NV.** Handbook of Mellin Transforms. Chapman and Hall/CRC. - Boca Raton - London - New York. — 2018. — 607 p.
- [26] Tezaurus nauchno-tehnicheskikh nauk / Pod.red. YU.I. SHemakina. M.: Voenizdat. — 1972. — 671 p.
- [27] **Glazunov NM, Driyanskiy VM, Komarova TN, Rybalko YUA.** [Tezaurus po teorii formal'nyh i algebraicheskikh grupp i ego realizacii na ES EHVМ] Thesaurus on the theory of formal and algebraic groups and its implementation on a computer [In Russian]. — Kiev: Institut kibernetiki NANU.1978. — 39 p.
- [28] **Bart K, Micevich AT, Petrina AM, Rigler SH, Sukacheva EV, Tosh V, Hadzilambru ZN.** Russko-nemeckij informacionno-poiskovyj tezaurus po robototekhnike [In Russian]. - M.: VINITI. 1991. - 318 p.
- [29] **Polyanin AD.** Handbook of Linear Partial Differential Equations for Engineers and Scientists, Chapman & Hall/CRC, 2002.
- [30] Mathematical encyclopedia [In Russian]. Matematicheskaja jenciklopedija / Ed. I.M. Vinogradov. V.2. D-Koo. M.: Soviet Encyclopedia, 1979. - 1104 p. [https://dic.academic.ru/contents.nsf/enc\\_mathematics](https://dic.academic.ru/contents.nsf/enc_mathematics).
- [31] **Vladimirov VS, Zharinov VV.** Uravneniya matematicheskoy fiziki. 2-e izd., stereotip. Uchebnik dlya vuzov. — M.: FIZMATLIT. 2004. — 400 c.
- [32] **Trikomi F.** Lectures on partial differential equations [In Russian]. Lekcii po uravnenijam v chastnyh proizvodnyh. Per. s ital. - M.: Komkniga, 2007. — 443 p.

## Сведения об авторах



**Муромский Александр Александрович<sup>1</sup>**, 1926 г. рождения, с.н.с., Вычислительный центр Федерального исследовательского центра «Информатика и управление» РАН (ФИЦ ИУ РАН), к.ф.-м.н., окончил механико-математический факультет МГУ им. М.В.Ломоносова и университет им. Н.Э. Баумана, долгие годы работал в ВИНТИ. Специалист в области математического анализа и информационных технологий. Автор более 50 работ в области функционального анализа и информационных технологий.

**Alexander Alexandrovich Muromskiy<sup>1</sup>** (b. 1926) senior researcher of CCAS (Federal Research Centre of Computer Science and Control RAS), PhD, graduated from mechanics and mathematics faculty of Lomonosov MSU and the university of N.E. Bauman, for many years worked in VINITI. He is a specialist in the field of the mathematical analysis and information technologies. He has published over 50 papers in the field of functional analysis and information technology.



**Тучкова Наталья Павловна**, 1955 г. рождения, с.н.с., Вычислительный центр Федерального исследовательского центра «Информатика и управление» РАН (ФИЦ ИУ РАН), к.ф.-м.н., окончила факультет вычислительной математики и кибернетики МГУ им. М.В.Ломоносова. Специалист в области алгоритмических языков и информационных технологий. Автор более 40 работ в области вычислительной математики и информационных технологий.

**Natalia Pavlovna Tuchkova** (b. 1955) senior researcher of CCAS (Federal Research Centre of Computer Science and Control RAS), PhD, graduated from CS faculty of Lomonosov MSU. She is a specialist in the field of programming languages and information technologies. She published more than 40 papers in the field of computational mathematics and information technology.

<sup>1</sup> Это статья, поступившая в редакцию в конце 2018, стала последней работой А.А. Муромского, который скоропостижно скончался в феврале 2019 года.

УДК 004.853, 004.55

## ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДЛЯ СОЗДАНИЯ СОВРЕМЕННЫХ ЭНЦИКЛОПЕДИЧЕСКИХ ПОРТАЛОВ

**Ю.В. Рогушина**

*Институт программных систем НАН Украины, Киев, Украина,  
ladamandraka2010@gmail.com*

### Аннотация

Обосновывается целесообразность разработки интеллектуальных, семантически структурированных информационных ресурсов Web, предусматривающих как их использование человеком, так и пригодность для автоматизированного анализа. Анализируются преимущества и недостатки технологии Wiki и перспективы её семантического расширения с использованием онтологического анализа. Предложена формальная модель онтологии Wiki-ресурса, на основе которой строится онтология задачи пользователя, предназначенная для использования во внешних интеллектуальных приложениях – например, при семантическом поиске. Описаны основные этапы построения этой онтологии. Рассматривается использование предложенных в работе моделей и методов на примере создания онлайн-энциклопедических изданий, объединяющих свободный доступ к материалам через среду Web с высоким уровнем доверия к контенту, разработанному экспертами. Сформированы базовые требования, предъявляемые к программному обеспечению таких проектов. Обосновывается необходимость семантизации ресурсов, анализируются основные направления развития функционала семантизированных электронных энциклопедий. На основе анализа выразительной мощности средств представления и обработки знаний, которые основываются на Wiki-технологиях, обосновывается необходимость расширения их методами искусственного интеллекта и технологиями Semantic Web. Для разработки портальной версии Большой украинской энциклопедии (е-ВУЕ) построена онтологическая модель энциклопедии. Описаны основные категории и семантические свойства типичных информационных объектов, которые используются в е-ВУЕ и могут применяться для поиска и интеграции информации.

**Ключевые слова:** Wiki-ресурс, онтология, информационный объект, семантическая разметка.

**Цитирование:** Рогушина, Ю.В. Использование онтологического анализа для создания современных энциклопедических порталов / Ю.В. Рогушина // Онтология проектирования. – 2019. – Т.9, №1(31). – С. 70-84. – DOI: 10.18287/2223-9537-2019-9-1-70-84.

### Введение

Главное направление в борьбе с информационным взрывом – переход от сохранения и обработки данных к накоплению и обработке знаний. Это определяет актуальность разработки инновационных технологий для поддержки функционирования распределённых интеллектуальных информационных систем (ИИС), обеспечивающих приобретение, хранение и использование знаний различных предметных областей (ПрО). При этом возникает проблема, связанная с наличием открытых и надёжных источников знаний, к которым могут обращаться ИИС.

Преобладающая часть современных ресурсов Web не является семантизированными, но количество информационных ресурсов (ИР), содержащих семантическую разметку и разнообразные метаописания, постоянно возрастает [1]. Пользователям сложно находить такие ресурсы в общей массе ИР, несмотря на то, что существует много поисковых систем, ориентированных именно на поиск структурированных ИР, например, для выявления, индексации

и запросов документов в формате RDF [2] или OWL [3]. Используя такие системы (например, Corese [4], ONTOSEARCH2 [5]), можно найти конкретный документ или онтологию, но сложно построить множество документов, которые соответствуют какой-либо конкретной задаче, так как непосредственные ссылки между такими ИР, как правило, отсутствуют или не являются наглядными.

Поэтому удобнее использовать базы знаний, построенные на Wiki-платформе [6]. Такие системы применяют стандартизированные средства для представления семантической разметки (с помощью системы категорий и свойств). Эти элементы можно легко распознавать даже в тех случаях, если разные информационные объекты (ИО) получены из разных ресурсов. В семантических Wiki-ресурсах всегда есть разнообразные средства для внутренней навигации и поиска, и это позволяет пользователю довольно быстро определить набор Wiki-страниц, связанных с его задачей. Важный фактор – наличие разнообразных семантических Wiki-ресурсов, количество, объём и качество которых постоянно увеличиваются. В случае, если информации в семантических Wiki недостаточно, их довольно легко дополнить сведениями с несемантизированных Wiki (например, из Википедий или Wiki-справочников). Из таких ресурсов можно получить меньше семантической информации, но в сочетании с семантизированными они позволяют довольно корректно описать произвольную проблему [7].

Следует отметить, что автоматизированное приобретение знаний значительно более эффективно осуществляется для тех ИР, которые имеют формализованную структуру и используют семантическую разметку контента (в отличие от естественных языковых или мультимедийных ИР, для которых извлечение знаний требует большего участия человека). В то же время тенденции развития ресурсов Web показывают, что в поиске источников знаний целесообразно ориентироваться на распространённые и понятные для пользователей формы представления информации.

Этим условиям удовлетворяют семантические Wiki-ресурсы, которые довольно легко интегрировать с разнообразными ИИС. Wiki представляет собой технологию коллективного создания и использования распределённых ресурсов. Она всё чаще воспринимается как новый тип коллаборативной технологии, которая может повлиять на управление знаниями, а также поддерживать их создание и совместное использование. Из различных программных средств для разработки Wiki-ресурсов наиболее широко используется MediaWiki. Именно на этом свободном программном обеспечении базируются многие всемирно известные проекты энциклопедий и справочников, такие как Wikipedia, Wikibooks, Wiktionary и Wikidata. Поэтому при создании Большой украинской энциклопедии e-ВУЕ было принято решение также ориентироваться на Wiki-технологию.

## 1 Семантические Wiki-ресурсы

Чтобы при разработке Wiki-ресурсов перейти от обработки данных к обработке и поиску знаний, используют семантические расширения. Сформированные на их основе ИР могут динамично обновляться всем сообществом пользователей, которые обеспечивают актуальность информации, имеют удобную и простую для понимания структуру, обеспечивают обработку информации на семантическом уровне, предоставляя технологическую платформу для группового управления знаниями. Одним из наиболее известных инструментов является *Semantic MediaWiki* (SMW) [8]. Эта надстройка над MediaWiki имеет высокую выразительную мощность, надёжную реализацию и удобный интерфейс пользователей. С её использованием реализован ряд успешных проектов. Следует отметить, что это были относительно небольшие тематические энциклопедии (например, энциклопедия Первой мировой войны) и порталы научных учреждений с однородным контентом. Однако знания, представленные в



Большой украинской энциклопедии, имеют значительно более сложную структуру, и поэтому возникает необходимость расширить базовые возможности SMW современными технологиями Semantic Web и средствами управления знаниями. Для этого вначале необходимо определить что именно можно представить встроенными средствами SMW.

SMW – это надстройка над инструментальным средством построения Wiki-сайта MediaWiki, которая позволяет интегрировать информацию из разных Wiki-страниц, осуществляя поиск на уровне знаний, и генерировать из Wiki-страниц онтологические структуры, которые могут использовать другие ИИС. Для организации знаний [3] в SMW используются такие механизмы, как категории, *семантические свойства* и *семантические запросы*.

*Семантические свойства* используются для привязывания данных к Wiki-страницам. Каждое свойство имеет тип, название и значение, а также собственную Wiki-страницу в специальном пространстве имен. Эта страница используется для того, чтобы задавать тип свойства, определять его место в иерархии свойств, а также документировать то, как это свойство необходимо использовать.

*Семантические запросы* позволяют интегрировать сведения из разных Wiki-страниц, осуществляя поиск на уровне знаний. В качестве параметров запросов используются не только категории, но и семантические свойства и их значения. Это значительно расширяет возможности таких запросов и обеспечивает целостность и актуальность информации.

В частности, в обычных, не семантизированных Wiki-ресурсах поиск ограничен только названиями и категориями страниц. Например, для того, чтобы найти людей, родившихся в 1800 году, необходимо создать отдельную категорию «Родившиеся в 1800 году» и присвоить её соответствующим страницам (в частности, такой подход реализован в Википедии). Такое решение является достаточно громоздким, и, что более важно, не позволяет использовать условия, например, найти людей, родившихся в интервале между 1750 и 1800 годами.

*Шаблоны* – это специальные Wiki-страницы, содержимое которых предназначено для встраивания в другие страницы. Использование шаблонов позволяет упростить и ускорить создание новых Wiki-страниц, а также обеспечить однотипное представление информации для пользователей. Значительный интерес представляет следующее: если в текст шаблона поместить семантическое свойство, то это свойство будут иметь все страницы, использующие шаблон.

Для любой ПрО, в том числе для энциклопедии, можно выделить *типичные* ИО, – объекты с подобной структурой и одинаковым набором семантических свойств. Создавая Wiki-ресурс, целесообразно разработать специальные шаблоны для таких ИО. Эти шаблоны упрощают построение Wiki-страниц и унифицируют визуализацию контента. Обычно такие шаблоны соответствуют одной или нескольким категориям или входят в состав страниц этих категорий. Шаблоны подкатегорий могут конкретизировать шаблоны для категорий более высокого уровня путём добавления семантических свойств, характерных только для этих подкатегорий.

Кроме того, такие шаблоны позволяют более эффективно и правильно вводить на страницах значения семантических свойств. Однако в использовании шаблонов ИО возникают определённые проблемы, связанные с их универсальностью: в различных экземплярах ИО могут быть определены не все значения семантических свойств, присущие этому типу ИО, а некоторые свойства могут иметь несколько различных значений. Разрабатывая шаблоны ИО, надо заранее предусмотреть такие ситуации. Это усложняет код шаблона, но обеспечивает корректное представление и обработку неполных данных.

В шаблонах SMW ситуации относительно неполноты и многозначности данных обрабатываются в отдельности, и потому необходимо заранее проанализировать семантику таких данных. Кроме того, если шаблон предусматривает визуализацию информации, связанной с



неполными и многозначными семантическими свойствами, то надо предусмотреть, чтобы в случае отсутствия такой информации не выводилось не только само значение, но и сопутствующая информация.

Использование в шаблонах семантических свойств с неполными и множественными значениями позволяет значительно уменьшить количество самих шаблонов, которые используются для описания типичных ИО. Например, в e-ВУЕ используется единый шаблон «Организация» для разных типов организаций, который содержит параметры, релевантные только для отдельных подтипов организаций. В этом шаблоне значение параметра «Вид медпомощи» может вводиться только для медицинских или ветеринарных учреждений, а значение параметра «Целевая аудитория» – только для издательств, СМИ и т.д. При меньшем количестве шаблонов при создании Wiki-страницы значительно проще выбрать соответствующий шаблон для создания статьи, а относительно небольшое количество самих шаблонов типичных ИО позволяет уделять больше внимания проверке и тестированию каждого из них.

Семантические значения, которые вводятся в шаблонах, могут использоваться в семантических запросах, которые позволяют находить Wiki-страницы по определённым требованиям, предъявляемым к этим значениям. Именно благодаря использованию шаблонов для типичных ИО можно достичь унификации в именах этих свойств, которая является значительной проблемой в разработке Wiki-ресурсов большого объёма со сложной и неоднородной структурой.

Чтобы преобразовать семантический Wiki-ресурс со сложной структурой и разнообразным гетерогенным контентом в распределённую базу знаний, к которой могут обращаться внешние ИИС, необходимо разработать средства интероперабельного представления его семантики. Предлагается использовать для этого онтологическое описание структуры Wiki-ресурса. Для этого необходимо сформировать онтологическую модель самого Wiki-ресурса, а также разработать методы её пополнения и использования для извлечения из Wiki-ресурса тех знаний, которые пертинентны той или иной задаче.

## 2 Онтологии и Semantic MediaWiki

Для разработки и поддержания сложной системы понятий семантического Wiki-ресурса, а также их свойств и отношений целесообразно применять онтологии и связанные с ними средства управления знаниями [9, 10]. Ряд таких возможностей предусмотрен непосредственно в SMW.

С точки зрения онтологического анализа, каждая Wiki-страница представляет собой онтологический элемент одного из RDF-классов – Thing, Class, ObjectProperty, DatatypeProperty, AnnotationProperty. Кроме того, каждая статья имеет собственный URI, что позволяет избежать путаницы между понятиями и HTML-страницами. Обычно статьи являются экземплярами классов онтологии OWL, категории – классами, а отношения – объектами свойствами онтологии.

Исходя из этого, с помощью специальной страницы ExportRDF для любой Wiki-страницы или набора страниц по запросу может генерироваться соответствующий OWL/RDF-файл [1, 11]. К сожалению, эта функция реализована в SMW неудачно и поддерживает ограниченный набор опций. Поскольку SMW совместима с моделью OWL DL [12], то существует возможность использования в Wiki внешних онтологий. Это возможно осуществить двумя путями: импорт онтологии позволяет создавать и модифицировать страницы в Wiki для представления отношений, заданных в некотором OWL DL-документе; а повторное использование словаря позволяет пользователям отображать Wiki-страницы на элементы существующих онтологий.

Функция импорта онтологии для чтения RDF-документов использует инструментальный RAP toolkit. Он извлекает RDF-утверждения, которые могут быть представлены в Wiki. Наименования статей импортированных элементов извлекаются с их меток (labels), или, в случае отсутствия метки, из идентификатора раздела их URI. Основной целью импорта является инициализация (первичная автоматическая загрузка) основы-шаблона для заполнения Wiki. Кроме того, импорт онтологии добавляет специальные аннотации, которые генерируют эквивалентные утверждения в экспорт OWL (owl:sameAs, owl:equivalentClass или owl:equivalentProperty). Импорт онтологий разрешён только администраторам сайта.

Импорт словаря позволяет пользователям идентифицировать элементы Wiki, указывая связь с элементами существующих онтологий. Например, Category:Person может непосредственно экспортироваться в класс foaf:Person словаря Friend-Of-A-Friend. Wiki-пользователи могут решать, какие страницы Wiki должны иметь внешнюю семантику, тем не менее набор имеющихся внешних элементов управляется только администраторами. Вводя в словарь Wiki некоторый новый элемент, они должны удостовериться в том, что повторное использование словаря соотносится с типами ограничений OWL DL. Например, внешние классы, такие, как foaf:Person, не могут быть импортированы в отношения.

Экспорт в OWL/RDF является средством обеспечения внешнего повторного использования данных из Wiki, но только практическое применение этой функции может показать качество сгенерированного RDF. Кроме того, SMW предоставляет сервис для поддержки запросов SPARQL. Система базируется на автономном RDF-сервере Joseki, синхронизированном с семантическим контентом Wiki.

### 3 Построение онтологии задачи на основе Wiki-ресурса

В ряде случаев для решения задачи пользователю нужна онтология, которая содержит знания о ПрО. Если пользователя не удовлетворяют встроенные в SMW средства построения онтологий, то он может использовать более сложный способ, при котором часть работы не может быть автоматизирована и требует его участия. Такая ситуация может возникнуть в тех случаях, если пользователю сложно построить формализованное описание ПрО, но он достаточно чётко представляет, какие именно сведения важны для его задачи. Такая онтология, в частности, может быть применена для персонифицированного поиска информации в Web, в рекомендующих системах, в задачах машинного обучения.

Для описания онтологий будем использовать формальную модель  $O = \langle X, R, F, T \rangle$ , более подробно описанную в [13], которая состоит из следующих элементов:

- $X = X_{cl} \cup X_{ind}$  – множество концептов онтологии, где  $X_{cl}$  – множество классов,  $X_{ind}$  – множество экземпляров классов;
- $R = r_{ter\_cl} \cup \{r_i\} \cup \{p_j\}$  – множество отношений между элементами онтологии, где  $r_{ter\_cl}$  – иерархические отношения между классами онтологии и свойствами классов;  $\{r_i\}$  – множество объектных свойств, которые устанавливают отношения между экземплярами классов;  $\{p_j\}$  – множество свойств данных, которые устанавливают отношения между экземплярами классов и значениями;
- $F$  – множество характеристик классов онтологии, экземпляров классов и их свойств, которые могут применяться для логического вывода (например, эквивалентность, отличие, отсутствие пересечения, область определения и область значения);
- $T$  – множество типов данных (например, строка, целое).

Формально построение онтологии задачи пользователя состоит в следующем: по онтологии ПрО  $O_{domain} = \langle X_{domain}, R_{domain}, F_{domain}, T_{domain} \rangle$  и набору Wiki-страниц  $W_{user}$ , семантиче-

ская разметка которых базируется на  $O_{\text{domain}}$ , строится «лёгкая» онтология задачи пользователя  $O_{\text{user}}$ , знания которой являются подмножеством знаний из  $O_{\text{domain}}$ . Онтология ПрО может иметь произвольную структуру, высокую выразительную мощность и быть сформирована как экспертами ПрО, так и с помощью средств получения онтологических знаний.

$O_{\text{user}} = \langle X_{\text{user}}, R_{\text{user}}, F_{\text{user}}, T_{\text{user}} \rangle$ , такая, что:

- $X_{\text{user}} \subseteq X_{\text{domain}}$ , то есть  $X_{\text{cl}_{\text{user}}} \subseteq X_{\text{cl}_{\text{domain}}}$ ,  $X_{\text{ind}_{\text{user}}} \subseteq X_{\text{ind}_{\text{domain}}}$ ;
- $R_{\text{user}} \subseteq R_{\text{domain}}$ , то является  $r_{\text{ier\_cl}_{\text{user}}} = r_{\text{ier\_cl}_{\text{domain}}}$ ,  $\{r_{\text{user}_j}\} \subseteq \{r_{\text{domain}_i}\}$ ,  $i = \overline{0, n}$ ,  $j = \overline{0, m}$ ,  $m \leq n$ ;
- $F_{\text{user}} = \emptyset$ ;
- $T_{\text{user}} \subseteq T_{\text{domain}}$ .

Такую работу целесообразно выполнять в том случае, если пользователь начинает работать над сложной проблемой, решение которой будет требовать информации на протяжении довольно значительного времени (значительно большего, чем время на построение собственной онтологии). Например, планируя исследования на несколько лет, целесообразно израсходовать несколько часов на то, чтобы в дальнейшем получать семантически отфильтрованные сведения.

### 3.1 Основные этапы построения онтологии задачи

*Этап 1.* Найти семантический Wiki-ресурс  $W$ , который по тематике соотносится с задачей пользователя или перекрывает более широкую ПрО. Проще всего использовать неспециализированные энциклопедии и справочники (такие, как e-ВУЕ), но, если пользователь располагает сведениями о более специализированных ресурсах, то их применение может увеличить эффективность работы.

*Этап 2.* Отобрать в этом Wiki-ресурсе множество страниц  $W_{\text{user}}$ , которые непосредственно связаны с задачей пользователя,  $W_{\text{user}} \subseteq W$ . Начать этот отбор можно с поиска Wiki-страниц, названия которых совпадают с ключевыми словами из описания задачи, а в дальнейшем воспользоваться одним или несколькими из следующих способов:

- с помощью встроенных средств навигации по Wiki-ресурсу переходить к страницам, соединённым с этими страницами семантическими свойствами (всеми или только теми, которые интересуют пользователя);
- воспользоваться семантическим поиском по выбранным свойствам;
- найти страницы тех же категорий.

На этом этапе пользователь может выполнить определённое количество работы самостоятельно, чтобы охарактеризовать ту информацию, которая ему нужна, и отвергнуть ту, которая не касается его текущей задачи (это может быть ценная информация, важная для ПрО в целом, но не нужная именно для решения текущей проблемы). От того, насколько точно будет выполнен отбор, зависит эффективность использования построенной онтологии в задаче пользователя: отсутствие нужной информации не разрешит находить соответствующие ресурсы, а наличие лишних страниц увеличит время обработки.

*Этап 3.* Проанализировать информацию из  $W_{\text{user}} = \{w_{\text{user}_i}\}$ ,  $i = \overline{1, s}$  для каждой страницы:

- информация о классах страницы (все или отобранные пользователем) позволяет пополнить  $X_{\text{cl}_{\text{user}}}$ , иерархические отношения между этими классами можно определить с помощью страниц этих категорий;
- имя самой страницы заносится в  $X_{\text{ind}_{\text{user}}}$ ;

- имена тех семантических свойств страниц, которые использованы на данной странице и область определения которых относится к типу «Страница» (все или отобранные пользователем), заносятся в  $\{r_{user}\}$ ;
- имена страниц, на которые данная страница ссылается с помощью семантических свойств типа «Страница» (все или отобранные пользователем), также заносятся в  $X_{ind_{user}}$ ;
- имена страниц, на которые данная страница ссылается с помощью гиперссылок (все или отобранные пользователем), также заносятся к  $X_{ind_{user}}$ ;
- если данная страница отсылает на другую страницу, то имя такой страницы рассматривается как синоним текущей страницы, заносится в  $X_{ind_{user}}$  и связывается отношением синонимии с именем текущей страницы.

При обработке несемантизированных Wiki-страниц алгоритм значительно сокращается.

### 3.2 Семантический поиск

Использование семантического поиска [14, 15] позволяет сократить время обработки, т.к. в этом случае надо обрабатывать не каждую страницу отдельно, а только результат запроса, который содержит важные для пользователя сведения в упорядоченном виде. В таком поиске в качестве условий могут быть заданы категории и условия, налагаемые на значения семантических свойств страниц.

```
{{#ask:
[[Категорія:персоналії]]
[[Рік народження::>1900]]
[[Рік народження::<1950]]
[[Місце народження::Україна]]
|?Рік народження
|?Місце народження
|?Alma mater
|?Напрями діяльності
|format=broadtable
|link=all
|headers=show
|searchlabel=... подальші результати
|class=sortable wikitable smwtable }}
```

Пользователь может вводить эти условия явным образом и определять, в каком формате получить результат – список, таблица (рисунок 1), облако тэгов (рисунок 2) и т.п.

Использование семантического поиска очень удобно как при генерации, так и при обновлении онтологий ПрО. Например, если надо составить онтологию рек и населённых пунктов определённого *Региона А*, то без семантического поиска надо обработать все страницы категорий *Города*, *Реки* и *Регион А*, а при использовании поиска – страницу, сгенерированную запросом с условиями, выделяющими на страницах категории *Реки* со значением свойства *Регион* информацию о *Городах* на берегу таких рек:

```
{{#ask: [[Категорія:Річки]]
[[Регіон::Україна]]
|?Міста на березі }}
```

Сам код запроса генерируется средствами SMW автоматически (следует отметить, что SMW позволяет вручную создавать значительно более сложные запросы, но это требует от



пользователя соответствующих навыков). Такие запросы можно встроить потом в Wiki-страницу, скопировав их код, и информация будет обновляться автоматически при изменении контента удовлетворяющих запросу страниц.

Пошук

Условия запроса

Знайти Приховати запит Показати вклучений

На запит `[[Категорія:персоналії]] [[Рік народження::>1900]] [[Рік народження::<1950]] [[Місце народження::Україна]]` було отримано відповідь із `SMWSQLStore3` за 0.0171 секунд.

Результати 1 – 40 (Попередня 100 | Наступна 100) (20 | 50 | 100 | 250 | 500) (JSON | CSV | RSS | RDF)

	Рік народження	Місце народження	Alma mater
Абаджян, Гаррій Артушевич	1939	Україна Запоріжжя (місто)	Харківський національний університет мистецтв імені Івана Котляревського Харків
Абалакін, Віктор Кузьмич	1930	Одеса Україна	Одеса Одеський національний університет імені Іллі Мечникова
Абашев-Константиновський Авраам Львович	1900	Біла Церква Київська область Україна	Національний медичний університет імені О. О. Богомольця Київ
	1945	Харків Україна	Всеросійський державний інститут кінематографії Москва

Страницы, соответствующие запросу

Значения семантических свойств

Рисунок 1 - Результаты семантического поиска в табличном виде

Пошук

Условия запроса

Знайти Приховати запит Показати вклучений

На запит `[[Категорія:персоналії]] [[Рік народження::>1900]] [[Рік народження::<1950]] [[Місце народження::Україна]]` було отримано відповідь із `SMWSQLStore3` за 0.0164 секунд.

Результати 1 – 20 (Попередня 20 | Наступна 20) (20 | 50 | 100 | 250 | 500) (JSON | CSV | RSS | RDF)

1900 1904 1905 1906 1908 1914 1930 1933 1937 1939 1944 1945 1948 1949 Сепаторія Івано-Франківськ Автономна Республіка Крим **Белівка** Біла Церква Всеросійський державний інститут кінематографії **Всесоюзний державний інститут кінематографії** **Вільшанка** Дніпро (місто) Житомирська область Запоріжжя (місто) Карлів університет **Київ** Київська область Київський національний університет театру, кіно і телебачення імені І. К. Карпенка-Карого Кропивницький Кіровоградська область Лисичанськ Луганська область **Маршинці** **Москва** Національна академія образотворчого мистецтва і архітектури Національна музична академія України ім. П. І. Чайковського Національна музична академія України імені П. І. Чайковського Національний гірничий університет України Національний медичний університет імені О. О. Богомольця Національний педагогічний університет імені М. П. Драгоманова Новоселицький район **Одеса** **Одеська національна музична академія імені А. В. Нежданової** Одеська область Одеський національний університет імені Іллі Мечникова Олександрівський район Париж **Паризька вища національна консерваторія музики й танцю** Поташ Прага Підлісне Самаркандський державний університет Сімферополь Тальнівський район **Україна** харків Харківський національний університет мистецтв імені Івана Котляревського Черкаська область Чернівецька область Чуднівський район Ясси **Яський університет**

Облако тэгов значений свойств

Рисунок 2 - Результаты семантического поиска в виде облака тэгов

Рассмотренный выше алгоритм позволяет построить онтологию задачи пользователя. Хотя существует возможность различной программной реализации, из-за того, что большинство операций нуждается во вмешательстве пользователя и принятии решения относительно



каждого понятия и отношения, на практике проще выполнять эти действия непосредственно с помощью редактора онтологий (например, Protégé [16]). Таким образом формируется онтология на языке OWL.

Основная проблема для пользователя при создании таких запросов заключается в необходимости правильно указать имена категорий и семантических свойств. Эту информацию можно получить, проанализировав код шаблонов типичных ИО, используемых на соответствующих Wiki-страницах. Но такой путь довольно сложен. Поэтому целесообразно при разработке семантических Wiki-ресурсов создавать специальную справочную страницу, описывающую вводимые в шаблонах типичных ИО семантические свойства и примеры их значений (рисунок 3).

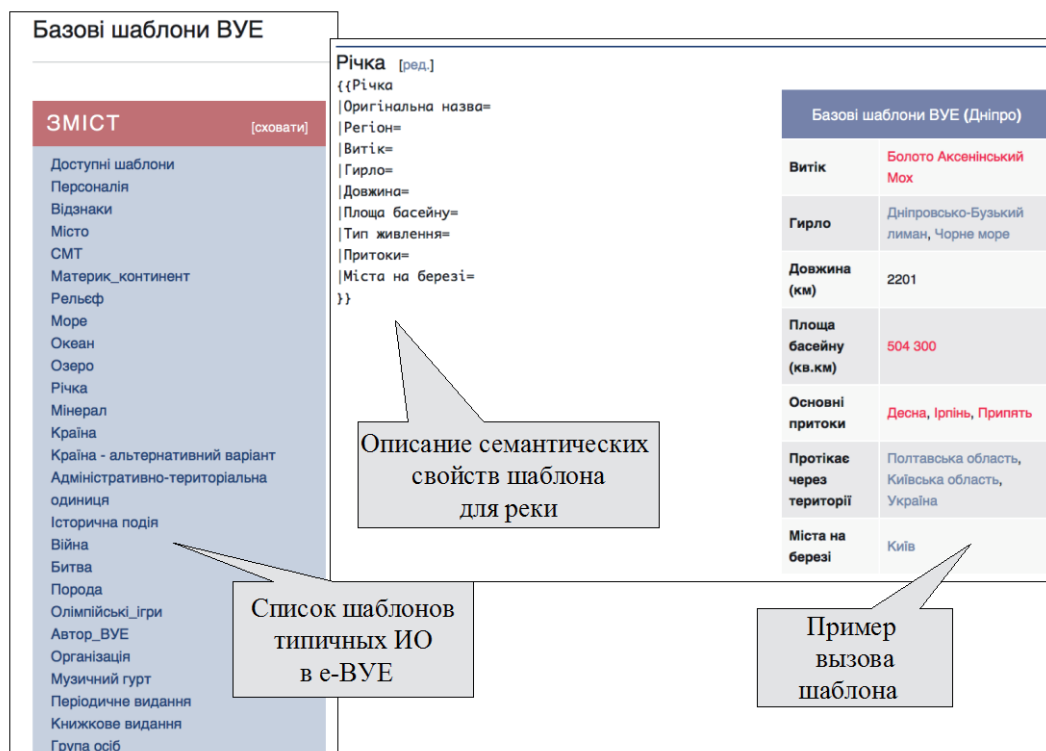


Рисунок 3 - Использование шаблонов типичных информационных объектов для построения онтологии ПрО.

### 3.3 Портальная версия Большой украинской энциклопедии (e-VUE) - семантический Wiki-ресурс

Сейчас в Украине активно ведутся работы по созданию Большой украинской энциклопедии. Качество информации, представленной в энциклопедии, обеспечивается ориентацией на рецензированные авторские статьи с оригинальным контентом, подготовленные специалистами в соответствующих областях, в которых представлены проверенные факты и признанные научным сообществом теории. e-VUE (vue.gov.ua) – портальная версия Большой украинской энциклопедии. Кроме текста, страницы e-VUE могут содержать другие типы контента (изображение, карты, видео, аудио и т.п.) и ссылки на доверенные источники. Для создания этого инновационного ИР, на базе современных знания-ориентированных технологий и оригинальных разработок проводится исследование соответствующих моделей и методов представления и обработки информации. При необходимости разрабатываются оригинальные программные решения, которые базируются на современных методах представления распределённых знаний (в частности, на технологиях Semantic Web).

Можно выделить следующие преимущества е-ВУЕ по сравнению с другими электронными справочниками и энциклопедиями:

- явное установление содержательных связей между страницами статей и их элементами;
- поиск информации по смыслу – по категориям и значениями семантических свойств страниц;
- возможность интегрировать информацию из разных статей и автоматизированно генерировать целостный контент;
- возможность экспорта знаний в форматах современных Web-технологий.

е-ВУЕ использует современные технологии и научные достижения в области управления знаниями, искусственного интеллекта, онтологического анализа, интеллектуального поиска и Semantic Web. Предполагается, что энциклопедия станет источником знаний не только для людей, но и для ИИС, которые смогут использовать сведения, экспортированные из е-ВУЕ в общепринятых форматах представления. Портальная версия Большой украинской энциклопедии использует свободное программное обеспечение MediaWiki версии 1.29.1. и его семантическое расширение SMW версии 2.5.5.

Каждая статья е-ВУЕ может быть отнесена к произвольному количеству категорий. Средства Wiki-среды позволяют явным образом указывать иерархические связи между такими категориями, которые могут отображать разные аспекты классификации статьи энциклопедии, учитывать специфику Про, условия публикации, использование материала и т.д. [17]. С точки зрения пользователей на верхнем уровне статьи подразделяются на три не пересекающиеся категории: «Персоналии», «Цивилизация» и «Природа». К каждой из этих категорий можно перейти непосредственно с главной страницы е-ВУЕ (рисунок 4).



Рисунок 4 - Категории верхнего уровня на главной странице е-ВУЕ

Кроме того, на этой же странице предусмотрен переход к иерархическому набору категорий, упорядоченному по областям знаний. В каждой из этих категорий есть соответствующие подкатегории. Например, для категории «Персоналии» – это «Учёные», «Лауреаты

Нобелевской премии» и т.д., для категории «Природа» – это «Географические объекты», «Гидронимы» и т.д. (рисунок 5).

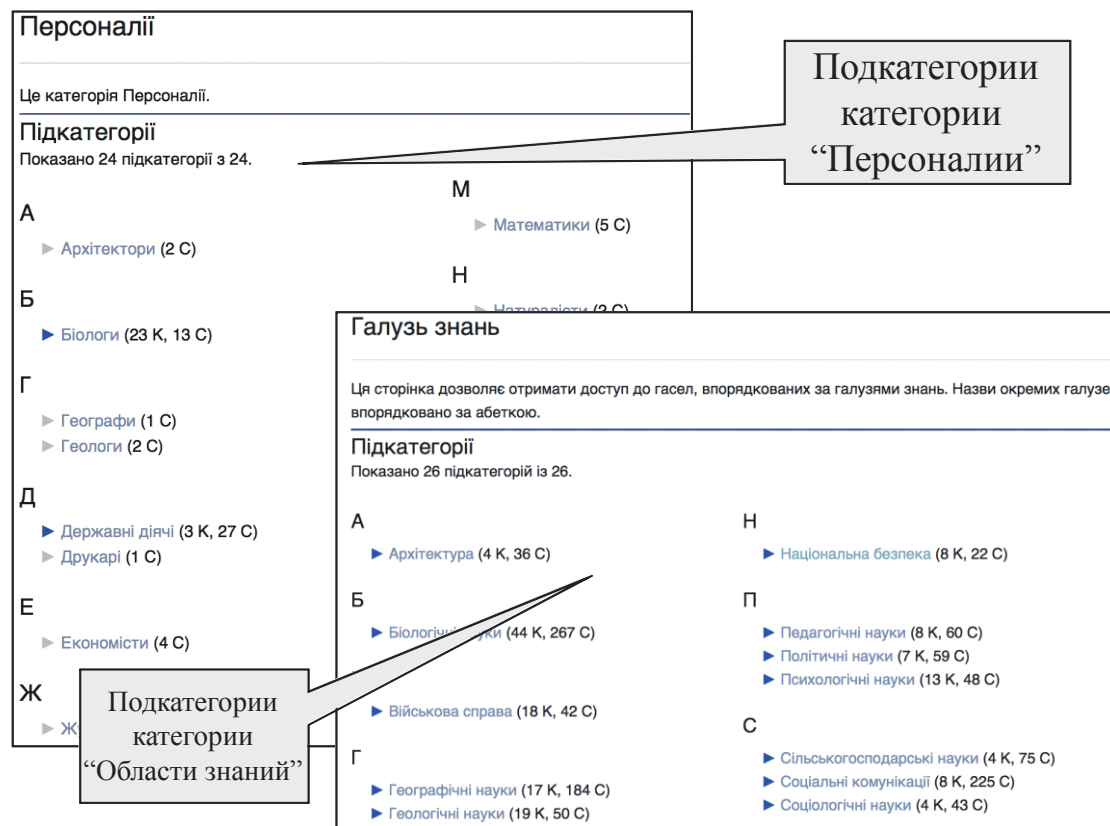


Рисунок 5 - Подкатегории областей знаний в e-VUE

#### 4 Онтологическая модель e-VUE

Для расширения функциональных возможностей портальной версии Большой украинской энциклопедии на основе её семантизации необходимо разработать методы построения полной, формализованной и однозначно интерпретируемой системы категорий и семантических свойств Wiki-страниц. Чтобы e-VUE была способная функционировать как распределённая база знаний, являющаяся источником полезной и проверенной информации как для людей, так и для внешних ИИС, необходимо создать и практически применить онтологическую модель [18] знаний энциклопедического издания. Онтологическая модель e-VUE формализует отношения между её основными объектами, их типами и свойствами. Эта модель должна удовлетворять требованиям со стороны средств анализа знаний и соответствовать специфическим ограничениям ПрО, корректно отображая её базовые закономерности. Использование этой модели как основы семантической разметки обеспечивает формирование и программную реализацию соответствующего набора иерархически связанных категорий, шаблонов типичных ИО, их семантических свойств и запросов, которые их используют. Наличие формальной модели позволяет предотвратить неоднозначную интерпретацию знаний разными разработчиками и пользователями портала.

Чтобы система категорий и семантических свойств была полной, непротиворечивой и пертинентной ПрО, целесообразно использовать существующие технологии представления и анализа знаний, ориентированные на Web-применение. Широко применяемый сегодня онто-

логический подход обеспечивает возможность визуализации знаний и их анализа специализированными инструментами. Кроме того, наличие онтологии ресурса для метаописания e-BUE значительно упрощает доступ к контенту внешних ИИС.

С помощью онтологий можно явным образом определить семантику типичных ИО Wiki-ресурса – их семантические свойства и отношение с другими ИО. Важно, что такое онтологическое представление позволяет проявлять и разрешать неоднозначные интерпретации и некорректное использование терминов, связанных с описанием ИО. Кроме того, онтология позволяет решить проблему унификации названий семантических свойств и категорий, которые используют разработчики. Онтология e-BUE (рисунок 6) определяет:

- иерархические отношения между категориями e-BUE;
- объектные свойства, связанные с семантическими свойствами страниц, которые отображают содержательные отношения между разными страницами энциклопедии (например, семантическое свойство «Место рождения» страниц категории «Персоналии» связывает их со страницами категорий «Страна» и «Город»);
- связи между категориями и шаблонами типичных ИО (одной категории может соответствовать несколько шаблонов типичных ИО (например, для категории «Персоналии» это шаблоны «Персоналия» и «Награды»), а один и тот же шаблон может использоваться на Wiki-страницах нескольких категорий (например, шаблон «Награды» используется на страницах категорий «Персоналии» и «Организации»);
- характеристики самих семантических свойств (например, симметричность, транзитивность и т.д.).

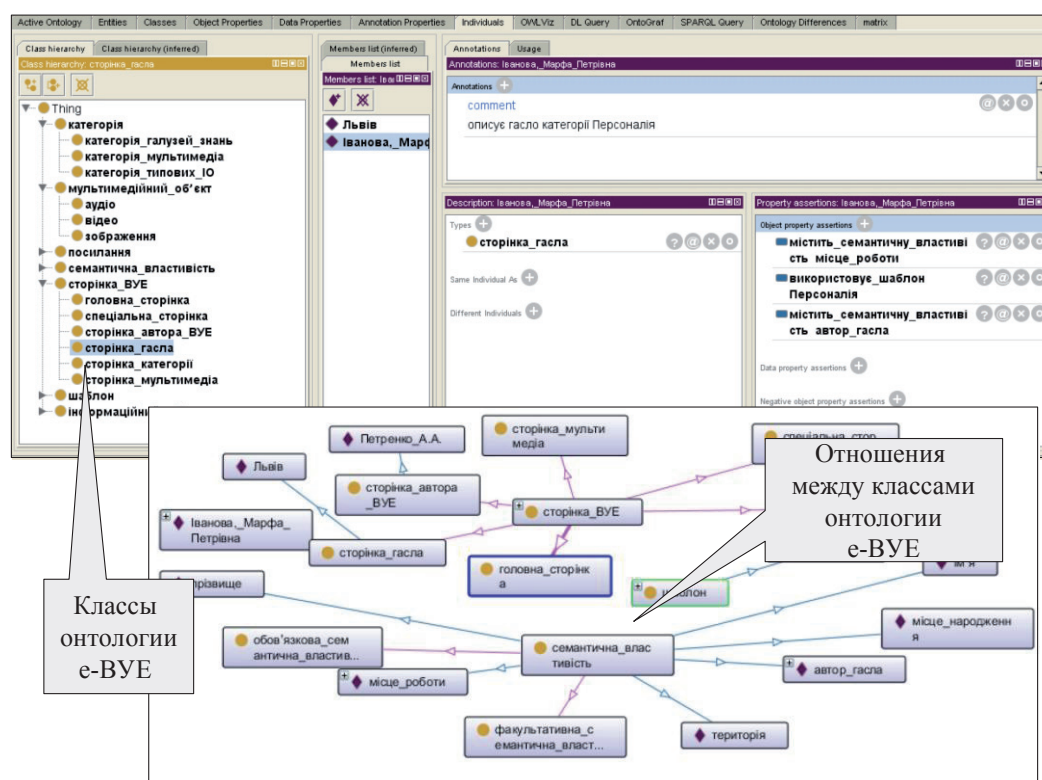


Рисунок 6 - Онтологическая модель e-BUE

Такая онтология в значительной мере зависит не только от специфики Про самого Wiki-ресурса, но и от особенностей его реализации. Поэтому для каждого Wiki-ресурса нужно сотрудничество экспертов Про с инженерами знаний, чтобы создавать оригинальную онтоло-



гию [19]. Наличие таких онтологий значительно облегчает установление семантических соответствий между различными энциклопедическими ресурсами.

Онтология е-ВУЕ является источником знаний о структуре Wiki-ресурса и помогает пользователям строить семантические запросы, обеспечивая сведения не только для правильного написания названий категорий и семантических свойств типичных ИО, но и для однозначного понимания связей между этими ИО.

## **Выводы**

Предложенные в работе модели и методы были апробированы при разработке портальной версии Большой украинской энциклопедии. Онтологический анализ обеспечил формализацию знаний о структуре знаний, содержащихся в энциклопедии. Семантизация е-ВУЕ позволяет легко создавать и экспортировать базы знаний, которые описывают определённые ПрО в общепринятых форматах (RDF). Интероперабельное представление знаний в онтологической модели обеспечивает корректное выполнение семантических запросов для формирования набора сведений, на основе которых может быть сгенерирована онтология ПрО, интересующая пользователя.

## **Благодарности**

Работа выполнена при проведении исследований по теме «Развитие информационного, функционального и программного обеспечения электронного варианта энциклопедических изданий» в рамках Программы информатизации Национальной академии наук Украины на 2018 год в Институте программных систем НАН Украины в сотрудничестве с Государственным научным учреждением «Энциклопедическое издательство».

## **Список источников**

- [1] **Lassila, O.** Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation / O. Lassila, R. Swick. — <http://www.w3.org/TR/REC-rdf-syntax>.
- [2] **OWL 2 Web Ontology Language Document Overview.** W3C. 2009. — <http://www.w3.org/TR/owl2-overview/>.
- [3] **Soumen, C.** Mining the Web: Discovering knowledge from hypertext data. Morgan Kaufmann, 2003. — 345 p.
- [4] **Corby, O.** Querying the Semantic Web with Corese search engine / O. Corby, R. Dieng-Kuntz, C. Faron-Zucker // Proc. ECAI-2004, IOS Press, 2004. — P. 705-709.
- [5] **Thomas, E.** ONTOSEARCH2: Searching ontologies semantically / E. Thomas, J.Z. Pan, D.H. Sleeman // Proc. OWLED-2007, CEUR Workshop Proceedings 258. CEUR-WS.org, 2007.
- [6] **Leuf B.** The Wiki way: collaboration and sharing on the Internet, 2001 / B. Leuf, W. Cunningham. — <http://www.citeulike.org/group/13847/article/7659081>.
- [7] **Krotzsch, M.** Semantic MediaWiki / M. Krotzsch, D. Vrandečić, M. Volkel. — <http://c.unik.no/images/6/6d/SemanticMW.pdf>.
- [8] **Semantic MediaWiki.** — [https://www.semantic-mediawiki.org/wiki/Semantic\\_MediaWiki](https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki).
- [9] **Gruber, T.R.** Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human-Computer Studies, 1995, V. 43, Issues 5-6. P. 907-928.
- [10] **Obr, L.** The evaluation of ontologies / L. Obr, W. Ceuster, I. Mani, S. Ra, B. Smith // In Semantic web: Revolutionizing Knowledge Discovery in the Life Sciences, New York: Springer Verlag, 2006, 139-158. — <https://philpapers.org/archive/OBRTEO-6.pdf>.
- [11] **OWL 2 Web Ontology Language Document Overview.** W3C. 2009. — <http://www.w3.org/TR/owl2-overview/>.
- [12] **Calvanese, D.** Tractable reasoning and efficient query answering in description logics: The DL-Lite family / Calvanese D., De Giacomo G., Lembo D., Lenzerini M., Rosati R. // JAR, 39(3), 2007. — P. 385-429.
- [13] **Rogushina, J.** Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies / J. Rogushina // International Journal of Mathematical Sciences and Computing (IJMSC), Vol.3, No.3, 2017. — P.50-58. — <http://www.mecspress.org/ijmsc/ijmsc-v3-n3/IJMSC-V3-N3-5.pdf>.



- [14] **Hendler, J.** Web 3.0: The dawn of semantic search / J. Hendler // *Computer*, 2010, 43(1), P.77-80.
- [15] **Baeza-Yates, R.** Next generation Web search / R. Baeza-Yates, A. Raghavan R. // S. Ceri and M. Brambilla, editors, *Search Computing*, Springer, 2010, P.11-23.
- [16] Protégé. – <http://protege.stanford.edu/>.
- [17] **Rogushina, J.V.** The Use of Ontological Knowledge for Semantic Search of Complex Information Objects / J.V. Rogushina // *Open semantic technologies for intelligent systems, OSTIS-2017, Minsk, 2017*, P.127-132.
- [18] **Rogushina, J.** Semantic Wiki resources and their use for the construction of personalized ontologies / J. Rogushina // *CEUR Workshop Proceedings 1631, 2016*, P.188-195.
- [19] **Missikoff, M.** The usable ontology: An environment for building and assessing a domain ontology / M. Missikoff, R. Navigli, P. Velardi // *International semantic web conference, 2002*, P. 39-53.

## USE OF ONTOLOGICAL ANALYSIS FOR CREATION OF ENCYCLOPEDIC PORTALS

**J.V. Rogushina**

*Institute of Software Systems of the National Academy of Sciences of Ukraine, Kyiv, Ukraine,  
ladamandraka2010@gmail.com*

### Abstract

We analyze the expediency of the development of the intelligent, semantically structured information Web resources that are oriented both on their use by humans and their suitability for automated analysis. The advantages and disadvantages of the Wiki technology and the perspectives of its semantic expansion with the use of ontological analysis are considered. Formal ontological model of the Wiki-resource that underlies the development of the user task ontology oriented for utilization by external intelligent applications (for example, for semantic search) is proposed. The main stages of construction of this ontology are described. The implementation of the proposed ontology-based approach is demonstrated on example of the development of on-line encyclopaedias combining free access to materials through the Web with a high level confidence in content developed by experts, is considered. We analyze the basic requirements for software that provides such projects. The necessity of semantization of such resources is substantiated and the main directions of development of the functional of semantic electronic encyclopedias are formulated. The article presents an analysis of the expressive power of the Wiki means for knowledge representation and processing. The need to expand them with artificial intelligence methods and Semantic Web technologies is stated. An ontological model of the encyclopedia serving for development of the portal version of the Great Ukrainian Encyclopedia is built. The main categories and semantic properties of typical information objects that are used in this Encyclopedia that can be applied for retrieval and integration of information are described.

**Key words:** *Wiki-resource, ontology, information object, semantic markup.*

**Citation:** *Rogushina JV. Use of ontological analysis for creation of encyclopedic portals [In Russian]. Ontology of designing. 2019. 9(1): 70-84. – DOI: 10.18287/2223-9537-2019-9-1-70-84.*

### Acknowledgment

The work was carried out as a part of a research “Development of information, functional and software of the electronic version of encyclopedic issues” as a part of the Informatization Program of the National Academy of Sciences of Ukraine for 2018 at the Institute of Software Systems of the NAS of Ukraine in collaboration with the State Scientific Institution “Encyclopedic Publishing House”.

### References

- [1] **Lassila O, Swick R.** Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation. — <http://www.w3.org/TR/REC-rdf-syntax>.
- [2] **OWL 2 Web Ontology Language Document Overview.** W3C. 2009. — <http://www.w3.org/TR/owl2-overview/>.

- [3] Soumen C. Mining the Web: Discovering knowledge from hypertext data. Morgan Kaufmann, 2003. — 345 p.
- [4] **Corby O, Dieng-Kuntz R, Faron-Zucker C.** Querying the Semantic Web with Corese search engine // Proc. ECAI-2004, IOS Press, 2004. — P. 705-709.
- [5] **Thomas E, Pan JZ, Sleeman DH.** ONTOSEARCH2: Searching ontologies semantically // Proc. OWLED-2007, CEUR Workshop Proceedings 258. CEUR-WS.org, 2007.
- [6] **Leuf B, Cunningham W.** The Wiki way: collaboration and sharing on the Internet, 2001. — <http://www.citeulike.org/group/13847/article/7659081>.
- [7] **Krotzsch M, Vrandeć D, Volkel M.** Semantic MediaWiki, — <http://c.unik.no/images/6/6d/SemanticMW.pdf>.
- [8] Semantic MediaWiki. — [https://www.semantic-mediawiki.org/wiki/Semantic\\_MediaWiki](https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki)
- [9] **Gruber TR.** Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human-Computer Studies, 1995, V. 43, Issues 5-6. P. 907-928.
- [10] **Obr L, Ceuster W, Mani I, Ra S, Smith B.** The evaluation of ontologies // In Semantic web: Revolutionizing Knowledge Discovery in the Life Sciences, New York: Springer Verlag, 2006, 139-158. — <https://philpapers.org/archive/OBRTEO-6.pdf>.
- [11] OWL 2 Web Ontology Language Document Overview. W3C. 2009. — <http://www.w3.org/TR/owl2-overview/>.
- [12] **Calvanese D, De Giacomo G, Lembo D, Lenzerini M, Rosati R.** Tractable reasoning and efficient query answering in description logics: The DL-Lite family // JAR, 2007; 39(3): 385-429.
- [13] **Rogushina J.** Analysis of Automated Matching of the Semantic Wiki Resources with Elements of Domain Ontologies // International Journal of Mathematical Sciences and Computing (IJMSC), 2017; 3(3): 50-58. — <http://www.mecs-press.org/ijmsc/ijmsc-v3-n3/IJMSC-V3-N3-5.pdf>.
- [14] **Hendler J.** Web 3.0: The dawn of semantic search // Computer, 2010, 43(1): 77-80.
- [15] **Baeza-Yates R, A. Raghavan R.** Next generation Web search // S. Ceri and M. Brambilla, editors, Search Computing, Springer, 2010, P.11-23.
- [16] Protégé. — <http://protege.stanford.edu/>.
- [17] **Rogushina JV.** The Use of Ontological Knowledge for Semantic Search of Complex Information Objects // Open semantic technologies for intelligent systems, OSTIS-2017, Minsk, 2017, P.127-132.
- [18] **Rogushina J.** Semantic Wiki resources and their use for the construction of personalized ontologies // CEUR Workshop Proceedings 1631, 2016, P.188-195.
- [20] **Missikoff M, Navigli R, Velardi P.** The usable ontology: An environment for building and assessing a domain ontology // International semantic web conference, 2002, P. 39-53.

## Сведения об авторе



**Рогущина Юлия Витальевна**, 1967 г. рождения. Окончила факультет кибернетики Киевского государственного университета им. Т.Г. Шевченко, кандидат физико-математических наук. Старший научный сотрудник Института программных систем НАН Украины. Автор более 200 научных публикаций, среди которых монографии и учебники в области онтологического анализа, семантического поиска, интеллектуальных агентов и менеджмента знаний. - <http://orcid.org/0000-0001-7958-2557>.

**Yulia Vitalyevna Rogushina**, born in 1967. Graduated from the Faculty of Cybernetics of Kiev State University named after TG Shevchenko, Candidate of Physical and Mathematical Sciences. Senior Researcher of the Institute of Software Systems of the National Academy of Sciences of Ukraine. Author of more than 200 scientific publications, including monographs and text-

books in the field of ontological analysis, semantic search, intelligent agents and knowledge management. - <http://orcid.org/0000-0001-7958-2557>.

УДК 681.3, 694.1

## ПРЕДПРОЕКТНЫЙ АНАЛИЗ ИНТЕЛЛЕКТУАЛЬНОГО ЖИЛОГО ДОМА ДЛЯ УСЛОВИЙ АРКТИКИ

Н.М. Боргест<sup>1,a</sup>, А.С. Галахарь<sup>2,b</sup>, М.В. Овсянников<sup>2,c</sup>, Р.О. Самсонов<sup>1,d</sup>

<sup>1</sup> Самарский национальный исследовательский университет им. академика С.П. Королева, Самара, Россия

<sup>a</sup>borgest@yandex.ru, <sup>d</sup>samsonov@ssau.ru

<sup>2</sup> Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия

<sup>b</sup>a.galakhar@bmstu.ru, <sup>c</sup>mvo50@mail.ru

### Аннотация

Сочетание неблагоприятных природных условий и малой населённости арктических пространств усложняет подходы к проектированию и строительству жилья для этих мест и определяет требования к автоматизированным средствам жизнеобеспечения, внутреннему устройству и надёжности жилых домов. В статье рассматривается проблема структурного синтеза интеллектуального жилого дома для арктических условий. Вводится классификация условий ограничения, накладываемых на конструктивные решения, принимаемые при проектировании. Ставится задача удовлетворения ограничений применительно к разработке конструктивных решений и автоматизации здания. В качестве цели назначается получение допустимого решения задачи синтеза на основании формализованных знаний об условиях эксплуатации, при котором целевая функция принимает наименьшее значение. Рассмотрены вопросы выбора целевых функций. Предложен подход к созданию новых привлекательных проектов интеллектуальных жилых домов для условий Крайнего Севера России, позволяющих обеспечить уют и безопасность с использованием последних достижений инженерной мысли в области автоматизации и надёжности автоматизированных систем. Привлекательное и уютное жильё поможет решить вопрос привлечения людей для обживания арктических пространств, их исследования, освоения и сохранения для потомков.

**Ключевые слова:** Арктика, проектный анализ, интеллектуальное здание, онтология, структурный синтез, удовлетворение ограничений.

**Цитирование:** Боргест, Н.М. Предпроектный анализ интеллектуального жилого дома для условий Арктики / Н.М. Боргест, А.С. Галахарь, М.В. Овсянников, Р.О. Самсонов // Онтология проектирования. – 2019. – Т. 9, №1(31). – С.85-100. – DOI: 10.18287/2223-9537-2019-9-1-85-100.

### Введение

В настоящее время большое внимание уделяется развитию Арктической зоны России [1], плотность населения которой, в целом, не превышает 0,5 человека на 1 км<sup>2</sup> вследствие неблагоприятных природных условий, объяснимых климатом: частая повторяемость отрицательных температур, недостаток или полное отсутствие (полярная ночь) солнечного света, сильные ветры [2]. Однако, даже при существующей плотности населения на севере России проживает около 2,5 млн. человек, что уже требует создания комфортной среды в арктических поселениях [3], особенно с учётом того, что перемещение в Арктику производств и населения будет продолжаться [4]. Недостаток удовлетворительных жилищных условий на севере сдерживает развитие этих территорий и отрицательно сказывается на здоровье местных жителей [5]. Для обеспечения приемлемого качества жизни людей в холодном климате в малоосвоенных местах необходимо строительство малоэтажных зданий, оснащённых современными техническими средствами автоматизированного мониторинга [6].

Строительство и обслуживание зданий на севере затруднено особенностями арктической окружающей среды, переменчивой погодой, коротким летом и ограниченными местными

ресурсами [5]. Чтобы улучшить жилищные условия на севере и обеспечить долгосрочное народнохозяйственное развитие Арктики, необходимы технологии и конструктивные решения, соответствующие условиям севера [5, 7]. При выработке подхода к проектированию жилых зданий для условий Арктики целесообразно учитывать не только особенности условий дальнейшей эксплуатации здания, но и запросы каждого заказчика, отражающие значимые условия жизни в доме (см. например, [8]).

## 1 Понятие «интеллектуальное здание»

Основным направлением уменьшения расходов на эксплуатацию жилых зданий в течение срока службы при сохранении качества жизни является внедрение средств автоматизации. Существуют 3 уровня автоматизации:

- 1) управление инженерным оборудованием здания;
- 2) объединение отдельных автоматизированных средств управления воедино;
- 3) управление эксплуатацией.

Понятие «интеллектуальное здание» (ИЗ) было введено в начале 1980-х для обозначения здания, имеющего все три уровня автоматизации [9]. По определению Вашингтонского института интеллектуальных зданий (Washington Intelligent Building Institute): «ИЗ — это такое здание, которое обеспечивает производительную и выгодную по затратам среду посредством совершенствования её основных составляющих: конструкций, устройств, служб и управления, а также взаимосвязей между ними» [10].

По определению Европейской группы интеллектуальных зданий (European Intelligent Building Group) — «это такое здание, которое увеличивает производительность жильцов здания, давая в то же время возможность выгодного управления ресурсами при наименьших затратах в течение срока службы на оборудование и средства обслуживания» [11].

Согласно строительным правилам Республики Казахстан: «3.14 интеллектуальное здание: Здание, оснащённое автоматизированной системой управления комплексом систем безопасности, жизнеобеспечения, информатизации, с возможностью объединения в систему диспетчеризации инженерного оборудования здания с единым центром мониторинга» [12]. Для дефиниции понятия ИЗ была предложена классификация определений, в основу которых положены [13]:

- *отдача от здания* (на ожиданиях и потребностях пользователей, но интеллектуальным средствам и электронным технологиям в них уделяется меньше внимания);
- *оснащённость здания вычислительными устройствами* (новейшей техникой и встроенным интеллектом в связи с запросами пользователей);
- *эксплуатационное удобство здания* (эксплуатационные качества здания).

В статье президента АВОК<sup>1</sup> [14] даны сразу несколько определений ИЗ, соответствующие различным классам предложенной классификации. Однако ИЗ для арктических условий должно, в первую очередь, обеспечивать эксплуатационное удобство, позволяющее обеспечить необходимое качество жизни, поэтому оно в большей степени соответствует определению Вашингтонского института интеллектуальных зданий.

## 2 Условия эксплуатации жилого дома

Возможны несколько классификаций условий эксплуатации жилого дома в Арктике:

- по климатическим параметрам;

---

<sup>1</sup>НП «АВОК» — Некоммерческое Партнёрство «Инженеры по отоплению, вентиляции, кондиционированию воздуха, теплоснабжению и строительной теплофизике». - <https://www.abok.ru/>.

- по удалённости от мест с населением не менее 25 тыс. человек;
- по грунтам основания;
- по рельефу местности;
- по удалённости от источников водоснабжения;
- по времени непрерывной эксплуатации;
- по сейсмичности места строительства.

Классификация условий эксплуатации по климатическим параметрам [2] позволяет установить условия теплопередачи с наружной поверхности здания.

Классификация по удалённости от мест с населением не менее 25 тыс. человек позволяет определить порядок утилизации бытовых отходов, требования к обеспечению надёжности здания и его инженерного оборудования, а также набору запасных частей, инструментов и принадлежностей (ЗИП), а также установить возможности оказания неотложной медицинской помощи. Выбор наименьшей численности населения обусловлен двумя причинами:

- современные мусороперерабатывающие заводы строятся в местах с численностью населения не менее 25 тыс. человек [15];
- согласно приказу Минздрава РФ [16] одна городская больница предусматривается в местах с численностью населения от 20 до 300 тыс. человек.

Классификация по грунтам основания, соответствующая ГОСТ 25100 [17, 18], позволяет выбрать наиболее приемлемую конструкцию фундамента, а также систем отопления, водоснабжения и водоотведения.

Условия эксплуатации зависят также от рельефа местности, т.к. рельефом определяется воздействие природных осадков на грунт и несущие конструкции дома. В Арктике представлены различные виды рельефа от равнинного до горного [19].

Удалённость от источников водоснабжения определяет требования к необходимому запасу воды и производительности системы водоснабжения.

Классификация условий по времени непрерывной эксплуатации здания в течение года [20, 21] позволяет определить требования к запасам топлива и предметов первой необходимости, а также к средствам автоматизации здания.

Классификация условий эксплуатации по сейсмичности места строительства [22] позволяет выбрать решения, обеспечивающие необходимую сейсмостойкость здания.

### 3 Классификация составных частей интеллектуального здания

Основными составными частями ИЗ для условий севера являются:

- *несущие и ограждающие конструкции* (фундамент, перекрытия, крыши, лестницы, стены, окна, двери, теплоизоляция и т.д.), которые содержат датчики для мониторинга технического состояния и обмена энергией с окружающей средой;
- *инженерное оборудование* (вентиляция, отопление, водоснабжение, водоотведение, электроснабжение, освещение и т.д.), каждый вид которого оснащается датчиками технического состояния и собственным устройством управления;
- *устройства управления* инженерным оборудованием здания, которые связаны через локальную сеть;
- *информационно-измерительная и управляющая система* (ИИУС) [23].

ИИУС объединяет в себе средства управления окружающей средой здания (переработка отходов, водоснабжение и водоотведение, энергосбережение, энергоснабжение, отопление), управления безопасностью (мониторинг и прогнозирование технического состояния инженерного оборудования и конструкций здания, автономное пожаротушение, защита от проникновения в дом посторонних лиц, видеонаблюдение), управления бытом (хранение запаса-



сов продовольствия и предметов первой необходимости, приготовление пищи, стирка и глажение одежды), управление средствами связи.

Особенностями арктической зимы, кроме низких температур, являются полное отсутствие солнечного света и ветреная погода с большим количеством осадков [2]. В арктических условиях строительные материалы и защитно-декоративные покрытия быстрее накапливают повреждения, а устранение дефектов обходится многократно дороже, чем в обжитых местах [6]. Поэтому принципиальным отличием жилого дома для условий Арктики являются строгие требования к надёжности как строительных конструкций, так и средств жизнеобеспечения — в условиях сурового климата и малой плотности населения возможности быстрого оказания помощи жильцам в случае выхода из строя систем связи, энергоснабжения, отопления, водоснабжения и водоотведения весьма ограничены. Основные задачи, решаемые при проектировании жилого дома для условий Арктики в связи с особенностями окружающей природной среды, и возможные способы решения представлены в таблице 1.

Таблица 1 – Особенности проектирования жилых зданий для условий Арктики

№	Особенности окружающей среды	Задачи проектировщика	Способы решения задач
1	Низкие зимние температуры	Обеспечить уютную среду внутри жилого здания	Энергоэффективная теплоизоляция помещений, бережливое отопление и вентиляция
2	Вечная мерзлота	Исключить влияние оттаивания вечной мерзлоты на несущую способность конструкций здания	Теплоизоляция вечномёрзлого грунта от нагреваемых конструкций здания
3	Снег	Избежать накопления большого количества снега на конструкциях и вблизи здания.	Автоматическое снегоудаление, остекление балконов, выбор конструкций и материалов, на которые не налипает снег
4	Недостаток солнечного света	Добиться необходимой освещённости при наименьших энергозатратах	Наибольшее использование солнечного света для освещения <sup>2</sup> с защитой от перегрева помещений, использование энергоэффективных осветительных приборов
5	Скудная природная растительность	Озеленение внутренних помещений здания	Устройство зимних садов [4]
6	Долгая зима	Создать возможности для полноценного отдыха в стенах жилого дома	Выделение помещений для занятий физической культурой, самообразования
7	Удалённость от обжитых мест	Уменьшить трудозатраты на удовлетворение бытовых нужд и обеспечение жизнедеятельности	Создание автономных систем управления безопасностью, надёжностью здания и его инженерного оборудования, автоматизация управления бытом

Для создания привлекательных условий жизни за полярным кругом жилые дома целесообразно проектировать в расчёте на проживание одной-двух семей. Большая автономность каждого жилого дома, высокая надёжность инженерного оборудования и обеспеченность современными средствами связи сделают жизнь в таких домах удобной и безопасной вне зависимости от количества зданий и инфраструктуры посёлка.

В качестве мест для постройки арктических жилых домов желательно выбирать возвышенные участки со скальным грунтом. Скальный грунт является наиболее прочным и не теряет несущую способность при повышении температуры. Возвышенное место позволяет использовать больше ветровой энергии для энергоснабжения дома, осуществлять водоотведение из всех помещений в водоочистное сооружение под действием только силы тяжести и избежать возможного подтопления при потеплении [24].

<sup>2</sup> Этот приём был впервые описан Витрувием [25].

Для обеспечения жилого дома электроэнергией и теплом при сильных ветрах основным источником энергии целесообразно выбрать автономную ветроэнергетическую установку. Дизельный электрогенератор или топливный элемент, рассчитанные на использование при длительной безветренной погоде или ремонте ветроэнергетической установки, должны иметь запас топлива, достаточный для обеспечения непрерывного энергоснабжения здания в течение времени между плановыми поставками топлива.

Объединение автономных источников энергоснабжения отдельных поселковых домов в единую сеть позволит повысить надёжность энергоснабжения посёлка в целом.

Предполагаемое устройство ИЗ для условий Арктики показано на рисунке 1. Для бесперебойной подачи в электрическую сеть дома переменного напряжения питания промышленной частоты ветроэнергетическая установка дополняется аккумулятором и инвертором (5). Для отопления дома энергетически выгодно использовать тепловой насос (4), охлаждающий грунт, на котором расположено здание. Тепло, отводимое от теплового насоса (4), используется для нагрева воды, используемой для бытовых нужд, а также в качестве теплоносителя системы водяного отопления. Чтобы сгладить неравномерность расхода электроэнергии на подачу водопроводной воды в течение дня, используются накопительные ёмкости для воды. Холодная вода накапливается в ёмкости (1), расположенной в чердачном помещении, и подаётся в трубы водоснабжения под действием силы тяжести. Получение и накопление горячей воды осуществляется в термостатированной ёмкости (3). Вентиляция в ИЗ для арктических условий должна соответствовать требованиям к энергосбережению, поэтому предусматривается естественная вентиляция с рекуперацией тепла и возможность включения принудительной вентиляции по показаниям датчиков  $\text{CO}_2$ . Управление вентиляцией и лифтами логично объединить, поскольку лифт является замкнутым помещением, естественный воздухообмен которого с окружающей средой затруднён. Управление климатом в жилом доме предполагает учёт влияния погодных условий, определяемых с использованием метеорологического датчика, на температуру и влажность во внутренних помещениях здания при поддержании температуры и влажности внутреннего воздуха в допустимых пределах [26, 27]. Для обеспечения требуемых показателей влажности и предохранения конструкций дома от разрушительного воздействия воздушных осадков управление климатом разумно совместить с управлением автоматической очисткой крыши от льда и снега и обогревом водостоков.

В силу малой плотности населения на севере России и большими расстояниями между отдельными населёнными местами для обеспечения пожарной безопасности жилого здания целесообразно использование автономной системы пожаротушения, которая приводилась бы в действие по показаниям датчиков изменения температуры и содержанию дыма и  $\text{CO}_2$  в воздухе в различных помещениях. Ёмкости со смесью для пожаротушения (2) энергетически выгодно разместить также в чердачном помещении и подавать в очаг возгорания по сети пожарных трубопроводов.

Автоматическая система автономного пожаротушения является одним из средств обеспечения безопасности здания. Средствами обеспечения безопасности являются также: камеры видеонаблюдения, позволяющие обеспечить защиту от посторонних лиц или приблизившихся к дому диких животных; датчики протечек и средства автоматического отключения воды, позволяющие предотвратить разрушение строительных конструкций в случаях неисправностей сетей водопровода, водоотведения или водяного отопления; датчики напряжённо-деформированного состояния конструкций, позволяющие обеспечить необходимую надёжность здания и его инженерного оборудования.

Для сбережения электроэнергии, расходуемой на освещение здания и прилегающего к нему участка, целесообразно использование автоматической системы управления освещением с учётом показаний сумеречных датчиков и датчиков движения.

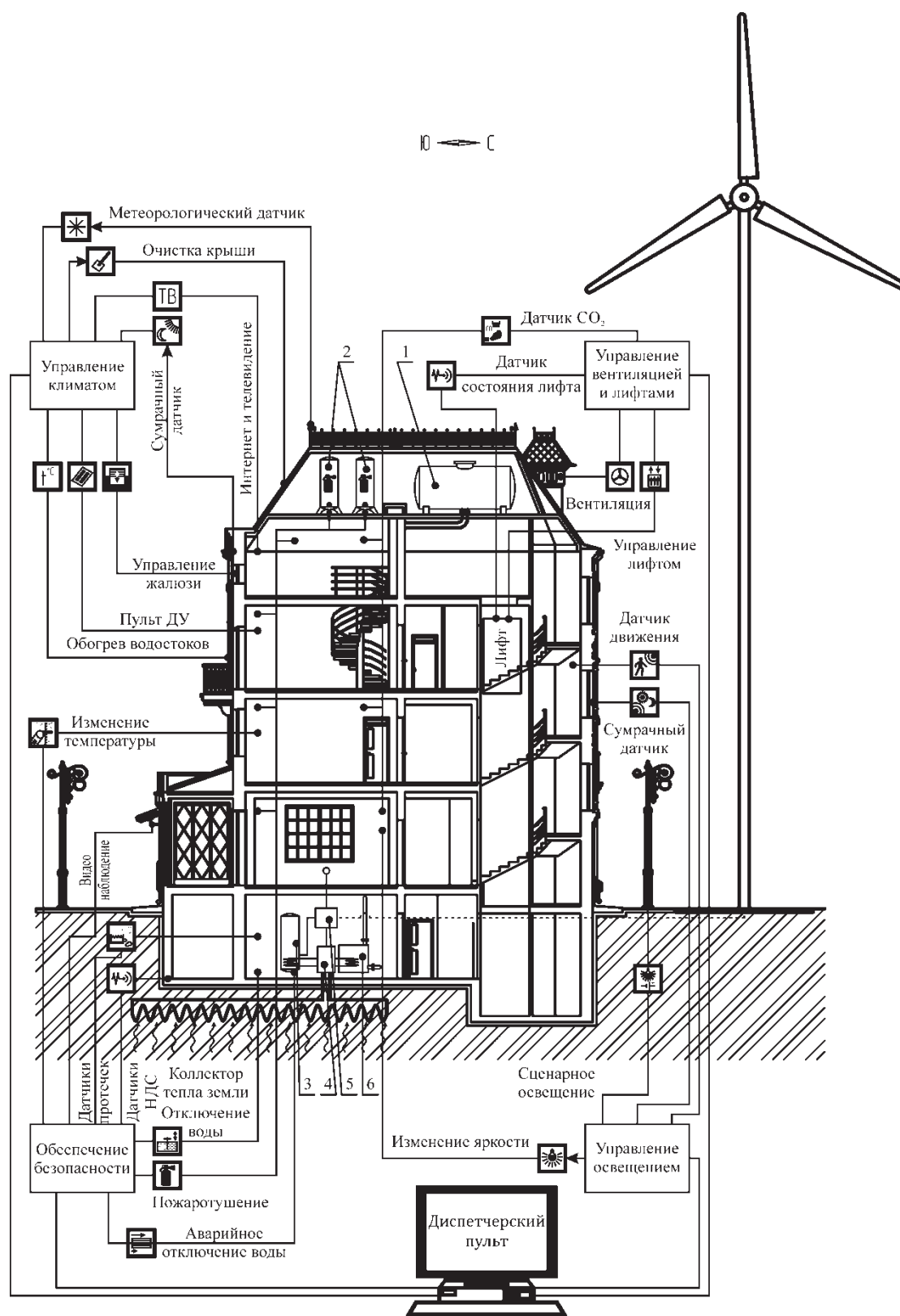


Рисунок 1 — Составные части интеллектуального жилого дома и его инженерное оборудование:  
 (1) ёмкость с запасом воды; (2) ёмкости со смесью для пожаротушения; (3) ёмкость с водонагревателем;  
 (4) тепловой насос; (5) аккумулятор и инвертор ветроэнергетической установки; (6) утилизатор тепла стоков

Для объединения отдельных автоматизированных средств управления инженерным оборудованием воедино предусмотрен диспетчерский пульт (см. рисунок 2). Диспетчерский пульт позволяет по локальной вычислительной сети собрать данные со всех датчиков систем мониторинга, обеспечивающих безопасность, а также управление климатом, освещением, вентиляцией и лифтами, согласовать работу отдельных автоматических систем, оценивать в реальном времени техническое состояние инженерных конструкций здания и показатели его надёжности.

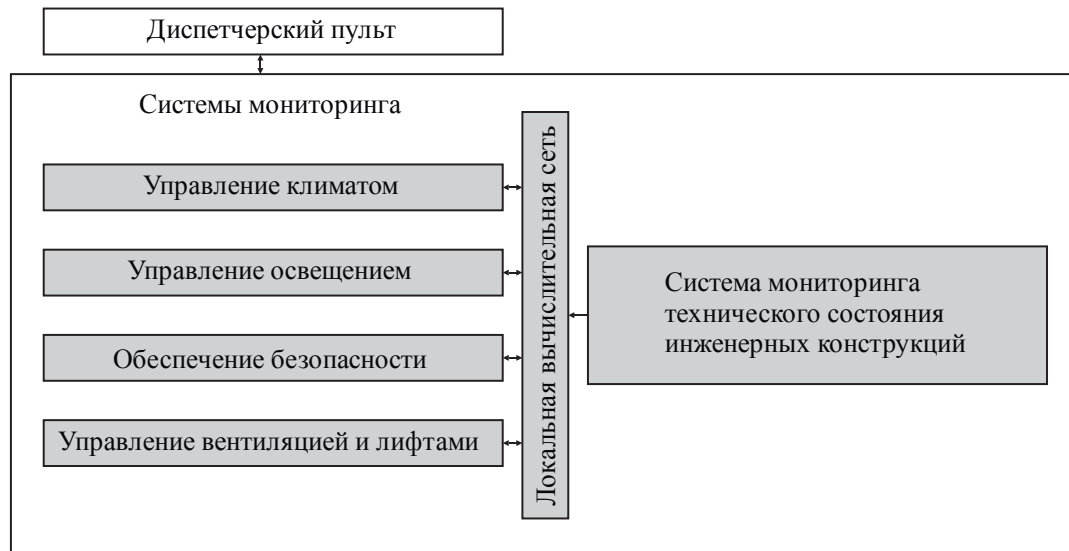


Рисунок 2 — Мониторинг технического состояния жилого дома

Оценка технического состояния инженерных конструкций по данным непрерывных наблюдений, осуществляемых с использованием показаний датчиков и диспетчерского пульта, позволяет предсказывать изменение технического состояния здания и его инженерного оборудования, а также оценивать риск его эксплуатации. Оценка риска эксплуатации в реальном времени позволяет планировать техническое обслуживание и ремонт (ТОиР), обеспечивающие необходимую надёжность с наименьшими затратами, а также состав и сроки пополнения набора ЗИП для поддержания всего инженерного оборудования в работоспособном состоянии (рисунок 3).

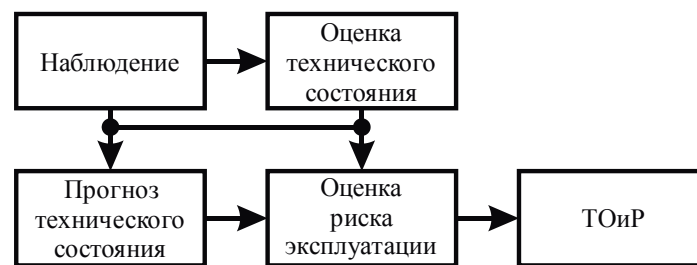


Рисунок 3 — Планирование ТОиР

Для оценки интеллектуальности здания было предложено использовать три разнородных показателя [28]: уровень техники; назначение; бережливость. При оценке интеллектуальности жилого дома для арктических условий предложенные показатели можно определить следующим образом.

*Уровень техники* — отношение риска эксплуатации здания, не оснащённого датчиками и системами автоматического управления, к риску эксплуатации такого здания, оснащённого необходимыми датчиками и системами автоматического управления.

*Назначение* — отношение количества средств управления, объединённых ИИУС, к общему количеству средств управления, которые могли бы быть установлены в здании.

*Бережливость* — отношение стоимости жизненного цикла (ЖЦ) здания, не оснащённого датчиками и системами автоматического управления, к стоимости ЖЦ такого здания, оснащённого необходимыми датчиками и системами автоматического управления.

Более развёрнутые показатели оценки ИЗ [29] для жилого дома в арктических условиях могут быть определены следующим образом:

- *дружественность к окружающей среде* — строение не наносит ущерба окружающей среде, способствует сбережению энергии и воды, безопасному удалению и переработке отходов;
- *использование пространства и приспособляемость* — количество целей, для которых можно использовать помещения здания, и трудоёмкость переоборудования помещений для целевого использования;
- *отдача на затраты* в течение всего срока службы;
- *здоровье и благополучие людей* — среднее количество дней нетрудоспособности в течение года у жильцов дома;
- *производительность труда и трудовая отдача* — величина, обратная трудоёмкости бытового самообслуживания жильцов здания;
- *охрана и обеспечение безопасности* в случае пожара, землетрясения, бедствия и повреждений конструкции — риск нанесения материального ущерба или вероятность гибели жильцов ИЗ в случае пожара, землетрясения, бедствия и повреждения конструкции;
- *культура*, соответствующая ожиданиям заказчика — соответствие планировочных решений народным бытовым обычаям предполагаемых жильцов;
- *полезные новые технические решения* — количество технических решений, впервые использованных в проекте ИЗ;
- *ход строительства и управления* — сроки и стоимость возведения ИЗ по плану-графику;
- *гигиена и санитария* — трудоёмкость поддержания чистоты и порядка в помещениях и на прилегающем к дому участке.

#### 4 Связь проектных решений с условиями эксплуатации

Проектные решения должны увязывать архитектуру с целевым использованием здания [30-32] и обеспечивать наименьшие общие расходы в течение всего срока его службы. Уменьшение расходов достигается за счёт:

- надлежащей теплоизоляции [33, 34], при которой суммарные затраты на теплоизоляцию и отопление здания являются наименьшими;
- создания запасов, достаточных для жизнеобеспечения между поставками, планируемыми исходя из возможности и стоимости осуществления грузоперевозок;
- первичной переработки бытовых отходов на месте и обеспечения места, достаточного для временного хранения отходов, которые необходимо направлять для переработки на мусороперерабатывающий завод;
- разумного выбора участка для постройки дома, позволяющего упростить конструкцию фундамента, получать больше энергии из имеющихся возобновляемых источников и расходовать меньше энергии на освещение помещений и водоотведение;
- выбора наименьшего количества диагностических параметров, достаточного для мониторинга технического состояния здания и его инженерного оборудования.

При большой удалённости от населённых мест проектные решения должны обеспечивать большую автономность жизнеобеспечения [35].



## 5 Основные показатели эксплуатации

Классификация основных показателей эксплуатации ИЗ, согласно [36], включает:

- *обязательные показатели*, которые необходимы для выполнения основных требований стандартов, строительных правил и определяют наименьшие цели;
- *желательные показатели*, которые определяют образцовые целевые эксплуатационные качества здания, с точки зрения пользователя, более требовательно, чем строительные правила или нормы;
- *предпочтительные показатели*, которые отражают цели и видение, существующие у заказчика;
- *неприменимые показатели*, применение которых не требуется или невозможно.

Основными показателями эксплуатации ИЗ для условий Арктики должны являться:

### 1) *энергетическая эффективность* здания [37].

Показатели энергетической эффективности здания [36]:

*Обязательные показатели:*

- энергия, потребляемая при эксплуатации здания в течение всего ЖЦ;
- использование дневного света для освещения основных помещений;
- обеспечение естественной вентиляции здания с учётом направления ветра.

*Предпочтительные показатели:*

- расположение здания на участке, позволяющее использовать как можно больше солнечного света [25, 36];
- использование возобновляемых источников энергии;
- уменьшение пикового потребления энергии;
- использование солнечного нагрева и естественного охлаждения для обеспечения комфортной температуры при наименьшем задействовании исполнительных механизмов;
- уменьшение теплопередачи через окна при обеспечении необходимого естественного теплообмена.

- 2) *риск*, равный произведению вероятности отказа какой-либо из частей здания на ущерб, который он может причинить.
- 3) *стоимость обслуживания* здания и его частей, которая в течение срока службы определяется транспортной доступностью, приспособляемостью здания к технологическим изменениям, управляемостью инженерного оборудования.
- 4) *стоимость ведения быта*, которая зависит от стоимости водоснабжения и водоотведения, утилизации твёрдых бытовых отходов, обеспеченности связью.

Кроме того, целесообразно учитывать экспертную оценку связи архитектурных решений с существующим культурным наследием.

## 6 Структурный синтез и онтологии проектных решений

Проектирование сложного объекта можно представить как совокупность структурного и параметрического синтеза. Задача структурного синтеза проектных решений, с точки зрения возможности формализации, относится к числу наиболее сложных. Это связано с тем, что с одной стороны, свойства синтезируемого объекта зависят от большого числа зачастую случайных, противоречивых, не до конца исследованных факторов. С другой стороны, при решении задачи синтеза часто приходится выбирать вариант из множества большой мощности. Основная проблема существующих подходов к автоматизации решения данного класса задач - это необходимость в трудоёмкой программной реализации алгоритмов поиска решений и поддержке разработанного программного средства для учёта всех изменений в методике

расчёта. Для решения подобных задач можно использовать метод синтеза конфигурации на основе метода удовлетворения ограничений и онтологического описания предметной области (ПрО), позволяющий инженеру описывать и решать задачи, не прибегая к программированию сложных алгоритмов [38, 39].

В ходе решения задач с использованием онтологий необходимо построить формальную онтологию исследуемой ПрО. При составлении онтологических моделей можно выделить 4 уровня онтологий: метаонтология, онтология верхнего уровня, онтология ПрО и прикладная онтология. Данная модель представляет атрибуты сущностей в виде бинарных отношений, что позволяет создать метаонтологию как совокупность классов объектов и отношений, а онтологию верхнего уровня - в форме классификаторов объектов и их свойств. Этот подход обеспечивает возможность применения универсальных алгоритмов на основе метода программирования в ограничениях для широкого класса задач структурного синтеза и управления конфигурациями. Применение таких моделей в информатике представлено в стандарте ISO/TR 9007:1987 «Системы обработки информации. Понятия и терминология для концептуальной модели базы данных».

Общая черта подходов, использующих бинарные модели — это представление элементов посредством реализации бинарных ассоциаций, т.е. предложений, в которых участвуют только два термина. Основные понятия атомарные, что позволяет явно задавать ограничения. Изменения в ПрО отражаются добавлением или удалением типов бинарных отношений, сущностей или имен сущностей, а также изменением ограничений. Это приводит к большей стабильности по сравнению с подходами, где ограничения жестко связываются с понятиями.

В качестве прототипа разрабатываемой метаонтологии ЖЦ использованы схемы и язык описания RDF и OWL и основные сущности и положения, заложенные в модель данных ISO 15926-2 «Системы промышленной автоматизации и интеграция. Интеграция данных ЖЦ для перерабатывающих предприятий, включая нефтяные и газовые производственные предприятия». Основное требование к разрабатываемой модели – возможность применения универсальных алгоритмов поиска решений методами удовлетворения ограничений и другими алгоритмами ИИ.

Базовые понятия метаонтологии включают сущности: Class, Abstract Object, Feature, Measure Unit, Class of information presentation, Possible individual, Physical object, Activity и типы отношений: subClassOf, Composition of individual, hasMeasureUnit, hasFeature, hasValue.

В качестве типов отношений в метаонтологии описываются те, из которых наследуются все остальные: род – вид; часть-целое; объект-свойство-значение; процесс-ресурс; и ряд других. Для всех отношений в метаонтологии должны быть определены формальные свойства.

В качестве языка описания выражений (функций интерпретаций, аксиом и ограничений) использован подход на основе языка моделирования данных “EXPRESS”, предложенного в стандарте ISO 10303.

Данная метаонтология должна послужить основой для создания онтологии ПрО.

К сущностям ISO 15926-2 и OWL добавлены:

*Feature* – класс для создания объектов-характеристик к индивидам *Possible individual*. С помощью индивидов *Feature* возможно будет осуществлять множественную связь объекта *Possible individual* с численными значениями, текстовыми значениями, другими индивидами, а также присоединять различные дополнительные данные, необходимые для полного понимания присоединяемого к *Possible individual* значения.

*Measure unit* – класс для создания размерных величин к значениям присоединяемым к индивидам *Feature*.

*hasFeature* — отношение, соединяющее *Feature* с *Possible individual*.

Таким образом, онтология ЖЦ ИЗ будет представлять собой классификаторы объектов (элементов) свойств, шкал измерений и процессов и множество конкретных объектов, связанных отношениями и функциями интерпретации, которые наследуются от базовых понятий метаонтологии.

## 7 Онтология ПрО

Анализ ЖЦ ИЗ позволил выявить множество процессов и объектов и их свойств, которые должны составить онтологию ПрО ЖЦ ИЗ.

### 7.1 Объекты

*Объекты ЖЦ ИЗ включают:* ИЗ; населённые пункты; энергоремонтные предприятия; строительные организации; вспомогательные предприятия; персонал; внешнюю среду (грунт; воздушная среда; водная среда; ..) и др.

*Элементы ИЗ:* несущие и ограждающие конструкции (фундамент, перекрытия, крыши, лестницы, стены, окна, двери, теплоизоляция и т.д.); инженерное оборудование (вентиляция, отопление, водоснабжение, водоотведение, электроснабжение, освещение и т.д.); устройства управления инженерным оборудованием здания; ИИУС и др.

Каждая категория может быть декомпозирована до уровня, обеспечивающего классификацию до конкретной номенклатуры элементов. Например, электроэнергетическая система (ЭЭС) состоит из элементов, которые можно разделить на три группы:

*Основные (силовые) элементы:*

- дизель-электрическая установка или топливный элемент;
- инвертор и аккумулятор ветроэнергетической установки;
- трансформаторы, автотрансформаторы, выпрямительные установки, преобразующие значения и вид тока и напряжения;
- линии электропередач, передающие электроэнергию на расстояние;
- коммутирующая аппаратура (выключатели, разъединители), предназначенные для изменения схемы ЭЭС и отключения поврежденных элементов;

*Измерительные элементы:* трансформаторы тока и напряжения, предназначенные для подключения измерительных приборов, средств управления и регулирования;

*Средства управления:* релейная защита, регуляторы, автоматика, телемеханика, связь, обеспечивающие оперативное и автоматическое управление схемой и работой ЭЭС.

### 7.2 Процессы

*ЖЦ ИЗ:* строительство ИЗ; монтаж оборудования ИЗ; ввод в эксплуатацию ИЗ; эксплуатация ИЗ; техническое обслуживание и модернизация ИЗ; утилизация ИЗ.

*Процессы внешней среды:* сейсмическая активность; движение воздушной среды; движение водной среды; прочие.

*Процессы жизнедеятельности персонала:* перемещение по ИЗ; дыхание; теплообмен; прочие.

### 7.3 Свойства элементов и объектов

Каждый уровень иерархии объектов (элементов) и процессов имеет существенно различающиеся особенности и свойства, которые должны учитываться в процессе эксплуатации. Однако свойства одного типа для разных объектов должны относиться к одному классу свойств. В соответствии с этим свойства объектов также имеют иерархическую структуру.

*Технико-экономические:* стоимость; затраты на обслуживание; экономичность; прочие.

*Физические:* масса; габаритные размеры; прочие.

*Эксплуатационные:* значения напряжений, мощностей и токов элементов; частоты, определяющие процесс производства, передачи, распределения и потребления электроэнергии; прочие.

*Показатели надёжности эксплуатации по состоянию:*

- параметр потока отказов — среднее количество отказов ремонтируемого изделия в единицу времени (обычно один год);
- время восстановления после отказа — среднее время вынужденного простоя, необходимого для отыскания и устранения одного отказа;
- коэффициент вынужденного простоя — определяет вероятность нахождения элемента или установки в вынужденном простое;
- коэффициент технического использования - характеризует вероятность нахождения установки в работоспособном состоянии (в работе или в резерве);
- риск.

*Показатели экологичности, эффективности и др.*

## 7.4 Шкалы (единицы измерения)

*Метрические:* денежные; физические; прочие.

*Лингвистические:* состояние поверхности (обледенелая, заснеженная, мокрая, сухая); скорость ветра (очень слабая, слабая, довольно слабая, умеренная, довольно сильная, сильная, очень сильная, исключительно сильная); порывы ветра (очень слабые, слабые, довольно слабые, умеренные, довольно сильные, сильные, очень сильные, исключительно сильные); другие.

На основе предметной онтологии могут быть разработаны онтологии задач, таких как синтез конфигурации ИЗ, планирование технического обслуживания, мониторинга и др.

## Заключение

Проектирование жилых домов для арктических условий как ИЗ позволит обеспечить уютную среду проживания при наименьших затратах в течение всего срока службы таких зданий. Предлагаемая структура онтологии ИЗ является основой структурной модели интеллектуальной среды для обмена информацией между субъектами ПрО и обеспечит поддержку функционирования ИЗ на современном уровне.

## Список источников

- [1] Указ Президента РФ от 02.05.2014 N 296 (ред. от 27.06.2017) «О сухопутных территориях Арктической зоны Российской Федерации». – <http://legalacts.ru/doc/ukaz-prezidenta-rf-ot-02052014-n-296/>.
- [2] СП 131.13330.2012. Строительная климатология. – М., 2012. – 124 с.
- [3] **Грицан, О.Е.** Благоустройство в арктическом климате. – 2018. – <https://goarctic.ru/live/blagoustroystvo-v-arkticheskom-klimate/>.
- [4] **Скупов, Б.** Архитектура высоких широт: от «мерзости запустения» к «застывшей музыке» городов и поселений. – 2015. – <https://ardexpert.ru/article/4496>.
- [5] **Daley, K.** Meeting the northern housing challenge. – 2017. – <https://www.canada.ca/en/polar-knowledge/publications/polarleads/vol1-no1-2016.html>.
- [6] **Варфоломеев, Ю.А.** Особенности проектирования и строительства малоэтажных домов в Арктике // Арктика и Север. – 2014. № 17. – С. 28-43.

- [7] **Аралова, Н.А.** Архитектурная концепция жилого комплекса в арктической зоне (на примере посёлка Зеленоборский Мурманской области) // Проблемы и достижения в науке и технике / Сборник научных трудов по итогам международной научно-практической конференции. – Омск, 2014. – С. 18.
- [8] **Капустин, П.В.** Онтологические вопросы в кастомизированном архитектурном онлайн проектировании персонализированных жилых домов / Капустин П. В., Канин Д. М., Чураков И. Л. // Онтология проектирования. – 2015. – Т. 5, № 3(17). – С. 256-277.
- [9] **Wigginton, M.** Intelligent Skins / M. Wigginton, J. Harris. – Oxford: Architectural Press, 2002. – 186 p.
- [10] **Caffrey, R.** The Intelligent Building-An ASHRAE Opportunity / R. Caffrey // ASHRAE Technical Data Bulletin. – 1985. – Vol. 4, № 1.
- [11] **Nguyen, T.A.** Energy Intelligent Buildings Based on User Activity: A Survey/ Nguyen T. A., Aiello M. // Energy and Buildings. – 2013. – Vol. 56. – P. 244-257.
- [12] СП РК 4.02-16-2005. Проектирование и строительство инженерных систем многоквартирных жилых домов. – Астана, 2005.
- [13] **Wang, S.** Intelligent Building and Building Automation. – New York: Routledge, 2009. – 248 p.
- [14] **Табунищников, Ю.А.** Интеллектуальные здания / Ю.А. Табунищников // АВОК. – 2001. № 3. – С.6-13.
- [15] Строительство мусороперерабатывающего завода с нуля. — <https://biztolk.ru/biznes-idei/proizvodstvo/stroitelstvo-musoropererabatyvayushhego-zavoda.html>.
- [16] Приказ министерства здравоохранения Российской Федерации от 27.02.2016 № 132н «О требованиях к размещению медицинских организаций государственной системы здравоохранения и муниципальной системы здравоохранения исходя из потребностей населения».
- [17] ГОСТ 25100—2011. Грунты. Классификация. – М.: Стандартинформ, 2013. – 44 с.
- [18] СП 25.13330.2012. Основания и фундаменты на вечномёрзлых грунтах. – М.: Стандартинформ, 2012. – 123 с.
- [19] Геология СССР (61 том). – М.: Госгеолтехиздат, 1944-1974.
- [20] СП 255.1325800.2016. Здания и сооружения. Правила эксплуатации. Основные положения. – М.: Стандартинформ, 2016. – 62 с.
- [21] **Асаул, А.Н.** Теория и практика малоэтажного жилищного строительства в России / А.Н. Асаул, Ю.Н. Казаков, Н.И. Пасяда, И.В. Денисова. – СПб.: Издательство «Гуманистика», 2005.
- [22] СП 14.13330.2018. Строительство в сейсмических районах. – М.: Стандартинформ, 2018. – 122 с.
- [23] **Петрова, И.Ю.** Проектирование информационно-измерительных и управляющих систем для интеллектуальных зданий. Направления дальнейшего развития. / Петрова И. Ю., Зарипова В. М., Лежнина Ю. А. // Вестник МГСУ. – 2015. — Т. 10, № 12. – С.147-159.
- [24] **Саревут, Т.О.** Принципы формирования среды обитания в арктическом регионе // Вестник МГСУ. – 2018. – Т. 13, № 2(113). – С.130-140.
- [25] **Витрувий.** Десять книг об архитектуре. – Изд. 2-е, исправл. изд. – М.: Едиториал УРСС, 2003. – 320 с.
- [26] ГОСТ 30494—2011. Здания жилые и общественные. Параметры микроклимата в помещениях. – М.: Стандартинформ, 2011.
- [27] СП 55.13330.2016. Здания жилые многоквартирные. Правила проектирования. – М.: Стандартинформ, 2016. – 40 с.
- [28] **Huang, Z.Y.** Evaluating Intelligent Residential Communities Using Multi-strategic Weighting Method in China // Energy and Buildings. – 2014. – Vol. 69. – P. 144-153.
- [29] **So, A.T.P.** A New Definition of Intelligent Buildings for Asia / So A.T.P., Wong A.C.W., Wong K.C. // The Intelligent Building Index Manual. – Hong Kong: Asian Institute of Intelligent Buildings, 2011.
- [30] **Bhatt, M.** Ontological modelling of form and function for architectural design / M. Bhatt, J. Hois, O. Kutz // Applied Ontology. — 2012. — Vol. 7 – P. 233-267.
- [31] **Dittmar, G.A.** Architecture as Dwelling and Building Design as Ontological Act / G.A. Dittmar. – 1998. – <https://ru.scribd.com/document/45763774/Architecture-as-Ontological-Act>.
- [32] **Hois, J.** Modular Ontologies for Architectural Design / J. Hois, M. Bhatt, O. Kutz // Formal Ontologies Meet Industry / Ferrario R., Oltramari A.— IOS Press, 2009.
- [33] **Петерсон, Ю.Н.** Сибирский сверхтеплоэкономичный умный дом эконом-класса // Сборник докладов III Всероссийской научной конференции с международным участием «Энерго- и ресурсоэффективность малоэтажных жилых зданий». – Новосибирск: Институт теплофизики им. С.С. Кутателадзе Сибирского отделения Российской академии наук, 2017. – С. 105-111.
- [34] **Мельник, П.** Тепловая защита и системы климатизации зданий в условиях арктического региона / П. Мельник, Н. Шилкин // Здания высоких технологий. – 2017. № 1. – С. 54-62.
- [35] **Барцев, С.И.** Автономный и экологически комфортный тип жилья для освоения Арктики / Барцев С. И., Дегерменджи А. Г., Охонин В. А., Тихомиров А. А. // Журнал Сибирского федерального университета. Серия: Биология. – 2018. – Т. 11, № 2. – С. 190-198.



- [36] **AlWaer, H.** Key Performance Indicators (KPIs) and Priority Settings in Using the Mutli-Attribute Approach for Assessing Sustainable Intelligent Buildings / AlWaer H., Clements-Croome D. J. // Building and Environment. – 2010. – Vol. 45, № 4. – P. 799-807.
  - [37] **Baumgärtel K.** An ontology framework for improving building energy performance by utilizing energy saving regulations. / K. Baumgärtel, M. Kadolsky, J.R. Scherer // Conference: 10th European Conference on Product & Process Modelling (ECPPM 2014), At Vienna, Austria. - DOI: 10.1201/b17396-86.
  - [38] **Бочкарев, С.В.** Структурный синтез сложного электротехнического оборудования на основе метода удовлетворения ограничений / С.В. Бочкарев, А.Б. Петрученок, М.В. Овсянников, С.А. Буханов // Электротехника. 2015. № 6. - С.63-69.
  - [39] **Van Beek P., Walsh T.** Principles of Constraint Programming and Constraint Processing: A Review //AI Magazine Volume 25 Number 4 (2004).
- 

## PRE-PROJECT ANALYSIS OF AN INTELLIGENT RESIDENTIAL BUILDING FOR ARCTIC CONDITIONS

**N.M. Borgest**<sup>1a</sup>, **A.S. Galakhar**<sup>2b</sup>, **M.V. Ovsyannikov**<sup>2c</sup>, **R.O. Samsonov**<sup>1d</sup>

<sup>1</sup> Samara National Research University named after academician S.P. Korolev, Samara, Russia

<sup>a</sup> [borgest@yandex.ru](mailto:borgest@yandex.ru), <sup>d</sup> [samsonov@ssau.ru](mailto:samsonov@ssau.ru)

<sup>2</sup> Bauman Moscow State Technical University, Moscow, Russia

<sup>b</sup> [a.galakhar@bmstu.ru](mailto:a.galakhar@bmstu.ru), <sup>c</sup> [mvo50@mail.ru](mailto:mvo50@mail.ru)

### Abstract

The combination of adverse natural conditions with low density of population in arctic areas complicates approaches to design and construction of housing accommodations for arctic locality and governs the requirements for interior arrangement, automated sustenance systems and reliability of residential buildings. The article addresses the problem of structural synthesis of an intelligent residential building for arctic conditions. A classification of construction constraints affecting the decisions made in design process is introduced. A task of satisfying these constraints in the process of structural design and building automation is introduced. The designated goal is to obtain an admissible solution of the synthesis problem on the basis of formalized knowledge of operating conditions, that minimizes the objective function's value. The aspects of selecting a target function are considered. An approach towards developing new compelling projects of intelligent residential buildings for the conditions of the Russian High North, capable of providing comfort and safety using the latest achievements of engineering in the field of automation and reliability of automated systems, is proposed. Compelling and comfortable housing will help the issue of convincing people to settle in the Arctic regions, research, develop and preserve them for posterity.

**Keywords:** Arctic, pre-project analysis, intelligent building, ontology, structural synthesis, satisfaction of constraints.

**Citation:** Borgest NM, Galakhar AS, Ovsyannikov MV, Samsonov RO. Pre-project analysis of an intelligent residential building for Arctic conditions [In Russian]. *Ontology of designing*. 2019; 9(1): 85-100. – DOI: 10.18287/2223-9537-2019-9-1-85-100.

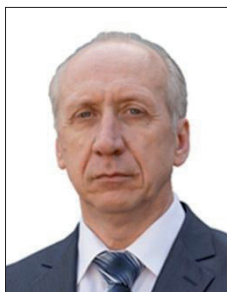
### References

- [1] Decree of the President of Russian Federation of 02.05.2014 No. 296 (ed. of 27.06.2017) "About land territory of the Arctic Zone of Russian federation" [In Russian]. – <http://legalacts.ru/doc/ukaz-prezidenta-rf-ot-02052014-n-296/>.
- [2] SP (Code of Practice) No. 131.13330.2012. Building climatology [In Russian]. - 2012.
- [3] **Gritsan OE.** Accomplishment in arctic climate [In Russian]. — 2018. — <https://goarctic.ru/live/blagoustroystvo-v-arkticheskom-klimате/>.
- [4] **Skupov B.** Architecture of high latitudes: from "abomination of desolation" to "poured music" of towns and settlements [In Russian]. — 2015. — <https://ardexpert.ru/article/4496>.
- [5] **Daley K.** Meeting the northern housing challenge. — 2017. — <https://www.canada.ca/en/polar-knowledge/publications/polarleads/vol1-no1-2016.html>.

- [6] **Varfolomeev YuA.** Specifics of design and construction of low-rise buildings in the Arctic [In Russian]. The Arctic and the North. — 2014; 17: 28-43.
- [7] **Aralova NA.** Architectural concept of housing estate in the Arctic Zone (by the example of the township of Zelenoborskyi in Murmansk Oblast) [In Russian]. In Proceedings of International Scientific and Practical Conference “Problems and advances in science and technology”. — Omsk, 2014. — P.18.
- [8] **Kapustin PV, Kanin DM, Churakov IL.** The Ontological questions of personalize homes customized architectural online designing [In Russian]. *Ontology of Designing*. — 2015; 5(3): 256-277. - DOI:10.18287/2223-9537-2015-5-3-256-277.
- [9] **Wigginton M, Harris J.** Intelligent Skins. — Oxford: Architectural Press, 2002. — 186 p.
- [10] **Caffrey R.** The Intelligent Building-An ASHRAE Opportunity // ASHRAE Technical Data Bulletin. — 1985. — Vol. 4, № 1.
- [11] **Nguyen TA.** Energy Intelligent Buildings Based on User Activity: A Survey/ Nguyen T. A., Aiello M. // *Energy and Buildings*. — 2013; 56: 244-257.
- [12] SP RK (Code of Practice of Republic of Kazakhstan) No. 4.02-16-2005. Design and construction of plumbing and heating systems of one-family dwellings [In Russian], 2005.
- [13] **Wang S.** Intelligent Building and Building Automation. — New York: Routledge, 2009. — 248 p.
- [14] **Tabunshchikov YuA.** Intelligent buildings [In Russian], AVOK. — 2001; 3: 6-13.
- [15] Construction of a new recycling plant [In Russian]. — <https://biztolk.ru/biznes-idei/proizvodstvo/stroitelstvo-mosoroperabatvayushhego-zavoda.html>.
- [16] Order of the Ministry of Public Health of Russian Federation of 27.02.2016 No132n “About requirements for siting of medical organizations of state health system and municipal health system based on needs of people” [In Russian].
- [17] GOST (State Standard) No. 25100—2011, Soils. Classification [In Russian]. 2013. - 44 p.
- [18] SP (Code of Practice) No. 25.13330.2012. Soil bases and foundations on permafrost soils [In Russian]. 2012. — 123 p.
- [19] *Geology of the USSR* [In Russian]. (61 Vol.) — Moscow: Gosgeoltechizdat, 1944-1974.
- [20] SP (Code of Practice) No. 255.1325800.2016. Buildings and structures. Operating rules. General provisions [In Russian]. - М.: Standardinform, 2016. - 62 p.
- [21] **Asaul AN, Kazakov YuN, Pasyada NI, Denisova IV.** The theory and the practice of low-rise house-building in Russia [In Russian]. — St. Petersburg: Publishing house “Gumanistika”, 2005.
- [22] SP (Code of Practice) No. 14.13330.2018. Seismic building design code [In Russian]. - Moscow: Standardinform, 2018. - 122 p.
- [23] **Petrova IYu, Zaripova VM, Lezhnina YuA.** Design of information-measuring and control systems for intelligent buildings. Trends of development [In Russian]. *Vestnik MGSU*. — 2015; 10(12): 147-159.
- [24] **Sarvut TO.** The principles of building habitat in the Arctic Region [In Russian]. *Vestnik MGSU*. — 2018; 13(2): 130-140.
- [25] **Vitruvius.** Desyat' knig ob architecture [Ten Books on Architecture] [In Russian]. Trans. Latin F.A. Petrovsky. Ed. 2<sup>nd</sup> corrected. — Moscow: Editorial URSS, 2003. — 320 p.
- [26] GOST (State Standard) No. 30494—2011, Residential and public buildings. Parameters of indoor microclimate [In Russian]. - Moscow: Standardinform, 2011.
- [27] SP (Code of Practice) No. 55.13330.2016. Preschool educational institution buildings. Design rules [In Russian]. - Moscow: Standardinform, 2016. - 40 p.
- [28] **Huang ZY.** Evaluating Intelligent Residential Communities Using Multi-strategic Weighting Method in China // *Energy and Buildings*. — 2014. — Vol. 69. — P. 144-153.
- [29] **So ATP, Wong ACW, Wong KC.** A New Definition of Intelligent Buildings for Asia // *The Intelligent Building Index Manual*. — Hong Kong: Asian Institute of Intelligent Buildings, 2011.
- [30] **Bhatt M, Hois J, Kutz O.** Ontological modelling of form and function for architectural design. *Applied Ontology*. — 2012; 7: 233-267.
- [31] **Dittmar GA.** Architecture as Dwelling and Building Design as Ontological Act. — 1998. — <https://ru.scribd.com/document/45763774/Architecture-as-Ontological-Act>.
- [32] **Hois J, Bhatt M, Kutz O.** Modular Ontologies for Architectural Design // *Formal Ontologies Meet Industry* / Ferrario R., Oltramari A.— IOS Press, 2009.
- [33] **Peterson YuN.** Siberian thermally super efficient intelligent building of economy class [In Russian]. In Proceedings of the Third All-Russian Scientific Conference with International Participation “Energy and Resource Efficiency of Low-rise Residential Buildings”. — Novosibirsk: The Institute of Thermal Physics named after S.S. Kutateladze of the Siberian Division of Russian Academy of Sciences, 2017. — P. 105-111.
- [34] **Melnik P, Shilkin N.** Thermal protection and climate systems of buildings in the conditions of the Arctic [In Russian]. *Sustainable Building Technologies*. — 2017; 1: 54-62.

- [35] **Bartsev SI, Degermendji AG, Okhonin VA, Tikhomirov AA.** Autonomous and ecologically comfortable type of dwelling for development of the Arctic [In Russian]. Journal of Siberian Federal University. Series: Biology. — 2018; 11(2): 190-198.
- [36] **AlWaer H, Clements-Croome DJ.** Key Performance Indicators (KPIs) and Priority Settings in Using the Mutli-Attribute Approach for Assessing Sustainable Intelligent Buildings // Building and Environment. – 2010; 45(4): 799-807.
- [37] **Baumgärtel K, Kadolsky M, Scherer JR.** An ontology framework for improving building energy performance by utilizing energy saving regulations // Conference: 10th European Conference on Product & Process Modelling (ECPPM 2014), At Vienna, Austria. - DOI: 10.1201/b17396-86.
- [38] **Bochkarev SV, Petrochenkov AB, Ovsyannikov MV, Bukhanov SA.** Structural synthesis of complex electrical equipment based on the method of satisfying restrictions [In Russian]. Electrical engineering. 2015; 6: 63-69.
- [39] **Van Beek P., Walsh T.** Principles of Constraint Programming and Constraint Processing: A Review //AI Magazine Volume 25 Number 4 (2004).

### Сведения об авторах



**Боргест Николай Михайлович**, 1954 г. рождения. Окончил Куйбышевский авиационный институт имени академика С.П. Королёва в 1978 г., к.т.н. (1985). Профессор кафедры конструкции и проектирования летательных аппаратов Самарского национального исследовательского университета имени академика С.П. Королёва, директор издательства «Новая техника», с.н.с. ИПУСС РАН. Член Международной ассоциации по онтологиям и их приложениям (IAOA), Российской ассоциации искусственного интеллекта (РАИИ). В списке научных трудов более 200 работ в области автоматизации проектирования и искусственного интеллекта.

**Nikolay Mikhailovich Borgest** (b.1954) graduated from the Kuibyshev Aviation Institute named after academician S.P. Korolev (Kuibyshev-city) in 1978, PhD (1985). He is a Professor at Samara National Research University named after academician S.P. Korolev, Director of «New Engineering» publishing house, Senior Research worker at ICCS RAS. He is a member of the International Association for Ontology and its Applications, a member of the Russian Association of Artificial Intelligence (RAII), co-author of more 200 scientific articles and abstracts in the field of CAD and AI.

**Галахарь Александр Сергеевич**, 1971 г. рождения. Окончил Московский государственный технический университет им. Н.Э. Баумана в 1995 г., к.т.н. (2012). Доцент кафедры «Компьютерные системы автоматизации производства» Московского государственного технического университета им. Н.Э. Баумана.

**Alexander Sergeevich Galakhar** (b.1971) graduated from Bauman Moscow State Technical University in 1995, PhD (2012). He is an Associate Professor in the School of Computer-driven Systems of Industrial Automation at Bauman Moscow State Technical University.



**Самсонов Роман Олегович**, 1961 г. рождения. Окончил Грозненский нефтяной институт им. акад. Миллионщикова (1983), к.т.н. (1999), д.т.н. (2008, специальность 05.13.01). Первый проректор Самарского национального исследовательского университета имени академика С.П. Королёва. В списке научных трудов более 100 работ, 10 книг и монографий, 17 патентов.

**Roman Olegovich Samsonov** (b.1961) graduated from the Grozny Oil Institute named after acad. Millionshchikov in 1983, Ph.D. (1999), Doctor of Technical Sciences (2008). He is the First Vice-Rector of the Samara National Research University named after Academician S.P. Korolev. He is a co-author of more than 100 papers, 10 books and monographs, 17 patents.

**Овсянников Михаил Владимирович**, 1948 г. рождения. Окончил Московский государственный технический университет им. Н.Э. Баумана в 1976 г., к.т.н. (1988). Доцент кафедры «Компьютерные системы автоматизации производства» Московского государственного технического университета им. Н.Э. Баумана.

**Mikhail Vladimirovich Ovsyannikov** (b.1948) graduated from Bauman Moscow State Technical University in 1976, PhD (1988). He is an Associate Professor in the School of Computer-driven Systems of Industrial Automation at Bauman Moscow State Technical University.



УДК 004.896, 004.414.23, 004.4'242

## МОДЕЛЬ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ПОСТРОЕНИЮ СВЯЗАННЫХ ДАННЫХ НА ОСНОВЕ РАЗНОРОДНЫХ ИСТОЧНИКОВ

**С.В. Лебедев**

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»  
им. В.И. Ульянова (Ленина) и ООО «Датафабрик», Санкт-Петербург, Россия  
lebedev.sv.etu@gmail.com

### Аннотация

Сегодня доступно большое число различных источников данных. Многие источники представляют значительную ценность для принятия обоснованных решений в различных предметных областях. Простота и эффективность использования данных зависят от возможности их интеграции в единую модель, последовательно и всесторонне описывающую интересующий предмет. Получению такой модели препятствует то, что элементы разных источников не связаны друг с другом, а сами данные представлены в различных форматах. Несмотря на то, что консорциум W3C предложил языки единообразного и связанного описания данных, до сих пор существует значительное число издателей, публикующих несвязанные данные. В настоящей работе предлагается подход, на основе которого может быть создан эффективный инструмент для обработки больших объемов несвязанных данных. Новым является автоматизация построения онтологии процесса связывания данных на основе отображения множества элементов онтологии, описывающей источник данных, во множество элементов онтологии предметной области. Полученная онтология процесса связывания может служить для специализаций этого процесса пользователем, для дальнейшей его оптимизации и реализации на различных вычислительных платформах. В частности, в работе продемонстрирована возможность генерации программного кода процесса связывания, исполняемого на высокопроизводительной масштабируемой платформе. Полученные результаты позволяют говорить о целесообразности развития предложенного подхода.

**Ключевые слова:** связанные данные, модель процесса связывания, автоматизация построения модели, генерация программного кода, масштабируемые решения.

**Цитирование:** Лебедев, С.В. Модель-ориентированный подход к построению связанных данных на основе разнородных источников / С.В. Лебедев // Онтология проектирования. – 2019. – Т.9, №1(31). – С.101-116. – DOI: 10.18287/2223-9537-2019-9-1-101-116.

### Введение

Консорциум W3C (*World Wide Web Consortium*) разрабатывает и внедряет стандарты Всемирной Паутины (*World Wide Web*). Деятельность консорциума включает технологическое поддержание процессов построения глобальной сети данных. Некоторое время эта деятельность развивалась в рамках построения так называемой Семантической Паутины (*Semantic Web*) [1], а позднее была поглощена деятельностью (W3C Data Activity) по созданию Паутины Данных (*Web of Data*) [2]. Одна из основных задач этой деятельности заключается в том, чтобы перейти от паутины документов к паутине данных и сделать данные доступными для машинной обработки.

Для решения этой задачи несколькими рабочими группами были созданы стандарты языков представления данных (RDF/RDFs, OWL) и написания запросов к данным (SPARQL). Одной из отличительных черт этих языков является ориентация на использование междуна-



родных идентификаторов ресурсов (IRI, *International Resource Identifiers*) [3], позволяющих однозначно идентифицировать элемент данных на множестве наборов данных. Перечисленные стандарты составляют технический аспект решения.

Организационный аспект представлен требованиями к публикуемым данным. Первая группа требований определяет так называемые Связанные Данные (*Linked Data*) [4]. Требования сформулированы в виде правил, включающих требование использовать языки RDF и SPARQL. Вторая группа требований сформулирована в виде рейтинга открытости связанных данных. Наивысший рейтинг в пять звёзд присваивается данным, которые распространяются по открытой лицензии, используют стандарты W3C (RDF и SPARQL), имеют связи с данными, предоставляемыми третьими лицами [4]. Чем большему числу требований соответствуют данные, тем они более доступны для машинной обработки.

Общее количество наборов *Открытых Связанных Данных* растёт [5], но не охватывает некоторые важные источники. Одним из значимых поставщиков наборов данных является правительство: данные поставляются различными государственными органами. Если обратиться к сайту Реестра наборов открытых данных, то окажется, что из 21468 наборов только 5 представлено в формате RDF [6]. Большинство же данных представлено в форматах CSV, XML и JSON. Основным назначением форматов CSV, XML и JSON является представление данных в виде, не зависящем от приложения, в котором данные были созданы или использованы. Основное их назначение – это *сериализация* данных для хранения или передачи (RDF может быть также записан в XML) [7]. Предназначение RDF/OWL – это формальное и целостное описание предметной области (ПрО). Отсюда, RDF/OWL можно рассматривать как средство представления единой модели данных разделяемого экспертами описания ПрО.

Таким образом, требуется решить задачу построения связанных данных на основе данных, записанных в форматах сериализации (CSV, XML и JSON).

## 1 Задача построения связанных данных

Преимущество использования связанных данных (помимо машинной обработки) состоит в том, что они содержат больше сведений о рассматриваемом предмете, и, следовательно, позволяют принять относительно него более обоснованное решение. Использование единых семантических форматов (RDF/RDFs, OWL, SPARQL) позволяет унифицировать и тем самым упростить доступ к данным. Но большинство поставщиков публикует несвязанные данные, и задача их связывания ложится на плечи потребителя.

Концептуально связывание данных предполагает переход от множества наборов данных, описанных с помощью разрозненных схем, к набору, описанному с помощью одной или множества связанных схем. В рамках рассматриваемых технологий целевая схема представлена в форме онтологии ПрО. Связывание включает решение минимум трёх подзадач:

- 1) построение экземпляров классов ПрО;
- 2) инициализацию примитивных свойств экземпляров;
- 3) связывание экземпляров через объектные свойства.

Примером связывания может служить построение расширенного реестра коммерческих компаний на основе разнородных источников. В качестве источников могут выступать:

- сведения из Единого государственного реестра юридических лиц (ЕГРЮЛ) [8] (основной источник; полный набор является платным, представлен в формате XML);
- открытые данные Федеральной налоговой службы (ФНС) [9], например, Реестр дисквалифицированных лиц (формат CSV), Единый реестр субъектов малого и среднего предпринимательства (формат XML), Сведения о среднесписочной численности работников организации (формат XML) и др.;



- сведения по заключенным контрактам (формат JSON) [10] и т.п.

Реестр, созданный на основе множества перечисленных источников, позволит получать детальный «портрет» каждой компании и принимать более обоснованные (чем в случае использования только ЕГРЮЛ) решения, например, в рамках мероприятия по проверке контрагента и соблюдения должной осмотрительности.

Схематично система связывания данных представлена на рисунке 1. Модель ПрО готовится экспертом. Содержание и структура модели определяются предметной логикой.



Рисунок 1 – Схема связывания данных из разнородных источников

На основе анализа работ [11-13], описывающих требования к инструментам построения связанных данных, и представленного варианта связывания данных сформулированы функциональные и нефункциональные требования к системе.

Основными функциональными требованиями, которые можно рассматривать как обязательный минимум для средства связывания данных, являются:

- возможность связывания разнородных документов: CSV (строки таблицы), XML (иерархия элементов, атрибуты элементов), JSON (массивы, структуры, пары имя-значение), онтологии в формате RDF/OWL (тройки субъект-предикат-объект);
- возможность использования нескольких источников для построения единой модели ПрО;
- поддержка функций, аргументами которых являются значения элементов источника данных, а значениями функций – элементы RDF тройки (субъект, объект, предикат).

К нефункциональным требованиям относятся:

- 1) простота подключения нового источника данных (подключение источника предполагает связывание модели источника данных и модели ПрО);
- 2) простота изменения модели ПрО;
- 3) возможность обработки больших объемов данных за ограниченное время;
- 4) возможность встраивания в имеющуюся вычислительную инфраструктуру.

Первое нефункциональное требование актуально в случае периодического появления новых источников и пересмотра схем данных уже подключенных источников.

Появление нового источника потребует расширения текущей модели ПрО. Требование актуально и в случае периодического пересмотра экспертного взгляда на ПрО. Инструмент изменения модели ПрО должен предоставлять абстракции уровня эксперта ПрО.

Размер данных и ограничения на время их обработки определяются ПрО. Следовательно, наиболее универсальным будет масштабируемое решение.

Четвёртое требование обусловлено тем, что задача связывания не является конечной: на этом этапе готовятся данные для последующего анализа и принятия решений.

Первые два нефункциональных требования определяют аспект взаимодействия пользователя с программным средством, третье и четвёртое требование – аспекты реализации.

Сформулированные требования позволяют сравнить известные инструменты связывания данных.

## 2 Инструменты связывания данных

Множество доступных инструментов можно условно разделить на две группы.

Первую составляют конвейеры обработки данных, реализующие принципы потокового программирования, изложенные, например, в [14]. Конвейер обработки данных представляет собой граф узлов-обработчиков, соединённых каналами передачи данных. Отдельно выделяют инструменты загрузки-преобразования-выгрузки данных (ETL, *Extract-Transform-Load*) [15]. Среди ETL-инструментов, поддерживающих работу с семантическими стандартами, можно выделить *UnifiedViews*<sup>1</sup> [16], *LinkedPipes*<sup>2</sup> [17], *Karma*<sup>3</sup> [18]. К преимуществам таких платформ можно отнести наличие расширенного функционала:

- извлечения (адаптеры к различным источникам данных, например, к базам данных, брокерам сообщений, веб-сервисам и т.п.);
- преобразования (фильтрация, переход от одной схемы к другой и т.п.);
- выгрузки данных (схожие с адаптерами для извлечения, но работающие на запись);
- наличие визуальных средств описания конвейера;
- предоставление (некоторыми платформами) возможности масштабирования.

К недостаткам подобных платформ можно отнести:

- невозможность глубокой интеграции в имеющийся вычислительный процесс;
- ограничения по масштабированию;
- отсутствие возможности интегрировать данные из нескольких источников в рамках одной операции (для преодоления этой трудности предлагается сначала обрабатывать каждый источник в отдельности, а затем связывать полученные результаты).

Вторую группу составляют узкоспециализированные программные решения, оформленные в виде Java-библиотек. К наиболее развитым инструментам относятся *RML-Mapper*<sup>4</sup> [12, 19] и *SPARQL-Generate*<sup>5</sup> [13, 20]. К достоинствам данных решений можно отнести более широкие возможности для связывания, включая возможность создания элементов узлов ETL платформ на их основе. Основным недостатком названных инструментов – это отсутствие возможностей для масштабирования.

## 3 Прототип платформы связывания данных на основе моделей

Исходным пунктом связывания данных является построение отображения элементов источника данных во множество элементов ПрО. В первой группе инструментов для этого используется графический интерфейс, во второй – специализированный язык. На основе этого отображения реализуется вычислительный процесс. В [12] выделены следующие варианты реализации процесса:

- процесс, управляемый отображениями (*mapping-driven*);
- процесс, управляемый данными (*data-driven*);
- смешанное управление.

Реализация, управляемая отображениями, наиболее прямолинейна, но может потребовать многократное прохождение элементов источника данных. Это объясняется тем, что отображения обрабатываются независимо друг от друга. Так, например, отображение одного и того же элемента данных на два разных элемента ПрО рассматриваются как независимые и обрабатываются за два прохода по одним и тем же элементам данных.

<sup>1</sup> <https://github.com/UnifiedViews>.

<sup>2</sup> <https://github.com/linkpipes>.

<sup>3</sup> <http://usc-isi-i2.github.io/karma/>.

<sup>4</sup> <https://github.com/RMLio/RML-Mapper>.

<sup>5</sup> <https://github.com/sparql-generate/sparql-generate>.

Реализация, управляемая данными, учитывает связи между отображениями и сокращает число проходов до минимума. Для примера из предыдущего абзаца потребуется один проход по элементам данных, так как учитывается, что разные элементы ПрО строятся на основе одного и того же элемента источника данных. Но такая реализация более сложна, так как требует анализа зависимостей между отображениями и соответствующей оптимизации процесса.

### 3.1 Подход к связыванию данных на основе моделей

В работе предлагается подход, который можно отнести к смешанным реализациям:

- процесс, управляемый отображениями, строит *модель* процесса связывания данных;
- полученная модель описывает процесс связывания данных, управляемый данными.

Полученная модель может быть использована как для непосредственной реализации, так и для дальнейшей оптимизации и/или настройки пользователем.

Общая схема описанного подхода представлена на рисунке 2.

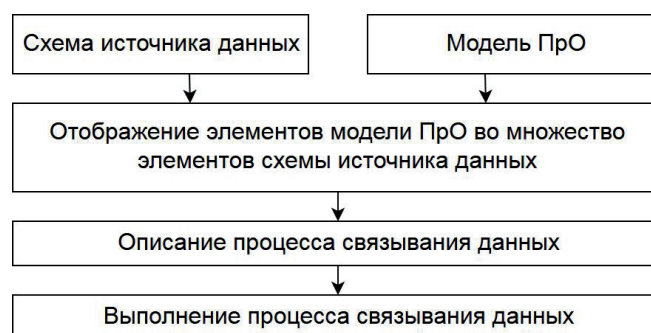


Рисунок 2 – Схема предложенного подхода к связыванию данных

### 3.2 Апробация подхода

Апробация подхода (*proof-of-concept*) имеет ограниченный функционал:

- работает с данными, которые хранятся в виде файлов в формате JSON;
- нет возможности подключения нескольких источников данных;
- нет возможности применения функций.

В качестве примера рассмотрено связывание данных о государственных контрактах (Федеральный закон о закупках 44-ФЗ). Данные по контрактам могут быть получены по адресу <https://clearspending.ru/opensdata/> в разделе *Contracts/44 Federal Law*. Описание контракта включает данные по самому контракту, по заказчику (*customer*), поставщикам (*suppliers*) и поставляемым продуктам (*products*). Для описания контракта используются регистрационный номер (*regNum*) и дата публикации (*publishDate*). Для описания заказчика – имя (*fullName*), ИНН (*inn*), адрес (*postalAddress*). Для описания поставщика – имя (*organizationName*), ИНН (*inn*), адрес (*postAddress*). Для описания продукта – имя (*name*) и цена (*price*).

Реализация предложенного подхода связывания данных представлена на рисунке 3 (детализирует схему на рисунке 2).

#### 3.2.1 Онтология абстрактной схемы данных

Онтология абстрактной схемы включает элементы, общие для любой схемы данных:

$AbstractScheme = \langle Atom, Collection, Container \rangle$ ,

где:

- *Atom* используется для представления атомарного элемента данных (пара имя-значение);
- *Collection* – для представления коллекции (массива) однотипных элементов данных;
- *Container* – для представления множества разнотипных элементов.

Элементы онтологии позволяют обращаться к различным элементам исходных данных с применением подходящего метода. Например, для коллекции таким методом будет перебор элементов.

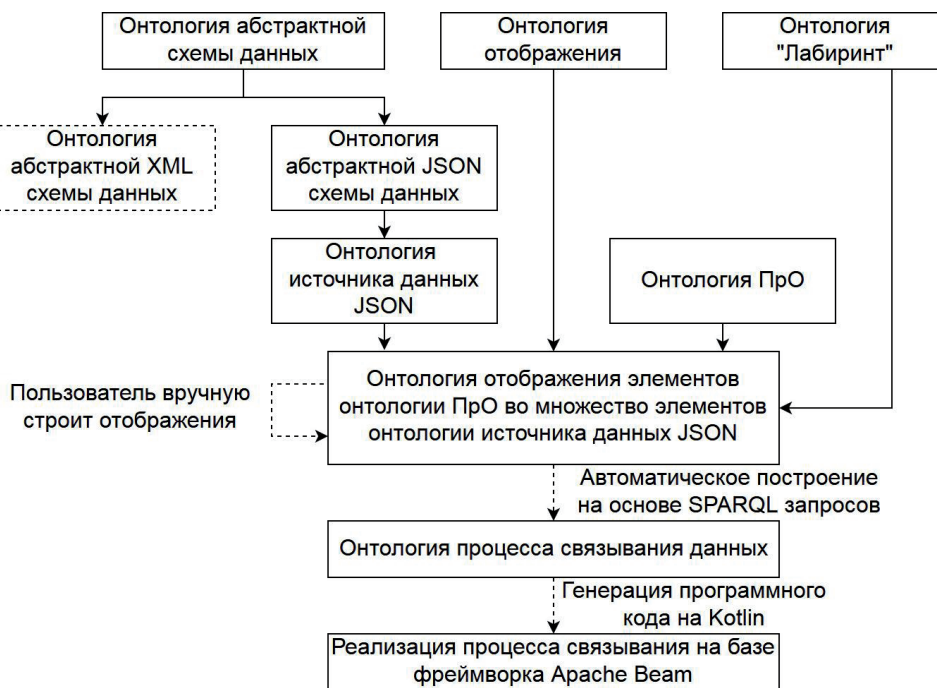


Рисунок 3 – Схема реализации предложенного подхода к связыванию данных

### 3.2.2 Онтология схемы данных заданного формата

Онтология абстрактной JSON-схемы наследует элементы онтологии абстрактной схемы и фиксирует элементы, характерные именно для данного формата:

*JsonAbstractScheme* = *<JSONDataElement, JSONArrayElement, JSONRecordElement>*.

Элементы данной онтологии непосредственно используются для описания источника данных. Указание на используемый формат данных (т.е. JSON) позволяет на этапе реализации процесса отображения выбрать инструменты для работы именно с данным форматом (например, для чтения данных из файла).

### 3.2.3 Онтология источника данных

Онтология источника данных описывает его структуру. На рисунке 4 представлен фрагмент онтологии источника данных для рассматриваемого примера. Все дальнейшие визуализации фрагментов онтологии выполнены с использованием инструмента *Ontodia* (<http://ontodia.org/>).

Онтология источника данных строится по следующему принципу. Имена элементов источника данных используется в качестве имён соответствующих классов. Например, класс *product* представляет одноимённый элемент JSON-документа. Те же имена используются в качестве имён свойств, которые связывают родительские элементы с дочерними элементами. Например, класс *product* связан с классом *name* свойством *name*, что отражает вложенность JSON-элемента *name* в JSON-элемент *product*.

Описание источника данных играет вспомогательную роль и является вторичным по отношению к самим данным. Поэтому к способу записи такого описания можно предъявлять требования, руководствуясь единственно удобством использования описания для построения модели процесса связывания данных.

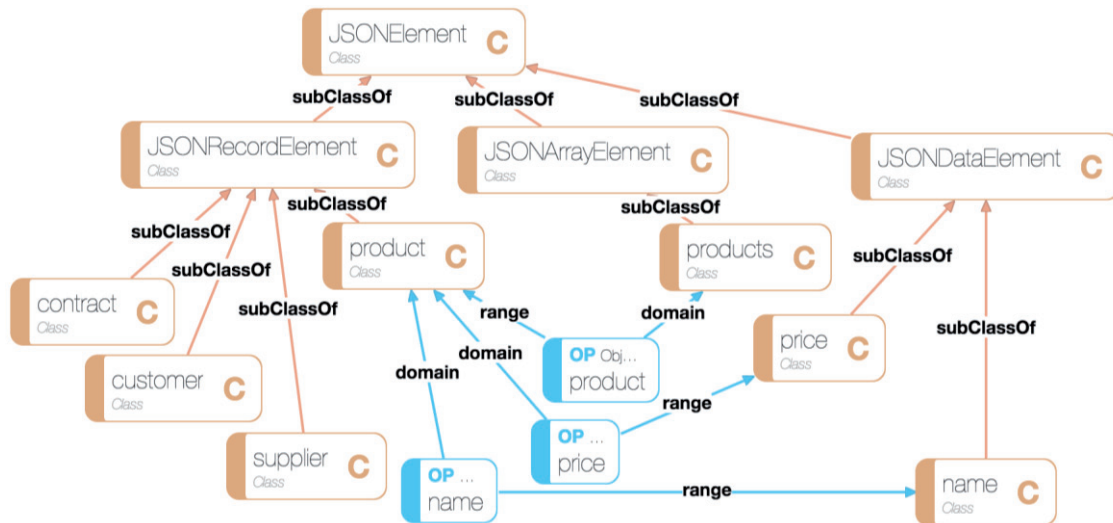


Рисунок 4 – Фрагмент онтологии источника данных

### 3.2.4 Онтология ПрО

Онтология ПрО фиксирует знание экспертов о некотором предмете. Онтология должна быть составлена, исходя из логики рассматриваемой ПрО. В рамках текущей реализации подхода единственным предъявляемым требованием к онтологии ПрО является указание области определения и области значения для всех свойств, которые должны быть включены в результаты отображения. Очевидно, что выполнение этого требования не предполагает понимания специфики процесса связывания данных. На рисунке 5 представлен фрагмент онтологии ПрО.

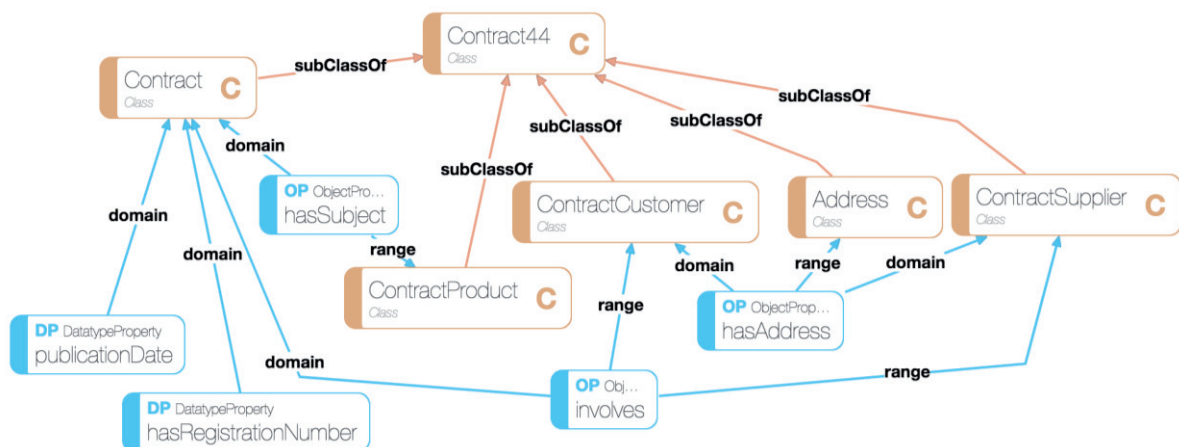


Рисунок 5 – Фрагмент онтологии ПрО

### 3.2.5 Онтология отображения

Онтология отображения включает только два элемента: *DataModel* используется в качестве родительского класса для всех классов источника данных, а *KnowledgeModel* использу-



ется в качестве родительского класса для всех классов ПрО. Наследование от перечисленных классов позволяет упростить разграничение элементов источника данных и ПрО на этапе работы с онтологией отображения.

### 3.2.6 Отображение элементов ПрО во множество элементов источника данных

Отображение элементов онтологии ПрО во множество элементов онтологии источника данных позволяет зафиксировать, какие элементы данных используются для порождения элементов онтологии ПрО. Отображения строятся пользователем. Результатом является онтология. Использование онтологии для фиксации конкретного отображения позволяет изолировать устанавливаемые отображения как от описания источника данных, так и от описания ПрО, и, соответственно, не ограничивает их повторное использование.

В текущей версии для установления отображения используются свойства *equivalentProperty* и *equivalentClass*, которые являются частью спецификации языка OWL.

На рисунках 6 и 7 показаны примеры таких отображений. При соблюдении требования указания областей определения и значения свойств отображение на рисунке 6 позволяет:

- 1) однозначно определить, на основе какого элемента данных должен строиться экземпляр класса *ContractProduct*;
- 2) какой элемент данных использовать для заполнения свойства *hasCost*; для этого сопоставляются области определений и значений эквивалентных свойств.

Свойство *equivalentClass* позволяет разрешить неоднозначность, возникающую в том случае, когда область определения свойства ПрО включает множество классов. В таком случае явно указывается, какой класс онтологии источника данных соответствует классу онтологии ПрО.

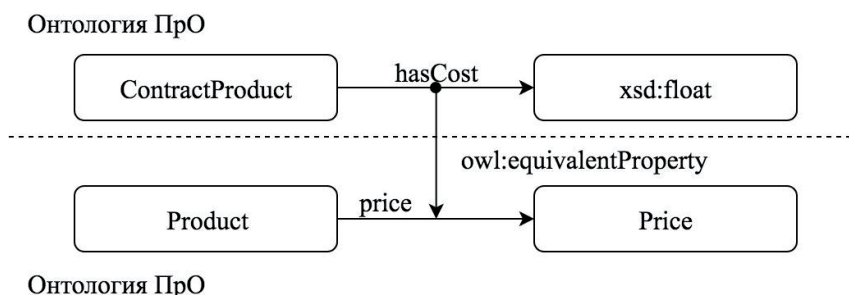


Рисунок 6 – Отображение на основе *equivalentProperty*

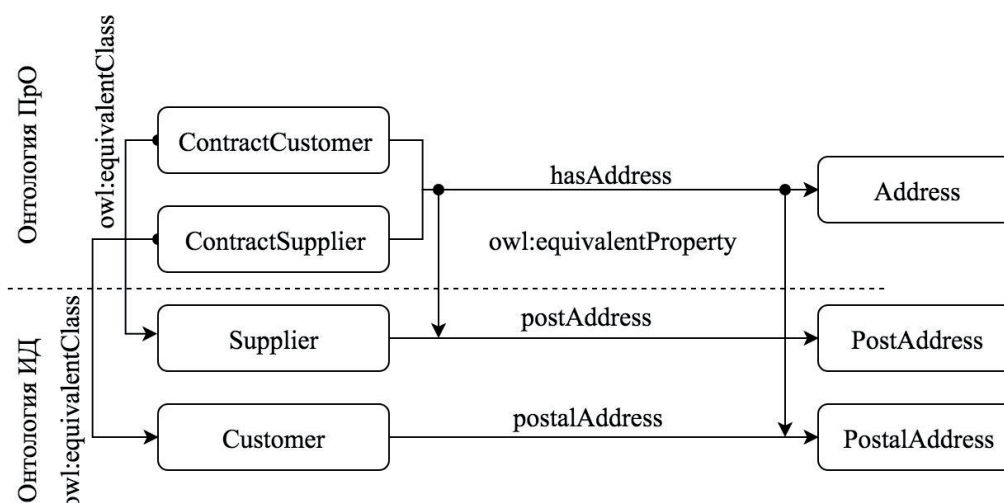


Рисунок 7 – Отображение на основе *equivalentProperty* и *equivalentClass*

### 3.2.7 Онтология «Лабиринт». Этапы построения модели процесса связывания

Модель процесса связывания данных строится на основе онтологии отображения элементов ПрО во множество элементов источника данных. Для реализации построения модели процесса связывания создана онтология «Лабиринт». Данная онтология включает как элементы, из которых строится модель процесса связывания, так и правила построения этой модели. Модель процесса связывания основана на модели потока данных (см. рисунок 8).

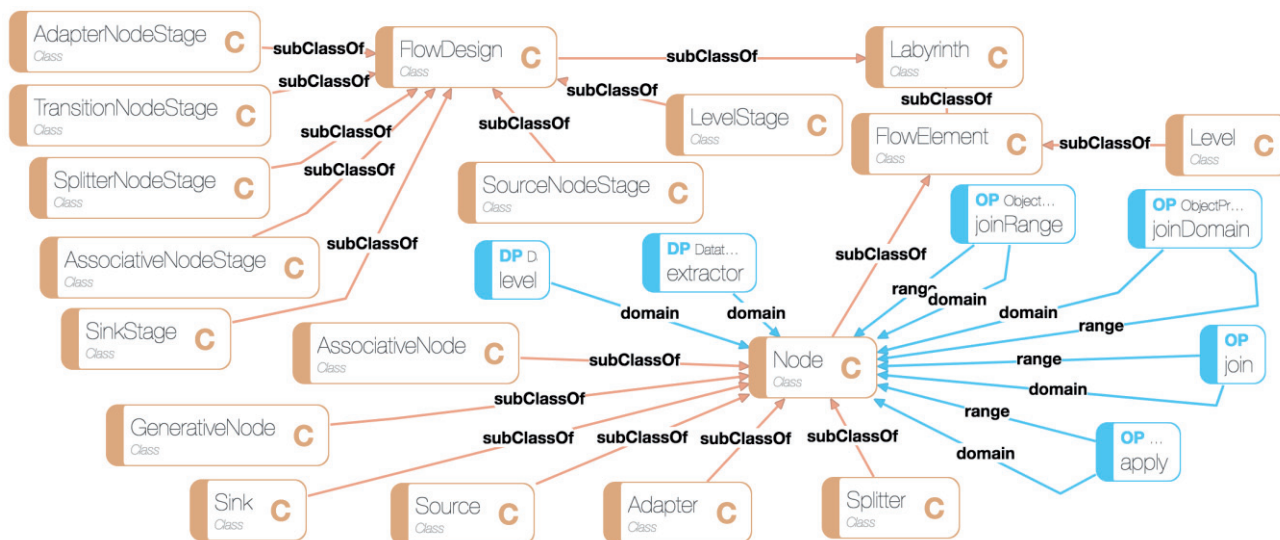


Рисунок 8 – Онтология «Лабиринт»: модель потока данных

Поток данных представляет собой граф, состоящий из узлов вычисления, которые сгруппированы по уровням. Каждый узел получает данные от узлов, лежащих на нижних уровнях. Узлы первого уровня получают данные из указанного источника данных. Онтология условно разделена на два подграфа: *FlowDesign* включает множество подклассов, представляющих этапы построения модели процесса связывания данных, а *FlowElement* включает подклассы, из экземпляров которых состоит модель процесса связывания. Основным классом модели процесса связывания данных является класс *Node* – узел потока данных, реализующий некоторую функцию. Узлы разделены на следующие типы:

- *Adapter*, *Splitter* – обеспечивают доступ к элементам источника данных: *Adapter* – это узел, который обеспечивает доступ к отдельным элементам источника данных, *Splitter* – к элементам массива;
- *GenerativeNode* и *AssociativeNode* – обеспечивают построение элементов ПрО: *GenerativeNode* отвечает за генерацию экземпляров классов, *AssociativeNode* – за генерацию отношений между экземплярами классов;
- *Source* обеспечивает доступ непосредственно к источнику данных;
- *Sink* обеспечивает запись результатов, полученных в отдельных узлах, в результирующую онтологию.

Каждый узел (т.е. экземпляр класса *Node*) описывается следующими свойствами:

- **объектные свойства:** *apply* указывает на узел, которому будут передаваться данные после обработки; *dataIn/dataOut* указывают на входные/выходные экземпляры классов онтологии источника данных; *join* указывает на узел типа *Sink*; *joinDomain/joinRange* указывают на узел типа *AssociativeNode*;
- **примитивные свойства:** *level* – номер уровня, к которому принадлежит узел; *extractor* – название элемента источника данных, который доступен для данного узла.

Таким образом, описываемая модель потока данных – это последовательность узлов, решающих следующие задачи:

- 1) чтение источника данных;
- 2) получение доступа к элементам иерархической структуры данных, полученной из источника данных;
- 3) порождение экземпляров классов ПрО на основе извлечённых данных;
- 4) связывание экземпляров классов отношениями;
- 5) запись полученных результатов.

Процесс связывания данных, описываемый предлагаемой моделью, можно представить как расщепление исходного источника данных на части и формирование на их основе множества потоков. Каждый поток несёт часть исходных данных. В соответствующих узлах эти данные преобразуются в элементы онтологии ПрО – экземпляры классов и отношения между экземплярами классов. В конце эти части собираются в единую структуру, формируя онтологическое описание ПрО.

### 3.2.8 Онтология «Лабиринт».

#### Правила для построения модели процесса связывания данных

Классы подграфа *FlowDesign* (см. рисунок 8) представляют собой элементы системы правил, которая используется для автоматического построения модели процесса связывания данных на основе онтологии отображения. Для этого онтология «Лабиринт» импортируется в построенную пользователем онтологию отображения.

Правила заданы в виде SPARQL-запросов: *WHERE*-часть SPARQL-запроса – antecedent, *CONSTRUCT*-часть – consequent. Запросы связаны с подклассами класса *FlowDesign* с помощью элементов языка SPIN<sup>6</sup>. TopBraid IDE<sup>7</sup> предоставляет удобный пользовательский интерфейс для построения и запуска таких правил.

На рисунке 9 представлен фрагмент интерфейса редактора и правило для построения уровней процесса связывания данных.

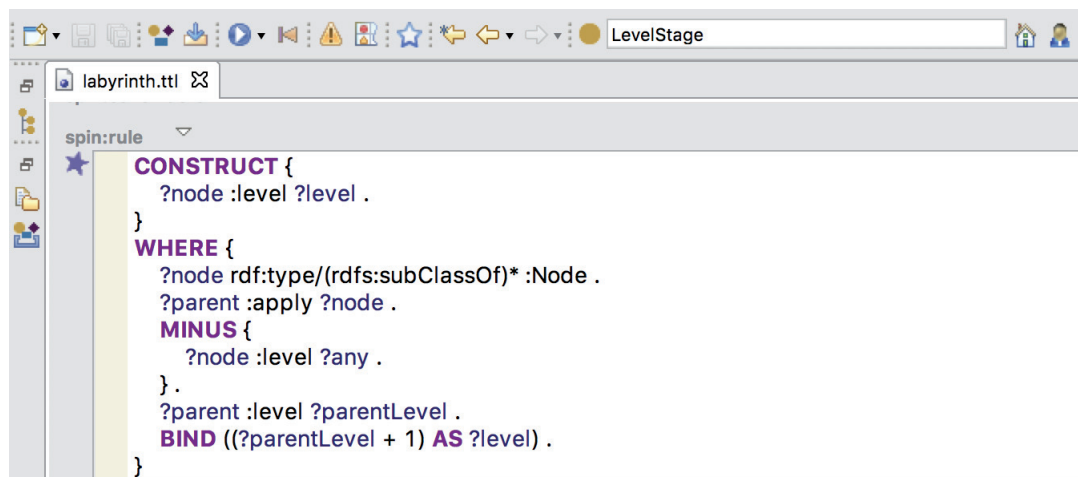


Рисунок 9 – Фрагмент интерфейса редактора с правилом построения уровней процесса связывания данных

О роли запроса каждого из подклассов *FlowDesign* (см. рисунок 8) можно судить по его названию. Например, класс *AssociativeNodeStage* отвечает за создание узлов, генерирующих отношения между двумя экземплярами.

<sup>6</sup> <http://spinrdf.org/>.

<sup>7</sup> <https://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/>.

### 3.2.9 Модель процесса связывания данных

На рисунке 10 представлена онтология процесса связывания данных – результат выполнения правил-запросов для онтологии отображения рассматриваемого примера (модель развернута сверху-вниз).

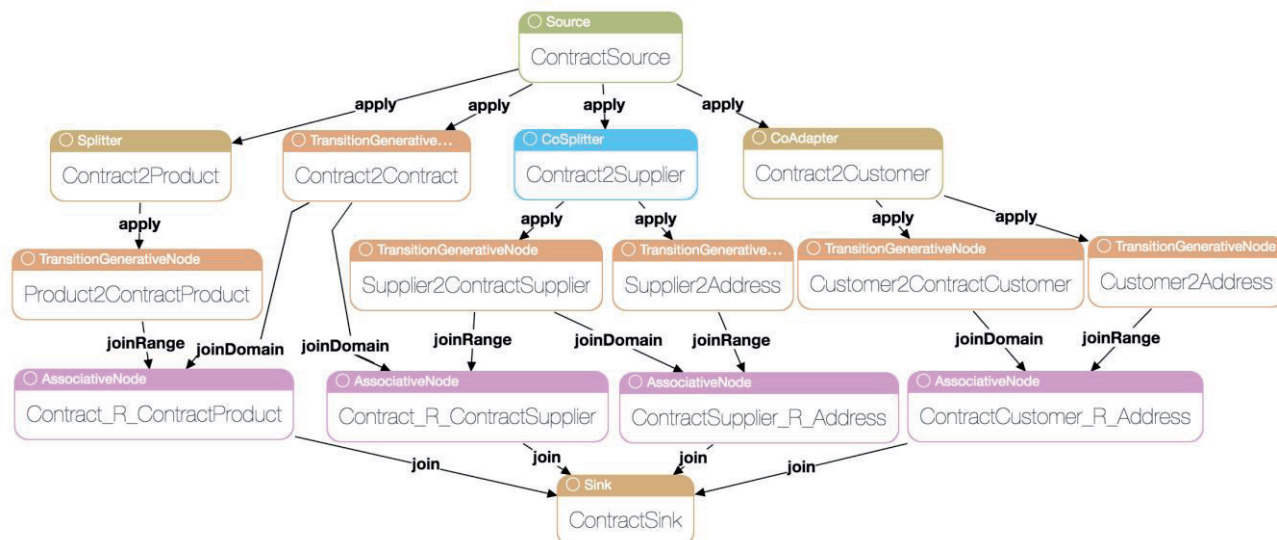


Рисунок 10 – Онтология процесса связывания данных

### 3.2.10 Генерация программного кода

На основе онтологии процесса связывания данных генерируется исполняемый код процесса. В качестве платформы для реализации выбрана библиотека *Apache Beam*<sup>8</sup>. Данную библиотеку можно рассматривать как своего рода общий интерфейс к различным реализациям высокопроизводительных вычислений на кластерах, таких как *Apache Spark*<sup>9</sup>, *Google Cloud Dataflow*<sup>10</sup>, *Apache Flink*<sup>11</sup> и др. *Apache Beam* на данный момент наиболее универсальное решение для реализации высокопроизводительных вычислений.

Процесс вычисления реализуется на языке программирования общего назначения *Kotlin*<sup>12</sup>. Для генерации кода по модели вычисления используется библиотека *KotlinPoet*<sup>13</sup>. Процесс генерации представляет собой итеративный процесс, состоящий из двух базовых действий: выполнение SPARQL-запроса к построенным онтологиям и генерация фрагмента программного кода на основе результатов запроса.

На рисунке 11 показана реализация процесса связывания на базе платформы *Google Cloud Dataflow*.

## 3.3 Оценка подхода

Предложенный подход сравнивался с инструментами *RML-Mapper* и *SPARQL-Generate*. Для приблизительной оценки того, как соотносится предлагаемое решение с иными решениями той же задачи, можно обратиться к [21].

<sup>8</sup> <https://beam.apache.org/>.

<sup>9</sup> <https://spark.apache.org/>.

<sup>10</sup> <https://cloud.google.com/dataflow/>.

<sup>11</sup> <https://flink.apache.org/>.

<sup>12</sup> <https://kotlinlang.org/>.

<sup>13</sup> <https://github.com/square/kotlinpoet>.

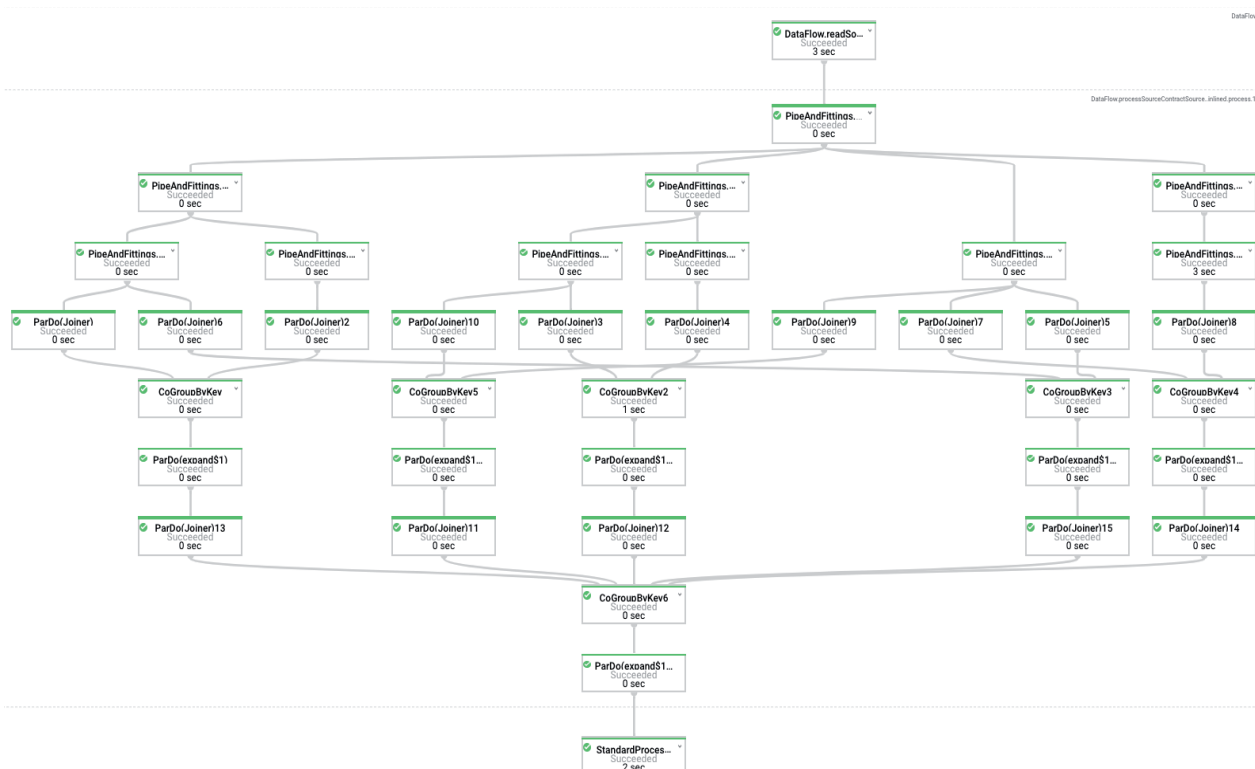


Рисунок 11 – Реализация процесса связывания данных на базе *Google Cloud Dataflow*

На рисунке 12 представлены результаты сравнения решений. Ось абсцисс фиксирует количество элементов в источнике данных. Под элементом понимается описание одного контракта. На оси ординат указано время выполнения запроса. DF (*Data Flow*) соответствует предлагаемому решению, запущенному на *Google Cloud Dataflow*; DF Elapsed Time – общее время обработки данных на базе *Google Cloud Dataflow*, включая время на построение кластера; DF Wall Time – процессорное время связывания данных; RML GCE (*Google Computation Engine*) Real Time – время связывания данных с помощью *RML-Mapper* на удалённой виртуальной машине; SG GCE Real Time – время связывания данных с помощью *SPARQL-Generate* на удалённой виртуальной машине.

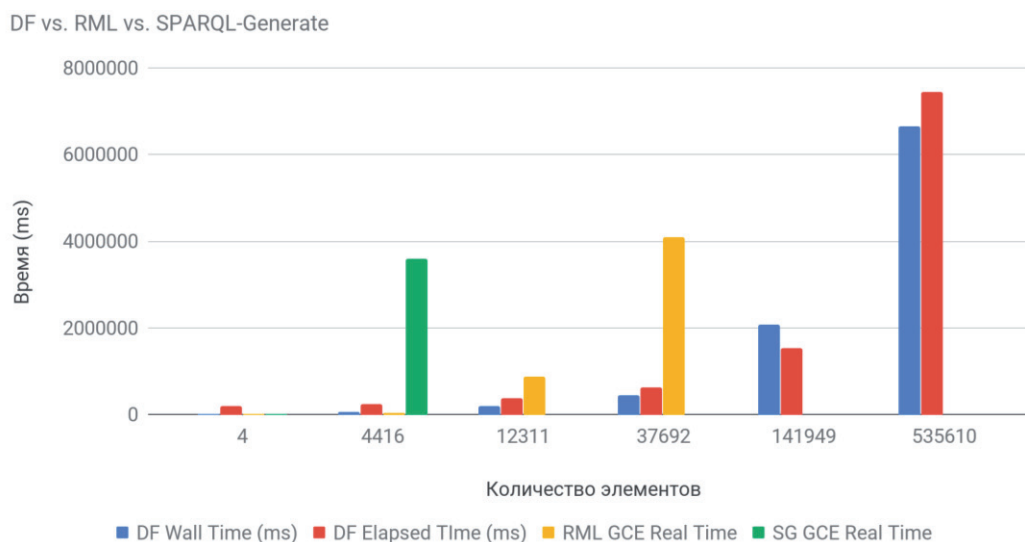


Рисунок 12 – Сравнение полученных результатов DF с *RML-Mapper* и *SPARQL-Generate*



Время для DF взято из пользовательского интерфейса, предоставляемого платформой *Google Cloud Dataflow*. Время для *RML-Mapper* и *SPARQL-Generate* измерялось с помощью команды *time* ОС Ubuntu. Команда выдаёт три значения времени – *real*, *user* и *sys*. В графике указано время *real*, т.е. время общей работы программы. Эквивалентным для него является время Wall Time.

Выполнение DF, кроме обработки 141949 элементов, было ограничено одной машиной типа *n1-standard-1* (по номенклатуре *Google Cloud Platform*): 1 процессор и 3.75 GB ОЗУ. При обработке 141949 использовалось 2 машины типа *n1-standard-1*, что видно из графика: процессорное время больше общего времени, т.к. обработка шла параллельно. Данный запуск был проведён без ограничений, масштабирование производилось автоматически.

В качестве GCE использовалась виртуальная машина типа *n1-standard-2* со следующими параметрами: 2 процессора и 7.5 GB ОЗУ.

Для 141949 и 535610 элементов не указано время преобразования для *RML-Mapper* и *SPARQL-Generate*, т.к. работа RML была прервана после нескольких часов ожидания, а *SPARQL-Generate* завершался с ошибкой из-за превышения доступной памяти.

По результатам сравнения можно сделать следующие выводы:

- при небольшом объёме данных предлагаемый подход проигрывает *RML-Mapper*, но при увеличении объёма превосходит его;
- реализация предложенного подхода на базе платформ высокопроизводительных вычислений позволяет воспользоваться реализованными в них средствами масштабирования вычислений (было продемонстрировано на наборе, состоящем из 141949 элементов).

## 4 Обсуждение

Простота рассмотренного сценария и функциональная ограниченность реализации не позволяют однозначно заявить о его преимуществах. Для полноценного сравнения решение должно реализовывать минимальный набор функциональных требований (см. первый раздел). Дополнительного обсуждения требует вопрос использования моделей для описания источника данных и процесса их связывания. Создание этих моделей требует определённых затрат, которые необходимо обосновать.

*RML-Mapper* и *SPARQL-Generate* используют специальные языки для фиксации отображений множества элементов источника данных во множество элементов ПрО: *RML-Mapper* использует язык RML [22], *SPARQL-Generate* [23] – расширение языка SPARQL. В указанных языках источник данных в явном виде не описывается. Для обращения к элементам источника данных используются строковые выражения. При этом если источник данных записан в формате JSON или XML, используются выражения на специальном языке. Например, для XML – это XPath. Такая строка не позволяет проверить тип хранимых данных или зафиксировать ограничения на возможные значения.

Напротив, явная онтологическая модель источника данных позволяет:

- документировать схему источника, что особенно полезно в тех случаях, когда элементы не имеют “дружественных” пользователю названий;
- добавлять ограничения (*owl:Restrictions*), например, на диапазон принимаемых элементов, и использовать эти знания для фильтрации неподходящих элементов данных;
- упростить повторное использование отображений (если известно, как источник данных связан с ПрО<sub>1</sub> и как ПрО<sub>1</sub> связана с ПрО<sub>2</sub>, то можно построить ПрО<sub>2</sub> на основе указанного источника).

Таким образом, при долговременном системном использовании разнородных источников этот подход будет более эффективным при создании средств, упрощающих построение и отображение онтологий.

## Заключение

В работе предложен подход для построения процесса связывания данных слабо-структурированных форматов в онтологическое представление, включающий:

- построение явных моделей источника данных и процесса преобразования;
- возможности автоматизации построения модели связывания данных и реализации процесса связывания данных на основе доступных масштабируемых высокопроизводительных платформ.

Развитие подхода должно быть направлено на более полное соответствие предложенным требованиям. Выполнение функциональных требований включает добавление возможности работы с несколькими источниками данных, возможности использовать функции для связывания данных, возможности работы с различными форматами. В части нефункциональных требований основная работа должна быть направлена на создание инструмента, упрощающего задачу построения онтологии источника данных и задачу построения отображения.

## Список источников

- [1] W3C Semantic Web Activity. – <https://www.w3.org/2001/sw/>.
- [2] W3C Data Activity – Building the Web of Data. – <https://www.w3.org/2013/data/>.
- [3] **Ishida, R.** An Introduction to Multilingual Web Addresses / R. Ishida. – <https://www.w3.org/International/articles/idn-and-iri/#problem>.
- [4] **Berners-Lee, T.J.** Linked Data – Design Issues / T.J. Berners-Lee. – <https://www.w3.org/DesignIssues/LinkedData>.
- [5] The Linked Open Data Cloud. – <https://lod-cloud.net/>.
- [6] Реестр наборов открытых данных. – <https://data.gov.ru/opendata>.
- [7] **Gonzalez, R.** RDF vs. XML / R. Gonzalez. – <https://www.cambridgesemantics.com/blog/semantic-university/learn-rdf/rdf-vs-xml/>.
- [8] Сведения о государственной регистрации юридических лиц, индивидуальных предпринимателей, крестьянских (фермерских) хозяйств. - <https://egrul.nalog.ru/>.
- [9] Открытые данные Федеральной налоговой службы Российской Федерации. - <https://www.nalog.ru/opendata/>.
- [10] Clearspending: Open data. - <https://clearspending.ru/opendata/>.
- [11] **Dimou, A.** What factors influence the design of a linked data generation algorithm? / A. Dimou, P. Heyvaert, B. De Meester, R. Verborgh // Proceedings of the 11th Workshop on Linked Data on the Web, LDOW (23 April, 2018, Lyon, France) – CEUR-WS.org, 2018. – 6 p.
- [12] **Dimou, A.** RML: A Generic Language for Integrated RDF Mappings of Heterogeneous Data / A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens, R. de Walle // Proceedings of the 7th Workshop on Linked Data on the Web, LDOW2014 (08 April, 2014, Seoul, Korea). – CEUR-WS.org, 2014. – Vol. 1184. – 5 p.
- [13] **Lefrançois, M.** A SPARQL Extension for Generating RDF from Heterogeneous Formats / M. Lefrançois, A. Zimmermann, N. Bakerally // 14th International Conference, ESWC (28 May – 1 June, 2017, Portorož, Slovenia) Springer, 2017. – P. 35–50.
- [14] **Morrison, J.P.** Flow-Based Programming: A New Approach to Application Development / J.P. Morrison □2nd ed. – Charleston: Createspace Independent, 2010. – 370 p.
- [15] **Vassiliadis, P.** A survey of extract–transform–load technology / P. Vassiliadis // International Journal of Data Warehousing and Mining (IJDWM). – Clayton: IGI Global, 2009. – Vol. 5, № 3. – P. 1–27.
- [16] **Knap, T.** UnifiedViews: An ETL Framework for Sustainable RDF Data Processing / T. Knap, M. Kukhar, B. Macháč, P. Škoda, J. Tomeš, J. Vojt // European Semantic Web Conference, ESWC (25-29 May, 2014, Anisaras, Crete, Greece) – Cham: Springer, Cham, 2014. – Vol. 475. – P. 379–383.
- [17] **Klímeck, J.** LinkedPipes ETL in use / J. Klímeck, P. Škoda // Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services, iiWAS '17 (4-6 December, Salzburg, Austria). – ACM Press, 2017. – P. 441–445.

- [18] **Gupta, S.** Karma: A System for Mapping Structured Sources into the Semantic Web / S. Gupta, P. Szekely, A.C. Knoblock, A. Goel, M. Taheriyani, M. Muslea // Extended Semantic Web Conference, ESWC (27-31 May, 2012, Heraklion, Crete, Greece). – Springer, 2015. – P. 430–434.
- [19] **Dimou, A.** Extending R2RML to a Source-independent Mapping Language for RDF / A. Dimou, M. Vander Sande, P. Colpaert, E. Mannens, R. de Walle // International Semantic Web Conference, Posters & Demos (21-25 October 2013, Sydney, Australia). – CEUR-WS.org, 2013 – Vol. 1035. – P. 237–240.
- [20] **Lefrançois, M.** Flexible RDF Generation from RDF and Heterogeneous Data Sources with SPARQL-Generate / M. Lefrançois, A. Zimmermann, N. Bakerally // European Knowledge Acquisition Workshop, EKAW (19–23 November, 2016, Bologna, Italy). – Springer, 2016. – P. 131–135.
- [21] **Kolchin, M.** A practical review of non-RDF to RDF converters / M. Kolchin. – <https://medium.com/datafabric/a-practical-review-of-non-rdf-to-rdf-converters-51686338927f>.
- [22] **Dimou, A.** RDF Mapping Language. Unofficial Draft / A. Dimou A., M. Vander Sande. – <http://rml.io/spec.html>.
- [23] SPARQL-Generate - Language Overview. – <https://ci.mines-stetienne.fr/sparql-generate/language.html>.

## MODEL-DRIVEN TRANSFORMATION OF HETEROGENEOUS SOURCES INTO LINKED DATA

S.V. Lebedev

*Saint Petersburg Electrotechnical University "LETI" and DataFabric Ltd., Saint Petersburg, Russia  
lebedev.sv.etu@gmail.com*

### Abstract

There are a lot of data sources that can be used to make a reasonable decision in the diversity of fields. But the linking of data from these sources into a solid view must come first. To represent linked data RDF language is used. Unfortunately, there are lots of those who publish data in non-RDF formats. Obviously, considering current data volumes integration task cannot be done in analytics' heads. A number of instruments to link heterogeneous data were proposed. The most flexible ones are based on special mapping languages. However, these solutions lack scaling and may not be effective for processing large data. In the presented paper, an approach for a possible solution is proposed. The idea of the approach is to abstract data linking process model separating it from a mapping as well as from a code implementation. The model is automatically generated based on a mapping of a set of source ontology elements into a set of domain ontology elements. The model of a process becomes available for separate manipulations relatively independent of other activities. The model can be used to optimize or to custom the linking process or can be treated as a specification for different implementations. As it based on the data-flow concepts it can be naturally translated into contemporary high-performance computation concepts. In other words, it is demonstrated how a process model can be implemented on one of such platforms. The obtained results make it reasonable to continue investigating and developing the proposed approach.

**Key words:** *linked data, linking process model, automation of model generation, program code generation, scalable solutions.*

**Citation:** *Lebedev SV. Model-driven transformation of heterogeneous sources into linked data [In Russian]. Ontology of designing. 2019. 9(1): 101-116. – DOI: 10.18287/2223-9537-2019-9-1-101-116.*

### References

- [1] W3C Semantic Web Activity. - <https://www.w3.org/2001/sw/>.
- [2] W3C Data Activity – Building the Web of Data. - <https://www.w3.org/2013/data/>.
- [3] **Ishida R.** An Introduction to Multilingual Web Addresses. - <https://www.w3.org/International/articles/idn-and-iri/#problem>.
- [4] **Berners-Lee T.J.** Linked Data – Design. - <https://www.w3.org/DesignIssues/LinkedData>.
- [5] The Linked Open Data Cloud. - <https://lod-cloud.net/>.
- [6] The register of sets of the open data [In Russian]. - <https://data.gov.ru/opendata>.

- [7] **Gonzalez R.** RDF vs. XML. - <https://www.cambridgesemantics.com/blog/semantic-university/learn-rdf/rdf-vs-xml/>.
- [8] Information about the state registration of legal entities, individual entrepreneurs, farms [In Russian]. - <https://egrul.nalog.ru/>.
- [9] Open data of the Federal Tax Service of the Russian Federation [In Russian]. - <https://www.nalog.ru/opendata/>.
- [10] Clearspending: Open data. - <https://clearspending.ru/opendata/>.
- [11] **Dimou A, Heyvaert P, De Meester B, Verborgh R.** What factors influence the design of a linked data generation algorithm? Proceedings of the 11th Workshop on Linked Data on the Web, LDOW (23 April, 2018, Lyon, France). CEUR-WS.org; 2018: 6.
- [12] **Dimou A, Vander Sande M, Colpaert P, Verborgh R, Mannens E, de Walle R.** RML: A Generic Language for Integrated RDF Mappings of Heterogeneous Data. Proceedings of the 7th Workshop on Linked Data on the Web, LDOW2014 (08 April, 2014, Seoul, Korea). CEUR-WS.org; 2014; 1184: 5.
- [13] **Lefrançois M, Zimmermann A, Bakerally N.** A SPARQL Extension for Generating RDF from Heterogeneous Formats. 14th International Conference, ESWC (28 May–1 June, 2017, Portorož, Slovenia). Springer; 2017: 35–50.
- [14] **Morrison JP.** Flow-Based Programming: A New Approach To Application Development, 2nd ed. Createspace Independent; 2010.
- [15] **Vassiliadis P.** A survey of extract–transform–load technology. International Journal of Data Warehousing and Mining (IJDWM). IGI Global; 2009; 5(3): 1–27.
- [16] **Knap T, Kukhar M, Macháč B, Škoda P, Tomeš J, Vojt J.** Unified Views: An ETL Framework for Sustainable RDF Data Processing. European Semantic Web Conference, ESWC (25–29 May, 2014, Anissaras, Crete, Greece). Springer, Cham; 2014; 475: 379–383.
- [17] **Klímeck J, Škoda P.** Linked Pipes ETL in use. Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services, iiWAS '17 (4–6 December, Salzburg, Austria). ACM Press; 2017: 441–445.
- [18] **Gupta S, Szekely P, Knoblock AC, Goel A, Taheriyan M, Muslea M.** Karma: A System for Mapping Structured Sources into the Semantic Web. Extended Semantic Web Conference, ESWC (27–31 May, 2012, Heraklion, Crete, Greece). Springer; 2015: 430–434.
- [19] **Dimou A, Vander Sande M, Colpaert P, Mannens E, de Walle R.** Extending R2RML to a Source-independent Mapping Language for RDF. International Semantic Web Conference, Posters & Demos (21–25 October 2013, Sydney, Australia). CEUR-WS.org; 2013; 1035: 237–240.
- [20] **Lefrançois M, Zimmermann A, Bakerally N.** Flexible RDF Generation from RDF and Heterogeneous Data Sources with SPARQL-Generate. European Knowledge Acquisition Workshop, EKAW (19–23 November, 2016, Bologna, Italy). Springer; 2016: 131–135.
- [21] **Kolchin M.** A practical review of non-RDF to RDF converters. - <https://medium.com/datafabric/a-practical-review-of-non-rdf-to-rdf-converters-51686338927f>.
- [22] **Dimou A, Vander Sande M.** RDF Mapping Language. Unofficial Draft. - <http://rml.io/spec.html>.
- [23] SPARQL-Generate - Language Overview. - <https://ci.mines-stetienne.fr/sparql-generate/language.html>.

## Сведения об авторе



**Лебедев Сергей Вячеславович**, 1984 г. рождения. Окончил Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина) (СПбГЭТУ ЛЭТИ) в 2010 г. Ассистент кафедры вычислительной техники СПбГЭТУ ЛЭТИ, программист в ООО «Датафабрик». В списке научных трудов более 10 работ в области многоагентных систем и технологий семантического Web.

**Sergey Vyacheslavovich Lebedev** (b. 1984) graduated from the Saint Petersburg Electrotechnical University "LETI" (ETU) in 2010. He is an assistant at ETU (Department of Computer Science and Engineering) and a programmer in DataFabric LLC. He is co-author of more than 10 publications in the field of multi-agent systems and semantic web technologies.

УДК 004.822

## KNOWLEDGE NET: МОДЕЛЬ И СИСТЕМА НАКОПЛЕНИЯ, ПРЕДСТАВЛЕНИЯ И ИСПОЛЬЗОВАНИЯ ЗНАНИЙ И ДАННЫХ

О.Н. Тушканова<sup>1,2</sup>, В.В. Самойлов<sup>3</sup>

<sup>1</sup> Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, Россия

<sup>2</sup> Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия  
tushkanova.on@gmail.com

<sup>3</sup> ООО «Информационные крылья», Санкт-Петербург, Россия  
vsamoilov@infowings.ru

### Аннотация

Описана разрабатываемая авторами модель Knowledge Net, предназначенная для формализации накопления, представления и использования знаний и данных единого информационного пространства интеллектуального предприятия, в том числе в рамках концепции Индустрия 4.0. В основу модели Knowledge Net положена графовая модель данных. Модель Knowledge Net является частью прикладных разработок цифровой платформы предприятия. Разработка программной системы, поддерживающей модель Knowledge Net, выполняется в концепции открытого программного кода. В работе приведён пример использования программного прототипа для описания объектов производственного предприятия. Ключевой особенностью Knowledge Net, отличающей разработку от известных аналогов, является комбинация таких свойств, как поддержка многоаспектности описания субъектов и объектов предприятия, возможность расширения сети понятий и сущностей вместе с развитием предприятия, динамическое формирование структуры аспектов и их свойств, множественность возможных типизаций сущностей, представленных моделью Knowledge Net, а также поддержка версии модели знаний и данных.

**Ключевые слова:** единое информационное пространство, модель знаний, представление знаний, аспект, онтология, граф знаний, графовая модель данных.

**Цитирование:** Тушканова, О.Н. Knowledge Net: модель и система накопления, представления и использования знаний и данных / О.Н. Тушканова, В.В. Самойлов // Онтология проектирования. – 2019. – Т. 9, №1(31). – С.117-131. – DOI: 10.18287/2223-9537-2019-9-1-117-131.

### Введение

Одной из основных черт современной концепции цифрового производства, называемой Индустрия 4.0, является сращивание промышленного производства, информационно-коммуникационных технологий и технологий искусственного интеллекта. Эта тенденция развития промышленности в настоящее время является одной из наиболее актуальных и часто обсуждаемых тем среди учёных и практиков всего мира. Правительства многих стран, включая Российскую Федерацию [1], объявили концепцию цифровой экономики ключевой стратегией развития в области высоких технологий.

По данным исследования [2] руководители крупнейших компаний в области промышленного производства, добычи нефти и газа, энергетики, горнодобывающей промышленности в Америке, Азии и Европе готовы инвестировать значительные средства в Индустрию 4.0 с целью цифровой трансформации их предприятий. В качестве основной цели инвестирования руководители чаще всего называют стремление превратить свои корпорации в сеть сильно связанных гибких компаний, функционирующих на основе знаний, которые будут способны удовлетворять требованиям нового цифрового мира.



Основная идея концепции Индустрии 4.0 направлена на развитие интеллектуальных фабрик, которые должны функционировать в рамках автоматизированных производств, управляемых программным обеспечением, а также на интеграцию всех информационных и программных компонентов на уровне предприятия в единое информационное пространство (ЕИП). Программные системы интеллектуальных фабрик должны полностью контролировать все физические процессы, происходящие на предприятии, эффективно взаимодействовать и сотрудничать друг с другом и с персоналом, а также быть способными принимать децентрализованные решения. Для большинства отраслей реализация таких преобразований является серьёзной проблемой, так как чаще всего информационные компоненты и программные системы исторически разрабатывались независимо друг от друга.

Обязательной компонентой интеллектуального предприятия, развивающегося в рамках концепции Индустрии 4.0, является система накопления, представления и использования данных и знаний, которая должна формировать семантически однозначное ЕИП для всех его подразделений, а также физических и виртуальных объектов производственной системы [3]. Такое пространство должно агрегировать информацию обо всех субъектах и объектах предприятия на всех этапах жизненного цикла (ЖЦ).

Особенность ЕИП состоит в том, что в нём разные подразделения могут рассматривать один и тот же объект или субъект с разных точек зрения. Например, отдел логистики могут интересовать массово-габаритные характеристики некоторого изделия, а отдел закупок – экономически выгодные аналоги сборочных единиц, которые обладают сходными функциональными возможностями. Это формирует специфическое требование к системе накопления знаний интеллектуального предприятия, сущность которого состоит в обеспечении возможности описания различных аспектов одного и того же объекта в рамках единой структуры представления знаний и данных информационного пространства. Следствием этого требования является вывод о том, что в структуре данных и знаний, используемой для совместной работы множества подразделений в ЕИП, не может быть единой типизации (таксономии) понятий онтологии.

Другой важной специфической чертой интеллектуального предприятия является необходимость постоянного развития и пополнения данных и знаний ЕИП. Действительно, вместе с развитием предприятия, совершенствованием используемых механизмов управления и соответствующей эволюцией программных компонент семантические модели объектов и субъектов также должны постоянно развиваться. Динамика моделей объектов и субъектов производственной системы определяется ЖЦ производства и производимых продуктов, и для каждого его этапа может потребоваться своя модель информации. Параллельно с эволюцией потребных моделей информации необходимо заботиться о поддержке целостности знаний и данных. Такое требование к динамической поддержке баз знаний и данных требует разработки соответствующих формальных методов и программных средств.

В соответствии с современными представлениями задача формирования и динамической поддержки ЕИП знаний и данных предприятия в рамках концепции Индустрии 4.0 возлагается на инфраструктурную компоненту системы управления предприятием, которую принято называть его цифровой платформой. Основной задачей этой платформы является интеграция гетерогенных компонент системы на всех этапах её ЖЦ.

Целью данной работы является описание предлагаемой модели накопления, представления и использования знаний и данных ЕИП производственной системы с учётом её многоаспектности, динамики и требований по поддержке её целостности, а также соответствующего программного прототипа. Эта модель является частью текущих прикладных разработок цифровой платформы предприятия. Соответствующий проект, программный прототип и описы-

ваемая модель получили название Knowledge Net. Разработки системы в данный момент выполняются в концепции открытого программного кода.

В разделе 1 обсуждаются наиболее близкие аналоги существующих систем представления знаний и данных. В разделе 2 пояснены некоторые понятия, важные для понимания принципов устройства модели Knowledge Net. В разделе 3 формулируются ключевые особенности разрабатываемой модели Knowledge Net. Раздел 4 посвящён общему описанию модели Knowledge Net и методики её применения. В разделе 5 кратко описано программное обеспечение, поддерживающее разработанную модель. В разделе 6 приведён пример описания производственного объекта модели Knowledge Net.

## 1 Краткий обзор систем семантического представления знаний и данных

В разделе представлены разработки, в которых использованы подходы, частично реализующие возможности описанной в данной работе модели и системы Knowledge Net.

В [4] предлагается концепция, получившая название Administrative Shell, которая должна поддерживать цифровое представление всей доступной информации об объектах производства, представляющих собой, например, некие механизмы, их части или программные компоненты. Цифровое представление объектов в рамках этой концепции реализуется с помощью таких семантических средств представления знаний как RDF, RDF-Schema и OWL. Использование RDF-концепции имеет несколько преимуществ, среди которых следует упомянуть встроенную возможность использования глобального уникального идентификатора (URI) объекта, возможность её простой интеграции с уже существующими базами знаний, описанными на RDF, а также возможность использовать язык SPARQL в качестве языка запросов. Следует отметить, что предложенная в [4] схема описания объектов не универсальна и на момент публикации работы поддерживала только некоторые типы объектов. Кроме того, использование этого подхода требует предварительного создания тезауруса и таксономии классов объектов. Это не подходит для развивающегося предприятия, так как в этом случае модель знаний должна эволюционировать вместе с развитием объектов, технологий и организационных структур предприятия в целом. В работе не описаны способы отслеживания истории изменений описаний объектов или схемы. Следует отметить, что языки RDF, RDF-Schema и OWL такую возможность не предусматривают.

В [5] описываются результаты сотрудничества между компанией Siemens и Оксфордским университетом, целью которого было облегчение разработки онтологий и их использования в приложениях компании. Итогом этого сотрудничества стала надстройка над Web Protege – программный прототип Siemens-Oxford Model Manager (SOMM). Этот инструмент рассматривается как помощник инженеров, имеющих небольшой опыт работы с семантическими технологиями, при создании онтологических моделей и заполнении их данными. Онтологии, созданные с помощью SOMM, рассматриваются авторами как ядро ЕИП компании Siemens. SOMM реализует фрагмент OWL 2 RL и поставляется вместе с компонентой поддержки схемы и рассуждений над данными. Основным недостатком такого онтологического подхода для создания ЕИП предприятия, который предлагается также в [6, 7], является трудоёмкость первоначального создания схемы, которая позволяла бы описать все объекты реального предприятия. В случае, если предприятие только начинает развиваться и набор используемых им объектов и компонентов постоянно меняется, разработка единой неизменяемой схемы практически невозможна или сопряжена с большими затратами. Более того, даже крупные компании часто не способны разработать свои онтологии настолько детально, чтобы стало возможно применение в них механизмов логического вывода.

В работе [8] описаны основные принципы устройства семантического ядра цифровой платформы xIRBIS-ML. Авторы позиционируют эту разработку как информационно-аналитическую систему, ориентированную на знания. По утверждению авторов, подсистема сбора и обработки документов системы xIRBIS-ML позволяет извлекать из текстовых документов графы понятий и сопоставлять их с существующими онтологиями предприятия. Однако в работе не описана степень участия эксперта для успешного выполнения операции. Отсутствует подробное описание технологий, используемых для создания и обработки онтологий, а также реализации возможности системой поддержки уникальных идентификаторов для всех сущностей системы и поддержки динамических изменений созданной схемы онтологии в процессе развития предприятия.

В работах [9, 10] описаны некоторые принципы новой технологии компонентного проектирования интеллектуальных систем, в частности предлагается использовать семантическую модель баз знаний и привлекать теорию графов для построения алгоритмов их обработки. В этих работах приводится подробное описание базового универсального языка семантических сетей (SC-кода), который, по мнению авторов [9, 10], должен стать стандартом для разработки унифицированных семантических сетей, обеспечивающих представление и интеграцию различных видов знаний в интеллектуальных системах нового поколения, а также описание этапов разработки таких интеллектуальных систем. Однако из текста работ не ясно, выполнялись ли авторами какие-либо прикладные программные разработки, направленные на реализацию описанного подхода и технологии.

Анализ исследований, посвящённых разработке методов и средств создания и поддержки ЕИП интеллектуального предприятия, показывает высокую заинтересованность индустрии в таких средствах. Но уровень их развития является недостаточным для активного внедрения результатов разработок.

## **2 Ключевые свойства модели Knowledge Net**

Семантическая модель накопления, представления и использования данных и знаний является ключевой компонентой интеллектуального предприятия, развиваемой в рамках концепции Индустрии 4.0 [3]. Основной задачей этой компоненты является предоставление сервиса *создания и поддержки* ЕИП знаний и данных интеллектуального предприятия для всех его подразделений. Можно говорить также о поддержке всего ЖЦ такого пространства.

Модель Knowledge Net разрабатывается в рамках исследовательского проекта, а соответствующее программное обеспечение является вариантом реализации этой модели. Ключевыми особенностями разрабатываемой модели ЕИП Knowledge Net, отличающими её от существующих аналогов являются:

- *учёт динамической природы модели* описания, представления и формализации знаний и данных: описание свойств и структуры представления моделей объектов, которые эволюционируют вместе с динамикой развития объектов и предприятия в целом;
- *поддержка многоаспектности модели* описания и представления знаний об объекте и характеризующих его данных: объект может иметь множество различных описаний и свойств, связанных с различными контекстами его использования или “точками зрения” на него (асpekтами) со стороны пользователей и задач;
- *графовая модель представления* пространства знаний и данных: аспекты и свойства объекта представляются в форме единой динамически расширяемой модели графа, который образует единую сеть для представления знаний и данных, называемую *метаграфом*, в котором реализация случаев использования знаний и данных сводится к поиску на этом метаграфе;

- *множественность возможных типизаций* (таксономий) объектов, представленных моделью Knowledge Net: формальные источники, как и эксперты-носители знаний, могут формировать различные описания “похожих” объектов, используя множество различных типизаций;
- *многообразие типов и структур данных*, которые могут быть представлены в графе Knowledge Net: свойства объектов, представленных в системе, могут быть описаны атрибутами любых типов, включая слабо структурированные (xml, json и т.п.) и неструктурированные данные, например, CAD-модели, PDF-файлы, текстовые описания, изображения, ссылки на другие объекты и их свойства и т.п.;
- *единая ссылочная сеть*: модель даёт возможность выполнения нечёткого поиска, выявления аналогов и кластеризации объектов, субъектов и аспектов-свойств;
- *поддержка специализированных механизмов фасетного поиска* объектов: поиск аналогов, поддержка поиска примеров объектов с заданными свойствами и характеристиками;
- *версионность* модели знаний и данных: возможность хранения истории *всех* изменений *всех* версий сущностей модели знаний Knowledge Net, которые происходили на всех этапах ЖЦ этих сущностей, как это, например, и реализуется в структурах представления знаний и данных типа Data Vault [11].

### 3 Используемая терминология

Под *субъектом* в данной работе понимается некоторый носитель, поставщик или пользователь знаний, активный элемент внешнего (по отношению к системе управления предприятием) мира. Субъект может представлять собой организацию или индивида, обладающего сознанием и волей, который способен к целесообразной деятельности, направленной на тот или иной объект. Примерами субъектов в бизнес-процессах организации являются поставщики, производители, разработчики продукта, сама организация, сервис или программа, осуществляющая добавление данных или знаний в модель.

*Объект* – это физический или виртуальный результат целенаправленной деятельности *субъекта*, предмет или явление реального или виртуального мира. Примерами объектов являются: товар, производимый предприятием, деталь, устройство, сборочная единица и т.д., услуга, сервис и т.п. Объекты могут обладать различными свойствами и могут быть представлены различными описаниями.

Ключевым понятием разрабатываемой модели является понятие *аспекта*. Согласно [12], под аспектом следует понимать “взгляд, точку зрения, с которой рассматривается какое-либо явление, понятие, перспектива”. *Аспектом* некоторого объекта будем называть одно из однотипных проявлений свойств его модели, общее для всех его описаний. Аспект обладает *наименованием*, возможно, *единицей измерения*, может характеризоваться *типами данных*, описывающих *свойства*, относящиеся к аспекту, и *доменами* их допустимых значений. В качестве домена допустимых значений аспекта могут выступать, например, *словари*, численные интервалы или формулы, задающие некоторые ограничения. Аспекты могут быть общими (глобальными) либо существовать только в некотором контексте, который задаётся определённым *субъектом*. Примерами аспектов могут выступать, например:

- *Вес* (вещественный тип, один из вариантов единицы измерения - килограмм, домен значений -  $[0; +\infty)$ );
- *Объём* (вещественный тип, один из вариантов единицы измерения - это литр, домен значений -  $[0; +\infty)$ );
- *Количество фаз переменного тока* (целый тип, единица измерения – штуки, домен значений - все целые неотрицательные числа);



- *Цвет* (строковый тип, нет единицы измерения, возможный домен значений – множество всех цветов) и т.д.

В рамках предлагаемой модели аспекты могут объединяться в иерархии. В этом случае дочерние аспекты будем называть *свойствами* по отношению к родительским аспектам. *Свойство аспекта* – это проявление некоторого дочернего аспекта в контексте родительского аспекта – владельца свойства. Например, *Вес* является самостоятельным аспектом, но он также может быть свойством других аспектов. Аспект *Массогабаритные показатели* может иметь свойства *брутто Вес* и *нетто Вес*, которые ссылаются на аспект *Вес*. Тот же аспект *Вес*, при этом, может входить и в свойство *максимальный для размещения Вес* в контексте родительского аспекта *Характеристики устройства хранения* (стеллажа или полки).

*Свойство объекта* – это проявление аспекта в том или ином контексте, задаваемом конкретным объектом. Свойство, в общем случае, обладает ссылкой на аспект, к которому оно относится, ролевым именем свойства (возможно пустым), мощностью множества допустимых значений и, возможно, конкретными значением. Мощность свойства означает допустимое число экземпляров значений этого свойства у объекта. Это понятие аналогично общепринятому понятию *мощность* отношений (кратность).

*Словарь* в контексте разрабатываемой системы – это допустимое множество значений некоторого аспекта, имеющего строковый тип. Примером словаря может выступать список возможных цветов для аспекта, описывающего цвет объекта; иерархия классификации объектов по функциональному назначению согласно стандарту [13] также может быть представлена с помощью словаря, а затем использована для указания функционального назначения объекта.

#### **4 Описание модели Knowledge Net и методики её применения**

Основными отличительными особенностями разрабатываемой модели Knowledge Net является поддержка многоаспектности описания сущностей, возможность расширения сети понятий и сущностей, динамического формирования структуры аспектов и их свойств. Выполнение этих требований подразумевает отсутствие жёсткой типизации сущностей, что приводит к сложности хранения данных и знаний в традиционной реляционной или объектной структуре. В связи с этим в основу модели Knowledge Net положена графовая модель данных. Под графовой моделью данных обычно понимают модель, в которой схема и экземпляры данных представляются в структурированном виде как ориентированный размеченный граф, а все манипуляции с данными и ограничения целостности данных могут быть выражены через операции на этом графе [14]. Примерами операций являются нахождение кратчайшего пути (например, при оценке семантической близости понятий онтологии, представленной графом), выделение компонент связности (при поиске независимых компонент знаний и в задачах их кластеризации) и т.д.

В качестве системы управления графовой базой данных выбрана СУБД OrientDB с открытым исходным кодом, которая объединяет в себе возможности документо-ориентированной и графовой базы данных. Кроме того, OrientDB поддерживает интерфейс объектно-ориентированной базы данных, который работает поверх документо-ориентированного слоя. OrientDB написана на языке Java и распространяется под лицензией Apache [15]. Важным преимуществом OrientDB является поддержка протокола Apache's Tinker Pop, что позволит в будущем реализовать импорт знаний из существующих RDF-хранилищ, например, из ресурсов Linked Data. Поддержка интерфейса объектно-ориентированной базы данных означает, что СУБД OrientDB позволяет задавать классы вершин и рёбер будущего графа, аналогичные классам в объектно-ориентированной модели.



На логическом уровне модели следует выделить некоторые основные сущности схемы данных, которые являются понятиями метаонтологии или метаграфа, а именно, *Субъект*, *Аспект*, *Объект*, *Единица измерения*. К вспомогательным сущностям будем относить такие понятия как *Запись в истории*, *Свойство аспекта*, *Справочник*, *Элемент Справочника* и некоторые другие.

Логическая модель данных, представленная графом (рисунок 1), является ядром модели Knowledge Net. Для представления этой логической модели в Orient DB были созданы следующие классы вершин: *Subject*, *Object*, *Aspect*, *Aspect Property*, *Object Property*, *Object Property Value*, *Measure*, *Measure Group*, *History Event*, *Reference Book Item*, *GUID* (англ. Global Unified Identifier) и др.

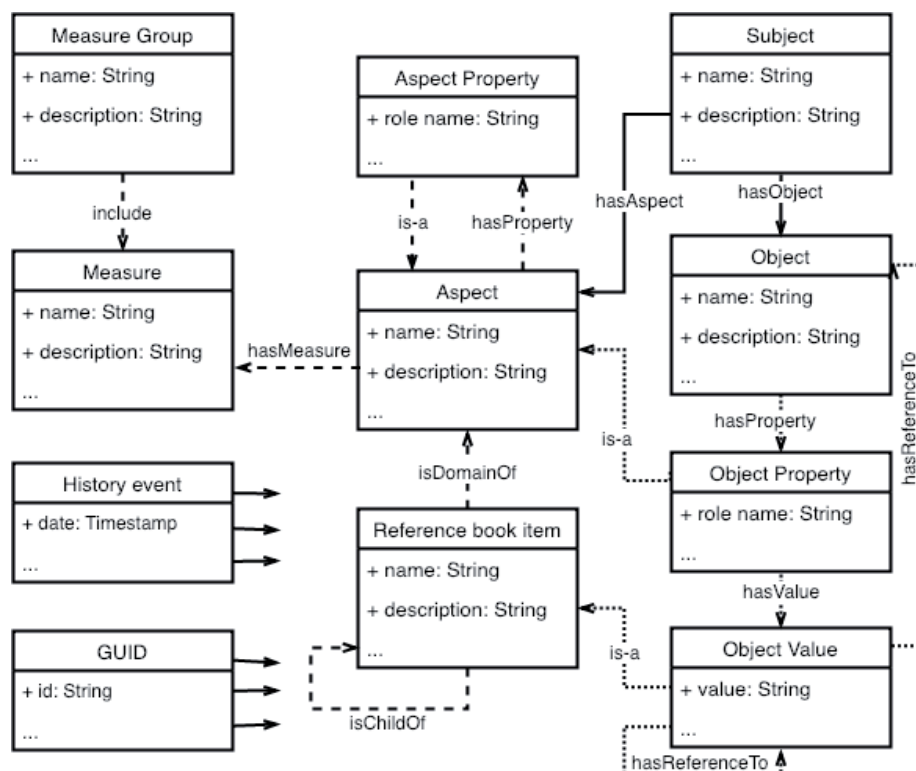


Рисунок 1 - Основные сущности графа знаний и связи между ними

Описанная структура модели Knowledge Net удобна тем, что она позволит организациям самостоятельно или с привлечением экспертов создавать графы знаний, которые могут быть положены в основу модели ЕИП интеллектуального предприятия.

Ключевыми вершинами этого графа знаний являются узлы *Aspect* и *Aspect Property*, предназначенные для описания аспектов и свойств этих аспектов. Сами свойства аспектов также являются аспектами, но с некоторыми дополнениями, например, с уточняющими ролевыми именами. Пара *аспект – свойство* чаще всего используется для группировки связанных аспектов. Например, свойствами аспекта *Геометрия объекта* могут быть аспекты *Длина*, *Ширина*, *Высота*.

Вершины типа *Measure* и *Measure Group* позволяют описывать единицы измерения, которые могут быть ассоциированы с аспектами. Узлы типа *Reference Item* служат для представления словарей, которые описывают допустимое множество значений некоторого строкового аспекта и их элементов. Штриховым пунктиром на рисунке 1 выделены отношения, которые являются основными при описании аспектов графа знаний предприятия.

Вершины типа *Subject* предназначены для описания основных свойств субъектов графа знаний. Объекты могут существовать только в контексте принадлежности некоторому субъекту, в то время как аспекты могут быть и глобальными, то есть не относиться к какому-либо субъекту. Узлы типа *Object*, *Object Property* и *Object Value* позволяют описывать объекты и их свойства. Каждый узел типа *Object Property* всегда связан с аспектом, который он реализует.

Описанный подход к представлению и хранению знаний и данных позволяет поддерживать многоаспектность описания объекта графа знаний, так как для добавления нового аспекта описания объекта, рассматривающего этот объект с другой точки зрения, потребуется всего лишь добавить новый узел графа *Object Property* и связать его с этим объектом. Точечным пунктиром на рисунке 1 выделены отношения, которые позволяют описывать взаимосвязи объектов и их свойств.

Граф знаний, который может быть разработан с использованием модели Knowledge Net, рассматривается авторами как общее информационное ядро (или часть такового) для множества систем в рамках как одной, так и нескольких организаций. Эти системы, с одной стороны, должны иметь возможность ссылаться внутри себя на сущности графа, а с другой стороны, иметь возможность обмениваться между собой этими ссылками без необходимости непосредственного доступа к графу знаний. Такие ссылки должны иметь глобальный контекст и обеспечить для любой внешней системы возможность получить по ним одинаковую непротиворечивую информацию из графа. В качестве ссылок на все типы сущностей для внешних и внутренних систем используется уникальный идентификатор *GUID*, представленный в виде отдельных узлов графа, связанных с каждой сущностью. В течение всего ЖЦ сущности в графе знаний её *GUID* остаётся неизменным. Наличие такого глобального идентификатора делает возможным хранение всей истории изменений всех сущностей в системе на протяжении всего их ЖЦ. Хранение истории изменений реализуется с помощью узлов типа *History Event* и нескольких вспомогательных типов узлов, не представленных на рисунке 1. Узлы *GUID* и *History Event* связаны с каждой сущностью графа (на рисунке 1 эти связи не представлены). Предполагается, что наполнение соответствующего графа предприятия в рамках модели Knowledge Net может выполняться как вручную, так и автоматически с привлечением методов машинного обучения для извлечения знаний.

## **5 Программное обеспечение, поддерживающее модель Knowledge Net**

Развитие модели Knowledge Net выполняется в рамках исследовательского проекта и поддерживается конкретными программными разработками. Программный код выложен в открытый доступ на хостинге GitHub [16].

Программная система, поддерживающая модель Knowledge Net, разрабатывается как веб-приложение с упрощённым пользовательским интерфейсом (рисунок 2), который позволяет вводить, искать и просматривать сущности разработанного графа знаний организации. На данный момент этот интерфейс является основным средством пользовательского доступа к описываемым сущностям. С ростом объёма данных основными средствами доступа к объектам станут API для внешних и внутренних подсистем предприятия, функционирующих в контексте единой платформы и использующих граф, а также специализированные “продвинутые” пользовательские интерфейсы с расширенными средствами поиска, навигации и отображения. После этого роль существующего пользовательского интерфейса станет скорее служебной, используемой при администрировании и контроле информационного пространства интеллектуального предприятия.

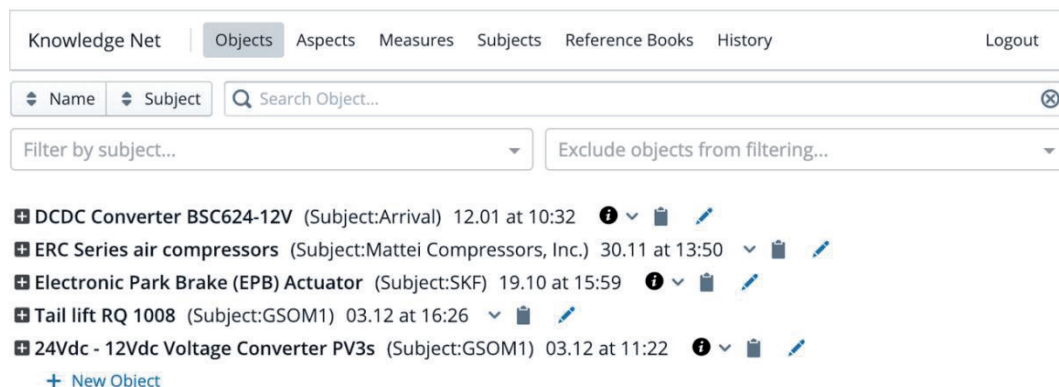


Рисунок 2 – Пользовательский интерфейс базовой версии программной системы модели Knowledge Net

На рисунке 2 представлена подсистема для ввода и отображения объектов организации. В верхней части страницы представлены кнопки переключения на другие разделы интерфейса, поддерживающие работу с аспектами (*Aspects*), единицами измерения (*Measures*), субъектами (*Subjects*), словарями (*Reference Books*) и историей изменений (*History*). Основная часть страницы подсистемы отведена под список объектов, каждый из которых может быть развернут для просмотра его свойств. Создание нового объекта выполняется на отдельной странице. Раздел работы с аспектами (рисунок 3) имеет схожую структуру.

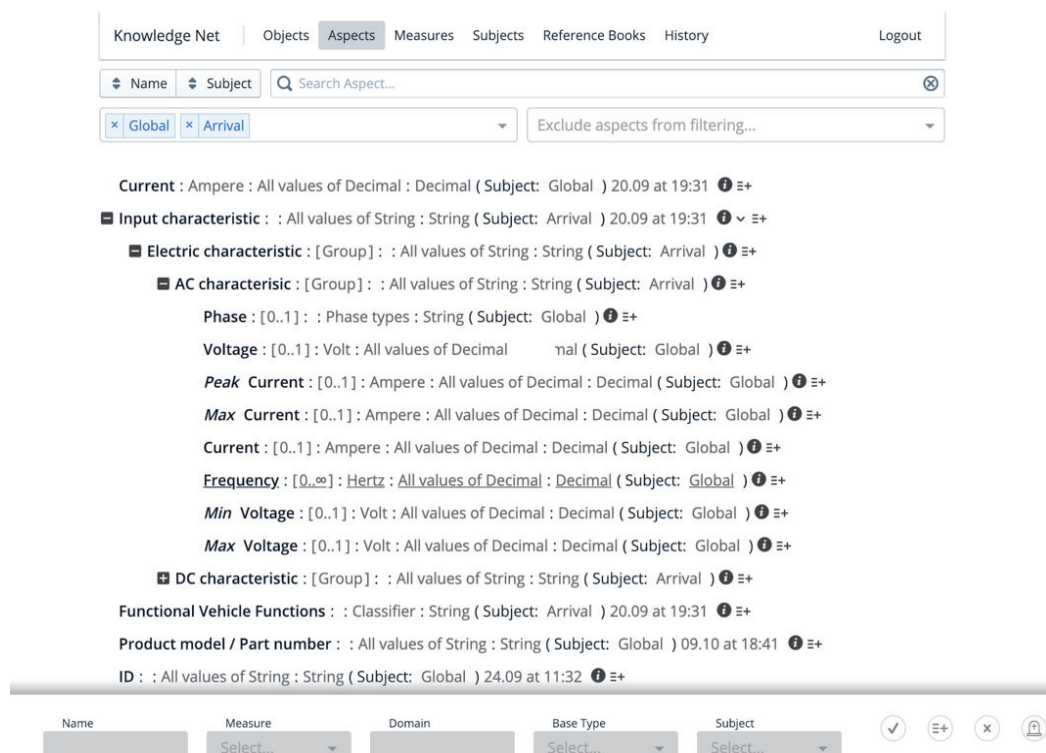


Рисунок 3 – Пользовательский интерфейс базовой версии для работы с аспектами

Разработанное программное обеспечение модели Knowledge Net, обладает следующими основными функциональными возможностями.

- Поддержка основных единиц измерения из общероссийского классификатора единиц измерения [17], а также механизма нечёткого поиска по ним.
- Создание уникального идентификатора (GUID) всех сущностей графа знаний внутри системы Knowledge Net.

- Просмотр и редактирование субъектов, что позволяет вносить в граф знаний информацию о субъектах организации (в первой версии системы субъекты могут обладать только уникальным именем, а интерфейс позволяет просмотреть все сущности, которые связаны с каждым субъектом).
- Создание, редактирование и просмотр аспектов, в том числе, поддержка ввода и редактирования наименования, единицы измерения, типа и домена значений. Аспекты могут иметь свойства, которые образуют деревья вложенности (дочернее свойство является фактически новой сущностью, представляющей собой комбинацию “ролевого” имени, по умолчанию пустого), ссылки-связи на другой аспект, который включается как дочерний, и мощности значения. Аспекты-свойства не могут в глобальном графе образовывать циклические ссылки по связям Аспект → Аспект-свойство → Аспект. Это ограничение проверяется при попытке сохранить новое свойство аспекта. Поиск аспектов-свойств осуществляется фасетным способом по наименованию, единицам измерения или типу.
- Поддержка аспектов, имеющих тип *связь* (значением для свойств, реализующих такие аспекты, может быть ссылка на другие сущности системы. В текущей версии программного обеспечения поддержаны ссылки на объекты и на значения других свойств).
- Фильтрация аспектов (в графе знаний количество используемых аспектов будет достаточно большим, при этом разнообразие будет по количеству субъектов, по тематике аспектов и по их структурной сложности).
- Нечёткий поиск по аспектам в графе (в текущей версии программного обеспечения используется свободно распространяемая библиотека для высокопроизводительного полнотекстового поиска Lucene, которая интегрирована в OrientDB).
- Просмотр и редактирование объектов, что позволяет описывать объекты организации и их свойства с использованием введённых заранее аспектов.
- Просмотр связей объекта и поиск связанных сущностей (разработанный интерфейс позволяет по запросу показать связанные с объектом или субъектом сущности).

Таким образом, в базовом программном обеспечении поддержаны основные особенности модели Knowledge Net, связанные с накоплением и представлением знаний, которые позволяют ввести семантические описания объектов в систему и начать использовать полученный граф как основу для ЕИП предприятия.

## 6 Пример описания объекта

Программное обеспечение, поддерживающее модель Knowledge Net, было апробировано на примере создания фрагмента графа знаний для компании-производителя электромобилей. Всего было обработано около 50 объектов разного уровня сложности, которые были отмечены компанией, как наиболее приоритетные. Основные аспекты (около 200), необходимые для описания этих объектов, были выделены и структурированы в рамках другого исследовательского проекта, результаты которого частично описаны в [18]. Такого количества объектов и описывающих их аспектов оказалось достаточно, для того чтобы сделать первые выводы о работоспособности модели.

С целью демонстрации возможностей модели и программного прототипа рассмотрим описание объекта *DCDC Converter BSC624-12V* – конвертера постоянного тока, – выполненное в системе Knowledge Net (рисунок 4).

Для описания объекта в основном использованы аспекты, обладающие свойствами, то есть родительские и дочерние аспекты с несколькими уровнями вложенности, объединенные в иерархии. Например, аспект *Geometry* обладает аспектами-свойствами *Height*, *Width*, *Length*, и все они использованы для описания свойств объекта *DCDC Converter BSC624-12V*.



При этом сам аспект *Geometry* в контексте объекта обладает ролевым именем *outer*, то есть он использован для описания внешней геометрии этого объекта.



Рисунок 4 – Описание объекта *DCDC Converter BSC624-12V* в системе Knowledge Net

На рисунке 5 представлена полная иерархия аспекта *Electric characteristic*. Для каждого аспекта указан тип и домен значений. Для аспектов-свойств в квадратных скобках также указана мощность. Мощность *Group* означает, что аспект по умолчанию может не обладать значением и выступает в качестве объединяющей характеристики для своих дочерних свойств.

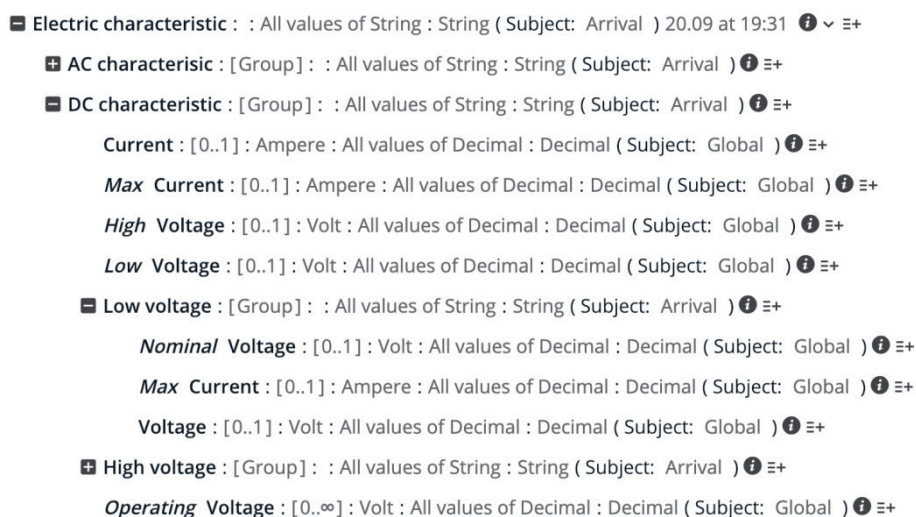


Рисунок 5 – Свойства аспекта *Electric characteristic*

Отметим, что при использовании некоторого аспекта со свойствами для описания объекта не все дочерние свойства аспекта обязательно должны быть заполнены. Например, для объекта *DCDC Converter BSC624-12V* не использован аспект *AC characteristic* и его свойства.



Аспект *Function core type* (см. рисунок 6) представляет собой пример аспекта с доменом значений *Словарь*. Этот словарь содержит иерархию классификации объектов по функциональному назначению согласно стандарту [13].

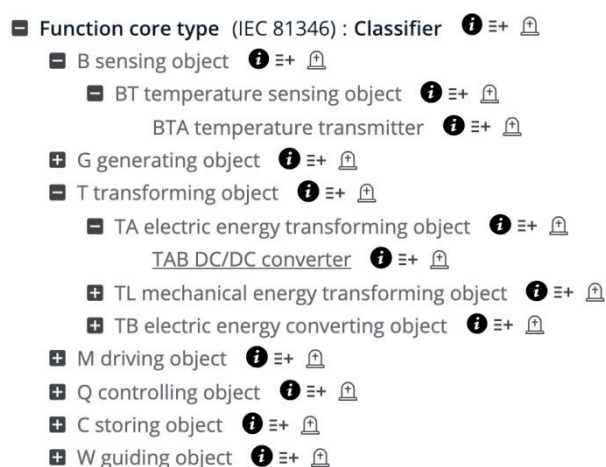


Рисунок 6 – Фрагмент словаря аспекта *Function core type*

Разработанный программный прототип положительно оценен специалистами в области формализации знаний.

## Заключение

Представлена семантическая модель накопления, представления и использования знаний и данных Knowledge Net и поддерживающий её программный прототип, который разрабатывается в концепции открытого программного кода. Структура модели Knowledge Net позволит организациям самостоятельно или с привлечением экспертов создавать графы знаний, которые могут быть положены в основу модели ЕИП предприятия, являющегося ключевой компонентой интеллектуального предприятия нового поколения, развиваемой в рамках концепции Индустрия 4.0.

Основными отличительными особенностями разрабатываемой модели Knowledge Net является поддержка многоаспектности, возможность расширения сети понятий и сущностей вместе с развитием предприятия и динамическое формирование структуры аспектов и их свойств. В основу модели Knowledge Net положена графовая модель данных. Программное обеспечение позволяет ввести семантические описания объектов в систему и использовать полученный в результате граф как основу для представления ЕИП предприятия. В последующих релизах программного обеспечения предполагается значительно расширить функциональные возможности системы Knowledge Net.

## Благодарности

Работа выполнена при финансовой поддержке компании ООО «Информационные крылья» в сотрудничестве с учёными Санкт-Петербургского института информатики и автоматизации РАН и Высшей школы менеджмента СПбГУ.

## Список источников

- [1] Государственная программа “Цифровая экономика Российской Федерации”. Утверждена распоряжением Правительства Российской Федерации от 28 июля 2017 г. №1632-р.

- [2] **Hanley, T.** The Industry 4.0 paradox. Executive summary / T. Hanley, A. Daecher, M. Cotteleer, B. Sniderman - <https://www2.deloitte.com/insights/us/en/focus/industry-4-0/challenges-on-path-to-digital-transformation/summary.html>.
- [3] **Adolphs, P.** Status report - Reference architecture model Industrie 4.0 (rami4.0) / P. Adolphs, H. Bedenbender, D. Dirzus, et al. – VDI, 2015. - [https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference\\_architectural\\_model\\_industrie\\_4.0\\_rami\\_4.0.pdf](https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf).
- [4] **Grangel-González, I.** Towards a Semantic Administrative Shell for Industry 4.0 Components / I. Grangel-González, L. Halilaj, G. Coskun, S. Auer, D. Collarana, M. Hofmeister // Proc. of IEEE Tenth International Conference on Semantic Computing (ICSC). – 2016. - P. 230-237.
- [5] **Kharlamov, E.** Capturing Industrial Information Models with Ontologies and Constraints / E. Kharlamov, B. Cuenca Grau, E. Jiménez-Ruiz, S. Lamparter, G. Mehdi, M. Ringsquandl, Y. Nenov, S. Grimm, M. Roshchin, I. Horrocks // Proc. of International Semantic Web Conference. – 2016. - P.325-343.
- [6] **Hauptert J., Schneider M.** The object memory server for semantic product memories. SemProM, Springer, 2013. - P.175–189.
- [7] **Kharlamov, E.** Optique: Ontology-Based Data Access Platform / E. Kharlamov, E. Jiménez-Ruiz, C. Pinkel, M. Rezk, M.G. Skjæveland, A. Soyly, G. Xiao, D. Zheleznyakov, M. Giese, I. Horrocks, A. Waaler // Proc. of ISWC (P&D). - 2015. - [http://ceur-ws.org/Vol-1486/paper\\_24.pdf](http://ceur-ws.org/Vol-1486/paper_24.pdf).
- [8] **Максимов, Н.В.** Семантическое ядро цифровой платформы / Н.В. Максимов, О.Л. Голицина, М.Г. Ганченкова, Д.В. Санатов, А.В. Разумов // Онтология проектирования. - 2018. - № 3 (29). - С. 412-426. – DOI: 10.18287/2223-9537-2018-8-3-412-426.
- [9] **Голенков, В.В.** Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 1: Принципы создания / В.В. Голенков, Н.А. Гулякина // Онтология проектирования. – 2014. - № 1(11). - С.42–64.
- [10] **Голенков, В.В.** Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 2: Унифицированные модели проектирования / В.В. Голенков, Н.А. Гулякина // Онтология проектирования. – 2014. - № 4(14). - С.34–53.
- [11] **Linstedt, D.** A short intro to #datavault 2.0 – <https://danlinstedt.com/allposts/datavaultcat/a-short-intro-to-datavault-2-0>.
- [12] Философский энциклопедический словарь (2010). – <http://philosophy.niv.ru/doc/dictionary/philosophy/index.htm>.
- [13] IEC 81346-2:2009. Industrial systems, installations and equipment and industrial products - Structuring principles and reference designations - Part 2: Classification of objects and codes for classes.
- [14] **Renzo A.** Survey of graph database models / A. Renz, C. Gutierrez // ACM Computing Surveys. – 2008. - №40 (1). - P.1–39.
- [15] OrientDB Community Edition. – <https://orientdb.org>.
- [16] Source Code of the Knowledge Net. – <https://github.com/InfoWings/Knowledge-Net/>.
- [17] ОК 015-94 (МК 002-97). Общероссийский классификатор единиц измерения (утв. Постановлением Госстандарта России от 26.12.1994 N 366).
- [18] **Kudryavtsev, D.** Mind Mapping and Spreadsheets in Collaborative Design of Knowledge Graphs / D. Kudryavtsev, T. Gavrilova, I. Leshcheva, A. Begler, M. Kubelskiy, O. Tushkanova // Joint Proceedings of the BIR 2018 Forum, Workshops and Doctoral Consortium, September 24-26 2018, Stockholm, Sweden. P. 82-93.

## KNOWLEDGE NET: MODEL AND SYSTEM FOR ACCUMULATION, REPRESENTATION AND USE OF KNOWLEDGE AND DATA

O.N. Tushkanova<sup>1,2</sup>, V.V. Samoylov<sup>3</sup>

<sup>1</sup> St. Petersburg Institute for Informatics and Automation of RAS, St. Petersburg, Russia

<sup>2</sup> Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia  
tushkanova.on@gmail.com

<sup>3</sup> Info Wings LLC, St. Petersburg, Russia  
vsamoilov@infowings.ru

### Abstract

The paper describes the Knowledge Net data and knowledge accumulation model developed by the authors for the company Info Wings LLC, proposed as a model for formalizing, presenting and using of knowledge and data for integrated information space of an intelligent enterprise, e.g. within the framework of Industry 4.0. The Knowledge Net is based on a graph data model and is a part of the ongoing application development of a digital enterprise platform. The code of a software system that supports the Knowledge Net model is open sourced. The paper provides an example of using a software prototype to describe objects of a manufacturing enterprise. A key feature of Knowledge Net, which distinguishes development from well-known counterparts, is the combination of properties such as support for multi-aspect descriptions of subjects and objects of an enterprise, the possibility of expanding the network of concepts and entities along with the development of the enterprise, dynamic formation of the structure of aspects and their properties, multiplicity of possible typifications presented by the Knowledge Net model, as well as support for the versioning knowledge and data models.

**Key words:** unified information space, knowledge model, knowledge representation, aspect, ontology, knowledge graph, graph data model.

**Citation:** Tushkanova ON, Samoylov VV. Knowledge Net: model and system for accumulation, representation and use of knowledge and data [In Russian]. *Ontology of designing*. 2019; 9(1): 117-131. - DOI: 10.18287/2223-9537-2019-9-1-117-131.

### Acknowledgment

The research was done in the interests and with the financial support of the company Info Wings LLC in collaboration with scientists from the St. Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences and the Graduate School of Management at St. Petersburg State University.

### References

- [1] State program “Digital economy of the Russian Federation” [In Russian]. Approved by the order of the Government of the Russian Federation dated July 28, 2017. №1632-r.
- [2] Hanley T, Daecher A, Cotteleur M, Sniderman B. The Industry 4.0 paradox. Executive summary. <https://www2.deloitte.com/insights/us/en/focus/industry-4-0/challenges-on-path-to-digital-transformation/summary.html>.
- [3] Adolfs P, Bedenbender H, Dirzus D. et al. Status Report - Reference architecture model Industrie 4.0 (rami4.0), 2015. - [https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference\\_architectural\\_model\\_industrie\\_4.0\\_rami\\_4.0.pdf](https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf).
- [4] Grangel-González I, Halilaj L, Coskun G, Auer S, Collarana D, Hofmeister M. Towards a Semantic Administrative Shell for Industry 4.0 Components. Proc. of IEEE Tenth International Conference on Semantic Computing (ICSC); 2016: 230-237.
- [5] Kharlamov E, Cuenca Grau B, Jiménez-Ruiz E, Lamparter S, Mehdi G, Ringsquandl M, Nenov Y, Grimm S, Roshchin M, Horrocks I. Capturing Industrial Information Models with Ontologies and Constraints. International Semantic Web Conference; 2016: 325-343.
- [6] Hauptert J, Schneider M. The object memory server for semantic product memories. SemProM, Springer; 2013: 175-189.

- [7] **Kharlamov E, Jiménez-Ruiz E, Pinkel C, Rezk M, Skjæveland MG, Soylu A, Xiao G, Zheleznyakov D, Giese M, Horrocks I, Waaler A.** Optique: Ontology-Based Data Access Platform. Proc. of ISWC (P&D); 2015. - [http://ceur-ws.org/Vol-1486/paper\\_24.pdf](http://ceur-ws.org/Vol-1486/paper_24.pdf).
- [8] **Maksimov NV, Golitsina OL, Ganchenkova MG, Sanatov DV, Razumov AV.** Semantic core of the digital platform [In Russian]. *Ontology of designing*. 2018; 3 (29): 412-426. – DOI: 10.18287/2223-9537-2018-8-3-412-426.
- [9] **Golenkov VV, Gulyakin NA.** The project is an open semantic technology component design of intelligent systems. Part 1: Creation principles [In Russian]. *Ontology of designing*. 2014; 1 (11): 42–64.
- [10] **Golenkov VV, Gulyakin NA.** The project of an open semantic technology component design of intelligent systems. Part 2: Unified design Models [In Russian]. *Ontology of designing*. 2014; 4(14): 34–53.
- [11] **Linstedt D.** A short intro to #datavault 2.0. - <https://danlinstedt.com/allposts/datavaultcat/a-short-intro-to-datavault-2-0>.
- [12] Philosophical Encyclopedic Dictionary (2010). [In Russian]. - <http://philosophy.niv.ru/doc/dictionary/philosophy/index.htm>.
- [13] IEC 81346-2:2009. Industrial systems, installations and equipment and industrial products -- Structuring principles and reference designations - Part 2: Classification of objects and codes for classes.
- [14] **Renzo A., Gutierrez C.** Survey of graph database models. *ACM Computing Surveys* 2018; 40 (1): 1–39.
- [15] OrientDB Community Edition. - <https://orientdb.org>.
- [16] Source Code of the Knowledge Net. - <https://github.com/InfoWings/Knowledge-Net>.
- [17] OK 015-94 (MK 002-97). All-Russian classifier of measurement units [In Russian]. (app. Resolution of state standard of Russia of 26.12.1994 N 366).
- [18] **Kudryavtsev D., Gavrilova T., Leshcheva I., Begler A., Kubelskiy M., Tushkanova O.** Mind Mapping and Spreadsheets in Collaborative Design of Knowledge Graphs Joint Proceedings of the BIR 2018 Forum, Workshops and Doctoral Consortium, September 24-26 2018, Stockholm, Sweden. - P.82-93.

## Сведения об авторах



**Тушканова Ольга Николаевна**, 1988 г. рождения. Окончила Южный федеральный университет в 2011 г. Научный сотрудник Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), доцент Санкт-Петербургского политехнического университета Петра Великого. Кандидат технических наук (2017). Количество печатных работ 28. Область научных интересов: интеллектуальный анализ данных, машинное обучение, большие данные, онтологии, извлечение знаний, семантические технологии.

**Olga Nikolaevna Tushkanova** (b. 1988). Researcher in St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences, St. Petersburg, assistant professor in Peter the Great St. Petersburg Polytechnic University, St. Petersburg. Received PhD in computer science from St. Petersburg Institute for Informatics and Automation of Russian Academy of

Sciences in 2017. Received MS degree in system analysis from Southern Federal University in Rostov-on-Don (2011). Main publications: 28. Current scientific interests: data mining, machine learning, Big Data, ontologies, knowledge engineering, semantic technologies.



**Самойлов Владимир Владимирович**, 1971 г. рождения. Окончил Высшее военное-морское инженерное училище им. Ф.Э. Дзержинского в 1993 г. Ведущий аналитик ООО “Информационные крылья”. В списке научных трудов более 30 работ в области много-агентных приложений, систем поддержки принятия решений, онтологий, механизмов машинного обучения и объединения гетерогенных данных из различных источников.

**Vladimir Vladimirovich Samoylov** (b. 1971) graduated from Higher Naval Engineering School named after F.E. Dzerzhinsky in 1993. Leading Analyst at Info Wings LLC. The list of scientific papers includes more than 30 works in the field of multi-agent applications, decision support systems, ontologies, machine learning mechanisms and combining heterogeneous data from various sources.

УДК 004.032.26, 004.822.2

## ОНТОЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ ГИБРИДНЫХ СЕМАНТИЧЕСКИ СОВМЕСТИМЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ НА ОСНОВЕ СМЫСЛОВОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

В.В. Голенков<sup>1,a</sup>, Н.А. Гулякина<sup>1,b</sup>, И.Т. Давыденко<sup>1,c</sup>, Д.В. Шункевич<sup>1,d</sup>, А.П. Еремеев<sup>2,e</sup>

<sup>1</sup>Белорусский государственный университет информатики и радиоэлектроники, Минск, Республика Беларусь

<sup>2</sup>Национальный исследовательский университет «МЭИ» (Московский энергетический институт),  
Москва, Россия

<sup>a</sup>golen@bsuir.by, <sup>b</sup>guliakina@bsuir.by, <sup>c</sup>ir.davydenko@gmail.com, <sup>d</sup>shunkevichdv@gmail.com, <sup>e</sup>YeremeevAP@mpei.ru

### Аннотация

Работа посвящена проблеме обеспечения семантической совместимости интеллектуальных систем. Показано, что обеспечение совместимости интеллектуальных систем и разработка соответствующих стандартов является ключевым направлением развития технологий проектирования интеллектуальных систем. Формально уточнено понятие смыслового представления информации в памяти интеллектуальной системы, которое обеспечивает однозначность представления информации с использованием заданного набора понятий. Показана возможность автоматической интеграции знаний в рамках смыслового представления знаний, которая сводится к склеиванию синонимичных знаков. Показана возможность автоматической интеграции различных моделей обработки знаний, если эти модели представляют собой коллективы агентов, ориентированных на обработку знаний, представленных в памяти интеллектуальных систем в смысловой форме, и взаимодействующих между собой через указанную память. Предложена Технология OSTIS, ориентированная на разработку семантических компьютерных систем. Предложена концепция Экосистемы OSTIS, представляющей собой коллектив взаимодействующих интеллектуальных систем, построенных по Технологии OSTIS и поддерживающих эволюцию и совместимость интеллектуальных систем в ходе их эксплуатации в рамках данной экосистемы. Рассмотрены примеры использования Технологии OSTIS при разработке прикладных интеллектуальных систем.

**Ключевые слова:** интеллектуальная система, язык представления знаний, модель интеллектуальной системы, компонентное проектирование, совместимость систем.

**Цитирование:** Голенков, В.В. Онтологическое проектирование гибридных семантически совместимых интеллектуальных систем на основе смыслового представления знаний / В.В. Голенков, Н.А. Гулякина, И.Т. Давыденко, Д.В. Шункевич, А.П. Еремеев // Онтология проектирования. – 2019. – Т. 9, №1(31). – С.132-151. – DOI: 10.18287/2223-9537-2019-9-1-132-151.

### Введение

До настоящего времени традиционные информационные технологии (ИТ) и технологии искусственного интеллекта (ИИ) развивались независимо друг от друга. Сейчас настало время фундаментального переосмысления опыта использования и эволюции традиционных ИТ и их интеграции с технологиями ИИ.

В данной работе рассмотрена ключевая на текущий момент проблема развития ИТ в целом и технологий ИИ в частности - проблема обеспечения информационной совместимости компьютерных систем (КС) и в том числе интеллектуальных систем (ИС).

Актуальность решения этой проблемы обусловлена тем, что информационная совместимость КС повысит уровень их обучаемости благодаря более эффективному восприятию опыта (знаний и навыков) от других КС. При этом появится возможность:



- расширить многообразие используемых в КС знаний и навыков без необходимости разработки специальных средств их согласования [1-3];
- создания коллективов КС, использующих универсальные принципы организации взаимодействия между КС на содержательном уровне;
- автоматизации процесса поддержки и восстановления информационной совместимости КС не только с другими КС, но и с их пользователями;
- сократить сроки разработки новых КС с помощью постоянно расширяемой библиотеки многократно используемых компонентов КС [4-6].

## 1 Предлагаемый подход к дальнейшей эволюции КС и технологий

Анализ проблем эволюции КС показывает, что несовместимость, дублирование и субъективизм используемых информационных ресурсов (ИР) и моделей их обработки присутствует в развитии: традиционных КС; технологий ИИ [7]; методов и средств информатизации научной и инженерной деятельности [8, 9].

Обеспечение совместимости ИР и моделей их обработки имеет разные аспекты:

- совместимость между различными компонентами КС, а также между целостными КС, входящими в коллективы КС;
- совместимость между различными КС и их пользователями;
- междисциплинарная совместимость, конвергенция различных областей знаний [10, 11];
- методы и средства постоянного мониторинга и восстановления совместимости в условиях интенсивной эволюции КС и их пользователей.

### 1.1. Направления эволюции КС

В эволюции КС можно выделить два направления. Первое направление – это расширение множества и многообразия задач, решаемых КС [1, 2, 12]. Второе – это повышение уровня их обучаемости и, как следствие, темпов их эволюции. Обучаемость КС определяется: *трудоёмкостью*, темпами расширения и совершенствования используемых знаний и навыков [13, 14]; *уровнем ограничений*, накладываемых на вид приобретаемых и используемых знаний и навыков.

В свою очередь, трудоёмкость и темпы расширения и совершенствования знаний и навыков КС определяются [15]:

- *гибкостью* – многообразием и трудоёмкостью возможных изменений, вносимых в систему в процессе пополнения системы новыми знаниями и навыками;
- *стратифицированностью* – разделением системы на достаточно независимые друг от друга уровни иерархии;
- *рефлексивностью* — способностью анализировать собственное состояние и свою деятельность;
- *гибридностью* – способностью приобретать и использовать широкое многообразие знаний и навыков;
- *уровнем самообучаемости* – уровнем активности, самостоятельности, целеустремленности в процессе обучения без учителя;
- *совместимостью* – трудоёмкостью интеграции с другими КС и с компонентами КС;
- способностью к *постоянному мониторингу* и поддержке своей совместимости как с другими КС, так и со своими пользователями в условиях эволюции КС и их пользователей.

Совместимость КС может рассматриваться в аспекте *глубокой интеграции* КС, что предполагает преобразование нескольких КС в одну целостную КС, и в аспекте преобразования нескольких КС в *коллектив взаимодействующих КС*.

Совместимость КС определяется:

- совместимостью различного вида информации (знаний), хранимой в памяти КС;
- совместимостью различных моделей решения задач;
- совместимостью встроенных подсистем, входящих в состав КС;
- совместимостью внешней информации с информацией, хранимой в памяти КС;
- коммуникационной совместимостью с пользователями и с другими КС.

## 1.2. Суть предлагаемого подхода

Суть предлагаемого подхода к решению проблем эволюции КС заключается в объединении указанных направлений эволюции КС и в трактовке проблемы обеспечения совместимости различных КС как ключевой проблемы эволюции КС, без решения которой невозможно:

- создавать коллективы КС, способные координировать свои действия при кооперативном расширении сложных задач;
- создавать гибридные КС, способные решать сложные задачи с использованием сочетания разных видов знаний и моделей решения задач;
- использовать компонентную методику проектирования КС.

Совместимость КС и их компонентов, а также совместимости КС с пользователями имеют разные трактовки этого термина:

- совместимость как один из факторов обучаемости, как *способность* к быстрому повышению уровня согласованности (интеграции, взаимопонимания);
- совместимость как характеристика достигнутого уровня согласованности.

## 1.3. Унификация КС

Главным фактором обеспечения совместимости различных видов знаний, различных моделей решения задач и различных КС в целом является унификация:

- представления информации в КС [16, 17];
  - принципов организации обработки информации в КС.
- Унификация представления информации в КС предполагает:
- *синтаксическую* унификацию информации – унификацию формы представления (кодирования) информации:
    - кодирование информации в памяти КС (внутреннее представление информации);
    - внешнее представление информации, обеспечивающее однозначность интерпретации информации разными пользователями и разными КС;
  - *семантическую* унификацию информации, в основе которой лежит согласование и точная спецификация используемых понятий с помощью системы формальных онтологий.

## 2 Смысловое представление информации в памяти КС

Объективным ориентиром для представления информации в памяти КС является формализация смысла представляемой информации [18-20].

Принципы смыслового представления информации основаны на:

- различиях внутреннего языка КС, используемого для хранения информации в памяти компьютера, и внешних языков КС, используемых для общения КС с пользователями и другими КС;
- упрощении синтаксиса внутреннего языка КС при обеспечении универсальности путём исключения из внутреннего универсального языка средств, обеспечивающих коммуникационную функцию языка (например, союзы, предлоги, склонения и другие).

## 2.1. Унификация внутреннего представления информации в КС

*Смысл* – это абстрактная знаковая конструкция, принадлежащая внутреннему языку КС, являющаяся инвариантом максимального класса семантически эквивалентных знаковых конструкций, принадлежащих разным языкам, и удовлетворяющая следующим требованиям:

- универсальность - возможность представления любой информации;
- отсутствие синонимии знаков;
- отсутствие дублирования информации в виде семантически эквивалентных текстов;
- отсутствие омонимичных знаков (в том числе местоимений);
- отсутствие у знаков внутренней структуры (атомарный характер знаков);
- отсутствие склонений, спряжений;
- отсутствие фрагментов знаковой конструкции, не являющихся знаками (разделителей, ограничителей, и т.д.);
- наличие знаков связей, компонентами которых могут быть любые знаки, с которыми знаки связей связываются синтаксически задаваемыми отношениями инцидентности.

Следствием указанных принципов смыслового представления информации в памяти КС является то, что знаки сущностей, входящие в смысловое представление информации, не являются именами (терминами) и, следовательно, не привязаны ни к какому естественному языку и не зависят от пристрастий различных авторов. Эти же принципы приводят к нелинейным знаковым конструкциям (к графовым структурам), что усложняет реализацию памяти КС, но существенно упрощает её логическую организацию (в частности, ассоциативный доступ). Нелинейность смыслового представления информации обусловлена тем, что:

- каждая описываемая сущность, т.е. сущность, имеющая соответствующий ей знак, может иметь неограниченное число связей с другими сущностями;
- каждая описываемая сущность в смысловом представлении имеет единственный знак, т.к. синонимия знаков здесь запрещена;
- все связи между описываемыми сущностями описываются связями между знаками этих сущностей.

Суть универсального смыслового представления информации можно сформулировать в виде следующих положений.

- Смысловая знаковая конструкция трактуется как множество знаков, взаимно-однозначно обозначающих различные сущности и множество связей между этими знаками.
- Каждая связь между знаками трактуется как множество знаков, связываемых этой связью, и как описание соответствующей связи, которая связывает денотаты указанных знаков или денотаты одних знаков непосредственно с другими знаками, или сами эти знаки.
- Денотатами знаков могут быть конкретные и произвольные, реальные и абстрактные, «внешние» и «внутренние» сущности, являющиеся множествами знаков, входящих в состав той же самой знаковой конструкции.

Ключевым свойством языка смыслового представления информации является однозначность представления информации в памяти каждой КС, т.е. отсутствие семантически эквивалентных знаковых конструкций, принадлежащих смысловому языку и хранимых в одной смысловой памяти. При этом логическая эквивалентность таких знаковых конструкций допускается и используется, например, для компактного представления некоторых знаний, хранимых в смысловой памяти.

Логически эквивалентные знаковые конструкции – это представление одного и того же знания, но с помощью разных наборов понятий. В отличие от этого семантически эквивалентные знаковые конструкции – это представление одного и того же знания с помощью одних и тех же понятий. Многообразие возможных вариантов представления одних и тех же знаний в памяти КС усложняет решение задач. Поэтому, исключив семантическую эквива-

лентность в смысловой памяти, необходимо стремиться к минимизации логической эквивалентности. Для этого необходимо грамотное построение системы используемых понятий в виде иерархической системы формальных онтологий [20, 21].

Важным этапом создания способа смыслового кодирования знаний был Универсальный Семантический Код (УСК) [18]. В качестве стандарта универсального смыслового представления информации в памяти КС предложен *sc-код* (*Semantic Computer Code*) [15]. В отличие от УСК В.В. Мартынова он специально ориентирован на кодирование информации в памяти компьютеров нового поколения, ориентированных на разработку семантически совместимых ИС и названных *семантическими ассоциативными компьютерами*. Таким образом, сутью предлагаемого смыслового представления информации является ориентация на формальную модель памяти нефоннеймановского компьютера, предназначенного для реализации ИС, использующих смысловое представление информации.

## 2.2. Синтаксис *sc*-кода

Универсальность *sc*-кода позволяет с его помощью описывать любые объекты. Таким объектом может быть любой язык коммуникации с пользователями (в том числе и естественный язык), а также сам *sc-код*. Синтаксис *sc*-кода представляется в виде соответствующей формальной онтологии, которая подробно рассмотрена в работе [22].

Семейство введённых классов объектов исследования трактуется как *Алфавит sc-кода*. Но, в отличие от других языков, классы синтаксически выделяемых элементарных фрагментов текстов *sc*-кода могут пересекаться. Так, например, *sc-элемент* может одновременно принадлежать и классу *sc-элементов*, и классу *sc-узлов*, а также одновременно принадлежать и классу *sc-элементов*, и классу *sc-коннекторов*, и классу *sc-дуг*, и классу *базовых sc-дуг*. Такая особенность *Алфавита sc-кода* даёт возможность строить синтаксически корректные *sc-тексты* (тексты *sc*-кода) в условиях неполноты исходных знаний о некоторых *sc-элементах*.

Отметим некоторые синтаксические особенности *sc*-кода.

- Тексты *sc*-кода являются *абстрактными* в том смысле, что они абстрагируются от конкретного варианта их кодирования в памяти КС. Кодирование текстов, в частности, зависит от варианта технической реализации памяти КС. Так, например, актуальной является аппаратная реализация ассоциативной нелинейной памяти, в которой реализуется структурная реконфигурация хранимой информации, в которой обработка информации сводится не к изменению состояния элементов памяти, а к изменению конфигурации связей между ними.
- Тексты *sc*-кода являются структурами *графоподобного вида*. Все графовые структуры легко представимы в *sc-коде* (неориентированные и ориентированные графы, мультиграфы, псевдографы, гиперграфы, сети и др.). В *sc-коде* представимы и связи между связями, связи между целыми структурами и многое другое. Таким образом, теория графов при соответствующем её расширении является основой описания синтаксиса *sc*-кода.

## 2.3. Семантика *sc*-кода

Простота синтаксиса *sc*-кода обусловлена следующими семантическими свойствами *sc*-текстов:

- все *sc-элементы*, то есть элементарные (атомарные) фрагменты *sc*-текстов, являются знаками различных описываемых сущностей, при этом каждая сущность, описываемая в тексте *sc*-кода, должна быть представлена своим знаком;
- никаких других знаков, кроме *sc-элементов*, *sc*-тексты не содержат;
- любая сущность может быть описана *sc-текстом*;

- все синтаксически выделяемые классы *sc-элементов* (т.е. все элементы *Алфавита sc-кода*) имеют чёткую семантическую интерпретацию – являются классами *sc-элементов*, каждый из которых обозначает сущность, имеющую общие одинаковые свойства со всеми другими сущностями, обозначаемыми другими *sc-элементами* этого же класса.

Денотационная семантика любой знаковой конструкции – это соответствие между множеством всех знаков, входящих в знаковую конструкцию, и множеством денотатов этих знаков, а также между множеством всех семантически значимых связей, связывающих знаки, и множеством соответствующих им связей, связывающих либо денотаты всех указанных знаков, либо денотаты некоторых из указанных знаков непосредственно с остальными знаками.

Формальное описание денотационной семантики *sc-кода* средствами *sc-кода* осуществляется в виде системы *формальных онтологий* верхнего уровня, представленных в виде текстов *sc-кода*. Указанная система формальных онтологий подробно рассмотрена в работе [22], а также представлена в базе знаний (БЗ) *Метасистемы IMS.ostis* [23].

### 3 Уточнение понятия семантической совместимости

Важнейшим этапом эволюции любой технологии является переход к компонентному проектированию на основе постоянно пополняемой библиотеки многократно используемых компонентов.

Основной проблемой для реализации компонентного проектирования являются:

- унификация компонентов по форме;
- разработка стандартов, обеспечивающих совместимость этих компонентов.

Для реализации компонентного проектирования БЗ требуется:

- универсальный язык представления знаний;
- универсальная процедура интеграции знаний в рамках указанного языка;
- разработка стандарта, обеспечивающего *семантическую совместимость* интегрируемых знаний (таким стандартом является согласованная система используемых понятий).

Для смыслового представления знаний нужны смысловые семантические координаты, роль которых выполняет используемая система понятий, которая описывается иерархической системой семантически связанных между собой *онтологий*.

Знания необходимо привести к общему «семантическому знаменателю», чем является постоянно уточняемая система понятий, специфицируемая в виде объединённой онтологии. Эта объединённая онтология стратифицируется на частные онтологии.

Один из принципов семантической совместимости новой информации с БЗ, в которую эта информация погружается, можно сформулировать следующим образом: все знаки, являющиеся новыми для воспринимающей БЗ, должны быть специфицированы через понятия, известные БЗ. Стандарт смыслового представления информации (*sc-код*) даёт возможность повысить уровень совместимости КС и формально уточнить понятие интеграции КС и их компонентов.

#### 3.1. Уточнение процесса понимания на основе смыслового представления информации

Очевидно, что формализация смыслового представления информации в памяти КС существенно упрощает уточнение того, как происходит процесс понимания новой информации, поступающей на вход КС, либо генерируемой в процессе обработки информации. Этот процесс можно разбить на три этапа:



- *трансляция* информации с некоторого внешнего языка на внутренний смысловой язык *sc-код* (этап отсутствует, если новая информация не вводится извне, а непосредственно генерируется в памяти КС);
- *погружение* новой информации, представленной в виде *sc-текста*, в текущее состояние ИР, хранимого в памяти КС и представленного также в виде *sc-текста*;
- *выравнивание* (согласование) понятий, используемых в новой вводимой извне или сгенерированной информационной конструкции, с понятиями, используемыми в текущем состоянии хранимого в памяти КС ИР.

Рассмотрим каждый из перечисленных этапов подробнее.

*Трансляция* информации с внешнего языка в *sc-код* упрощается, поскольку:

- средствами *sc-кода* можно описать *синтаксис* внешнего языка, т.к. универсальность *sc-кода* позволяет с его помощью и с любой степенью детализации описывать любые объекты, в том числе и такие сложные системы внешней среды КС, как внешние языки;
- процесс *синтаксического анализа* исходного текста внешнего языка можно выполнить путём манипуляции текстами *sc-кода* и в результате получить описание структуры исходного текста, имеющее достаточную полноту (детализацию) для последующей генерации семантически эквивалентного ему текста *sc-кода*;
- средствами *sc-кода* можно описать *семантику* внешнего языка, трактуя её как описание свойств морфизмов между *sc-текстами*, описывающими синтаксическую структуру исходных внешних текстов, и *sc-текстами*, которые семантически эквивалентны этим исходным текстам;
- процесс генерации *sc-текста*, *семантически эквивалентного* исходному внешнему тексту, также можно выполнить путём манипуляции *sc-текстами*.

*Погружение* (интеграция) нового сгенерированного *sc-текста* в текущее состояние *sc-текста* (например, в состав БЗ, представленной в *sc-коде*) сводится к *склеиванию* некоторых *sc-элементов* нового *sc-текста* с синонимичными им *sc-элементами*, входящими в состав заданного *sc-текста*. Задача погружения нового *sc-текста* сводится к задаче построения множества пар синонимичных *sc-элементов*, один из которых входит в состав нового погружаемого *sc-текста*, а второй – в состав заданного *sc-текста*. Установление пар синонимичных *sc-элементов* на начальном этапе осуществляется путём поиска пар *sc-элементов*, у которых совпадают согласованные внешние имена. На следующем этапе синонимичные *sc-элементы* выявляются путём логических рассуждений.

Для упрощения установления пар синонимичных *sc-элементов* некоторые высказывания о несуществовании, о существовании и единственности, о существовании заданного конечного числа структур заданного вида можно переформулировать с явным введением отношения синонимии *sc-элементов*. Так, например, вместо утверждения: «Для каждой пары точек существует единственная проходящая через них прямая» можно использовать следующую формулировку: «Если прямые  $pi$  и  $pj$  проходят через точки  $ti$  и  $tj$ , то либо  $pi = pj$ , либо  $ti = tj$ , либо  $ti \notin pi$ , либо  $ti \notin pj$ , либо  $tj \notin pi$ , либо  $tj \notin pj$ ». Формальная запись данного утверждения на языке SCg, являющемся графическим вариантом внешнего отображения текстов *sc-кода*, представлена на рисунке 1. Подробное описание примера погружения *sc-текста* в БЗ, представленную также в *sc-коде*, приведено в [15].

*Выравнивание понятий*, используемых в новом интегрируемом *sc-тексте*, с понятиями, используемыми в заданном интегрирующем *sc-тексте*, осуществляется следующим образом.

- Заданный интегрирующий *sc-текст* должен явно содержать информацию о текущем состоянии использования:
  - каждого известного понятия, используемого либо непосредственно в БЗ, либо внешними субъектами, информация от которых может поступать на вход БЗ;

- каждого внешнего знака (чаще всего термина, имени), соответствующего каждому используемому понятию, а также некоторым общеизвестным сущностям, которые не являются понятиями.
- Интегрируемый текст должен:
  - максимально возможным образом использовать согласованные понятия и соответствующие им согласованные внешние знаки (термины, имена);
  - включать в себя определения всех понятий, которые являются новыми, неизвестными в интегрирующем тексте (при этом в определении должны использоваться только те понятия, которые известны интегрирующему тексту).
- Для решения задачи *выравнивания* используемых понятий для текущего состояния БЗ и для нового вводимого в эту БЗ текста все используемые в БЗ понятия делятся на согласованные, устаревшие, устаревающие, отклонённые, вводимые новые понятия.

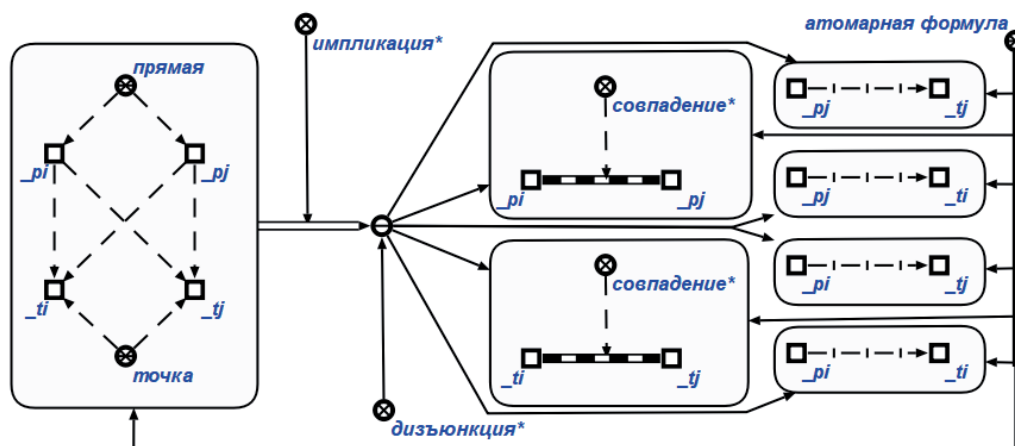


Рисунок 1 – Пример формальной записи логического утверждения

Таким образом, процесс выравнивания понятий, целью которого является сведение всех понятий, используемых в интегрируемом *sc*-тексте, к согласованным понятиям БЗ, осуществляется в условиях постоянного изменения статуса используемых понятий и постоянного увеличения числа таких понятий.

### 3.2. Унификация и совместимость различных моделей решения задач

Предлагаемый подход к повышению уровня совместимости (интегрируемости) различных *моделей решения задач* заключается в следующем [24]:

- вся информация, хранящаяся в памяти каждого *решателя задач*, представляется в форме смыслового представления этой информации (в *sc*-коде);
- решение каждой задачи осуществляется коллективом агентов, работающих над общей для них смысловой памятью и выполняющих интерпретацию хранимых в этой же памяти навыков (указанные агенты названы *sc*-агентами);
- интеграция двух разных моделей решения задач сводится:
  - к объединению памяти моделей;
  - к интеграции *sc*-текстов, хранимых в памяти моделей, путём взаимного погружения этих *sc*-текстов друг в друга;
  - к объединению множеств агентов, входящих в составы моделей.

Унификация моделей решения задач путём приведения этих моделей к виду *sc*-моделей повышает уровень совместимости этих моделей благодаря наличию прозрачной процедуры интеграции *sc*-текстов и тривиальной процедуры объединения множеств *sc*-агентов. Просто-

та процедуры объединения множеств *sc*-агентов, соответствующих разным моделям решения задач, обусловлена тем, что непосредственного взаимодействия между этими агентами нет, а инициирование каждого из них определяется самим агентом, а также текущим состоянием хранимой в памяти информации. Основными преимуществами использования *многоагентного подхода* [1, 25] являются автономность агентов и децентрализация обработки вносимых изменений. Многоагентная обработка БЗ имеет ряд недостатков:

- знания агента представляются при помощи узкоспециализированных языков, не предназначенных для представления знаний в широком смысле и онтологий в частности;
- в большинстве современных многоагентных систем взаимодействие агентов осуществляется путём обмена сообщениями непосредственно от агента к агенту;
- среда, с которой взаимодействуют агенты, уточняется отдельно разработчиком для каждой многоагентной системы, что приводит к несовместимости многоагентных систем.

Перечисленные недостатки предлагается устранять за счёт использования следующих принципов:

- коммуникацию агентов осуществлять путём спецификации действий, выполняемых агентами и направленных на решение задач;
- в роли внешней среды для агентов должна выступать общая память КС;
- спецификацию каждого агента описывать средствами языка представления знаний;
- синхронизацию деятельности агентов осуществлять на уровне выполняемых ими процессов;
- каждый информационный процесс в любой момент времени должен иметь ассоциативный доступ к необходимым фрагментам БЗ, хранящейся в общей памяти.

### 3.3. Семантическая совместимость КС

Уровень совместимости КС определяется трудоёмкостью реализации процедур интеграции знаний этих систем, а также трудоёмкостью и глубиной интеграции входящих в эти системы *решателей задач*.

Совместимые КС – это КС, для которых существует автоматически выполняемая процедура их интеграции, в рамках которой каждая исходная КС в процессе своего функционирования может свободно использовать знания, входящие в состав другой исходной КС.

Целостная КС – это решатель задач, интегрировавший несколько моделей решения задач и обладающий средствами взаимодействия с внешней средой.

Чтобы повысить уровень совместимости КС, необходимо преобразовать их к виду многоагентных систем, работающих с общей смысловой памятью, в которой информация представлена текстами *sc*-кода. Такие унифицированные КС далеко не всегда целесообразно объединять в более крупные КС. Иногда целесообразнее их объединять в *коллективы взаимодействующих КС*. Но при создании таких коллективов КС унификация и совместимость таких систем также важны. Противоречия между КС, входящими в коллектив, можно обнаруживать путём анализа виртуальной объединённой БЗ этого коллектива. Непротиворечивость указанной виртуальной БЗ можно считать одним из критериев семантической совместимости систем, входящих в соответствующий коллектив.

Ключевым отличием рассматриваемого подхода к разработке КС по сравнению с существующими [5, 7, 26] является обеспечение совместимости ИС и возможность автоматизации процесса их интеграции. Решение проблемы совместимости в таком контексте позволит обеспечить переход от современных ИС к гибридным ИС, способным наращивать и совершенствовать свои функциональные возможности. Подробно модели, методы и средства разработки гибридных БЗ на основе предложенного подхода рассмотрены в [21], модели, методы и средства разработки гибридных решателей задач - в [27].

### 3.4. Достоинства смыслового представления информации

Переход к *смысловому представлению информации* в памяти КС целесообразен по следующим причинам:

- смысловое представление информации есть объективный способ представления информации;
- в рамках смыслового представления существенно упрощается процедура интеграции знаний и погружения новых знаний в БЗ;
- упрощается процедура приведения различного вида знаний к общему виду;
- упрощается процедура интеграции различных решателей задач и КС;
- упрощается автоматизация процесса поддержки семантической совместимости для КС в условиях их постоянного совершенствования;
- на основе предложенного смыслового представления информации упрощается интеграция различных дисциплин в области ИИ.

## 4 Семантические КС и технологии

Переход к смысловому представлению информации в памяти КС фактически преобразует современные КС в *семантические КС*, которые являются этапом их эволюции, направленным на обеспечение высокого уровня обучаемости и совместимости КС.

Архитектура *семантических КС* практически совпадает с архитектурой ИС, основанных на знаниях. Отличие заключается в том, что в *семантических КС* БЗ имеет смысловое представление, а интерпретатор знаний и навыков представляет собой коллектив агентов, осуществляющих обработку БЗ.

Предлагаемая технология разработки семантических КС названа *Технологией OSTIS* (Open Semantic Technology for Intelligent Systems) [23]. В основе этой технологии лежит *sc*-код – разработанный стандарт смыслового представления информации в памяти КС. *Технология OSTIS* – это:

- *стандарт* семантических КС, обеспечивающий семантическую совместимость систем;
- *методы построения* таких КС и их совершенствования в процессе их эксплуатации;
- *средства построения* и совершенствования этих систем (языковые средства, библиотеки типовых технических решений, инструментальные средства).

Принципы, лежащие в основе *Технологии OSTIS* [22]:

- ориентация на смысловое однозначное представление знаний в виде семантических сетей, имеющих базовую теоретико-множественную интерпретацию;
- использование ассоциативной графодинамической модели памяти;
- применение агентно-ориентированной модели обработки знаний;
- реализация в виде интеллектуальной *Метасистемы IMS.ostis* [23].

Архитектура КС, разрабатываемых по *Технологии OSTIS*, чётко стратифицирована на две подсистемы:

- БЗ, которая представляет собой полную семантическую модель ИС (*sc*-модель ИС или *sc*-модель БЗ ИС);
- базовый универсальный интерпретатор семантической модели ИС, хранимой в ее памяти (интерпретатор *sc*-модели БЗ ИС).

При наличии эффективного варианта реализации интерпретатора *sc*-моделей БЗ разработка *ostis*-системы сводится к проектированию *sc*-модели БЗ разрабатываемой системы [21], которая включает в себя:

- *sc*-модель интегрированного решателя задач разрабатываемой *ostis*-системы [27];
- *sc*-модель интегрированного интерфейса разрабатываемой *ostis*-системы;

- часть *sc*-модели БЗ разрабатываемой *ostis*-системы, которая не входит ни в *sc*-модель интегрированного решателя задач *ostis*-системы, ни в *sc*-модель интегрированного интерфейса *ostis*-системы.

Проектирование решателя задач ИС заключается в проектировании знаний специального вида — навыков и спецификаций агентов, осуществляющих интерпретацию этих навыков при решении конкретных задач. Проектирование интерфейса ИС сводится к проектированию знаний, представляющих собой семантическую модель встроенной ИС, ориентированной на решение интерфейсных задач. Важнейшее место в комплексе средств разработки ИС занимает встраиваемая типовая ИС комплексной поддержки проектирования БЗ.

Реализация универсального интерпретатора *sc*-моделей КС может иметь большое число вариантов — как программно, так и аппаратно реализованных. Логическая архитектура универсального интерпретатора *sc*-моделей КС обеспечивает независимость проектируемых КС от многообразия вариантов реализации интерпретатора их моделей.

Аппаратно реализованный интерпретатор семантических моделей (*sc*-моделей) КС представляет собой *семантический ассоциативный компьютер*, который имеет ряд существенных отличий от современных традиционных компьютеров и решений в области хранения и обработки семантических сетей [28].

## 5 Экосистема OSTIS

*Технология OSTIS* ориентирована на разработку семантических КС, обладающих высоким уровнем обучаемости и высоким уровнем семантической *совместимости*. Для реализации этого необходима среда, социотехническая инфраструктура - *Экосистемой OSTIS*, которая представляет собой коллектив взаимодействующих (через сеть Интернет):

- семантических КС, построенных по стандартной технологии *OSTIS* (*ostis*-системы);
- пользователей *ostis*-систем;
- КС, не являющихся *ostis*-системами, но рассматриваемых ими в качестве дополнительных ИР или сервисов.

Задачами *Экосистемы OSTIS* являются:

- оперативное внедрение всех согласованных изменений стандарта *ostis*-систем;
- поддержка высокого уровня взаимопонимания систем, входящих в *Экосистему OSTIS*, и их пользователей;
- корпоративное решение различных сложных задач, требующих координации деятельности нескольких *ostis*-систем и пользователей.

Основное назначение *Экосистемы OSTIS* — обеспечить совместимость КС, входящих в неё как на этапе их разработки, так и в ходе их эксплуатации. *Экосистема OSTIS* является формой реализации, совершенствования и применения *Технологии OSTIS* и, следовательно, является формой создания, развития, самоорганизации семантически совместимых КС и включает в себя необходимые для этого ресурсы — информационные, технологические, кадровые, организационные, инфраструктурные.

## 6 Примеры использования Технологии OSTIS

### 6.1. Системы автоматизации рецептурного производства

Одним из важнейших направлений применения технологий ИИ является сфера автоматизации производства [31, 32]. *Технология OSTIS* активно используется для разработки систем автоматизации производства и информационного обслуживания сотрудников предпри-



ятия ОАО «Савушкин продукт» (г. Брест). Важной частью системы является формальная модель стандарта рецептурного производства ISA-88, представленная в виде совокупности sc-моделей предметных областей (ПрО) и их онтологий (рисунок 2).

*Предметная область моделей процедурного управления оборудованием рецептурного производства*  
⇒ основной идентификатор\*:

Предметная область моделей процедурного управления оборудованием рецептурного производства ...

⇐ частная предметная область\*:

Предметная область предприятий рецептурного производства

⇐ ключевой sc-элемент\*:

Раздел. Предметная область моделей процедурного управления оборудованием рецептурных производств

⇒ максимальный класс объектов исследования\*:

процедурный элемент

⇒ исследуемое отношение\*:

порядок исполнения\*

⇒ немаксимальный класс объектов исследования\*:

- процедура ячейки процесса
- процедура технологической установки
- операция
- атомарный процедурный элемент
- процедурный элемент процесса
- процедурный элемент оборудования
- процедура
- рецептурный процесс
- рецептурная производственная операция
- рецептурная фаза
- процедура процессной ячейки оборудования
- процесс оборудования
- производственная операция оборудования
- фаза оборудования

Рисунок 2 – Фрагмент формальной модели стандарта ISA-88

Формализация стандарта позволяет, с одной стороны, строить формальные модели предприятия, соответствующие данному стандарту (рисунок 3), с другой стороны – автоматизировать эволюцию описания стандарта и обеспечить информационную поддержку сотрудников предприятия. Подробно результаты применения *Технологии OSTIS* для автоматизации рецептурного производства описаны в работе [31].

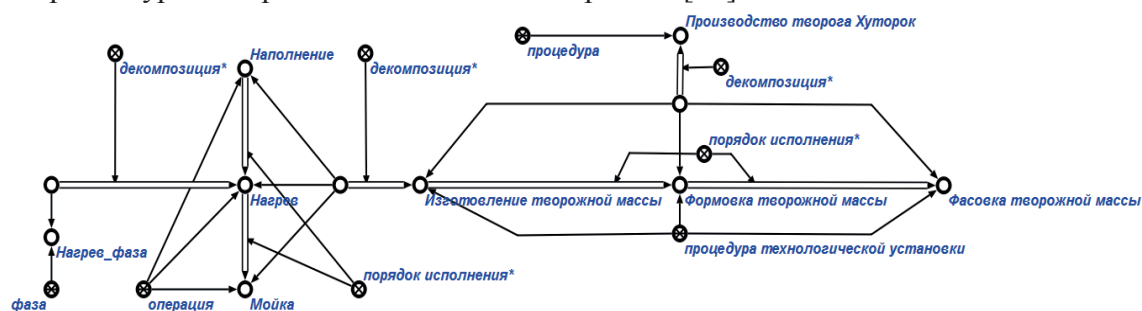


Рисунок 3 – Фрагмент модели производства [31]

## 6.2. Интеграция нейросетевых моделей с БЗ

Одним из перспективных вариантов гибридизации на основе *Технологии OSTIS* является интеграция нейросетевых моделей с БЗ [32]. Такого рода интеграция оказывается актуальной при решении задач, связанных с принятием решений на основе результатов работы систем машинного зрения. Вариант архитектуры такой системы приведён на рисунке 4. На рисунке 5 показан пример представления обученных нейронных сетей в БЗ. Такое представление позволяет системе автоматически выбирать и применять уже обученную сеть в зависимости от класса текущей задачи.



Рисунок 4 – Архитектура системы принятий решений

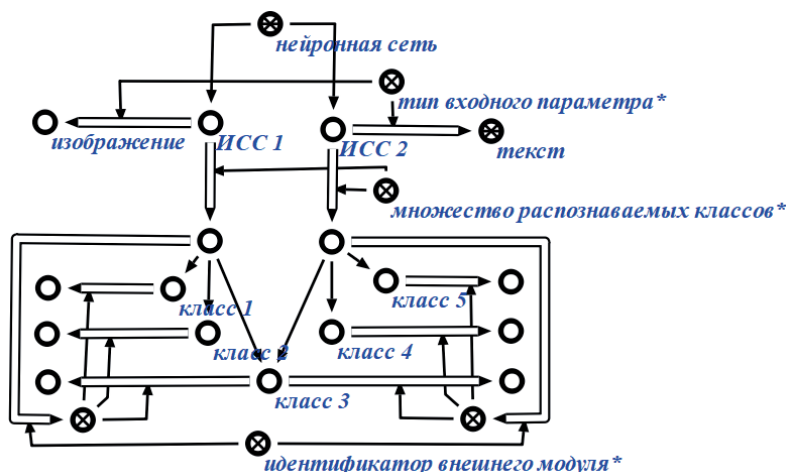


Рисунок 5 – Представление нейронной сети в базе знаний

### 6.3. Интеллектуальные обучающие системы

Одним из направлений использования *Технологии OSTIS* являются интеллектуальные обучающие системы. БЗ такой системы содержит в структурированном виде информацию о понятиях и объектах соответствующей ПрО, классах решаемых задач и способах их решения, вспомогательные материалы, такие как примеры и упражнения, и многое другое. На рисунке 6 показан верхний уровень иерархии ПрО для обучающей системы по геометрии Евклида.

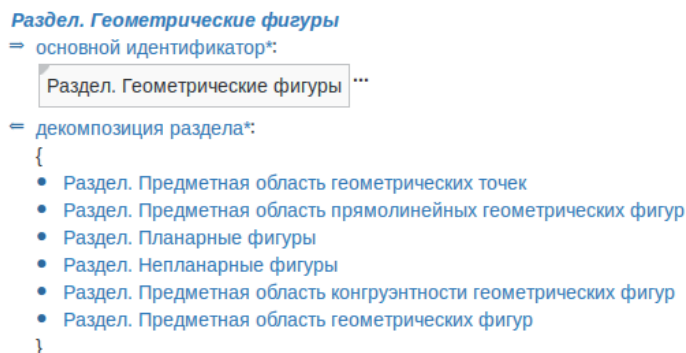


Рисунок 6 – Фрагмент структуры БЗ обучающей системы по геометрии Евклида

Система включает решатель задач и редактор геометрических чертежей, позволяющий пользователю создавать условия задач в привычном для него виде. На рисунках 7 и 8 показаны соответственно пример работы чертёжника и результат трансляции изображённого чертежа в память системы.

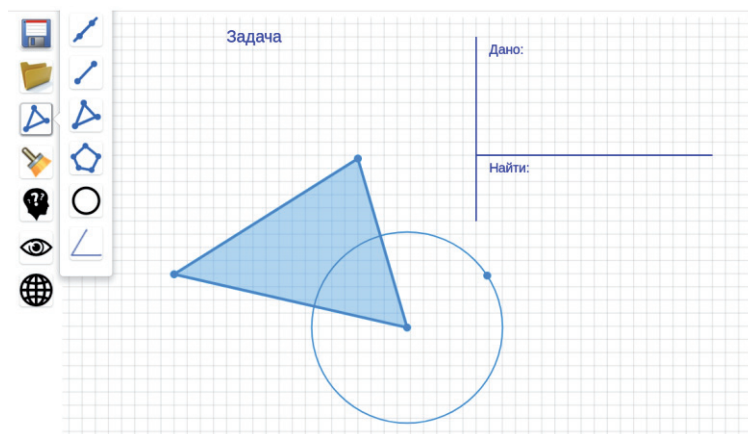


Рисунок 7 – Пример работы редактора геометрических чертежей

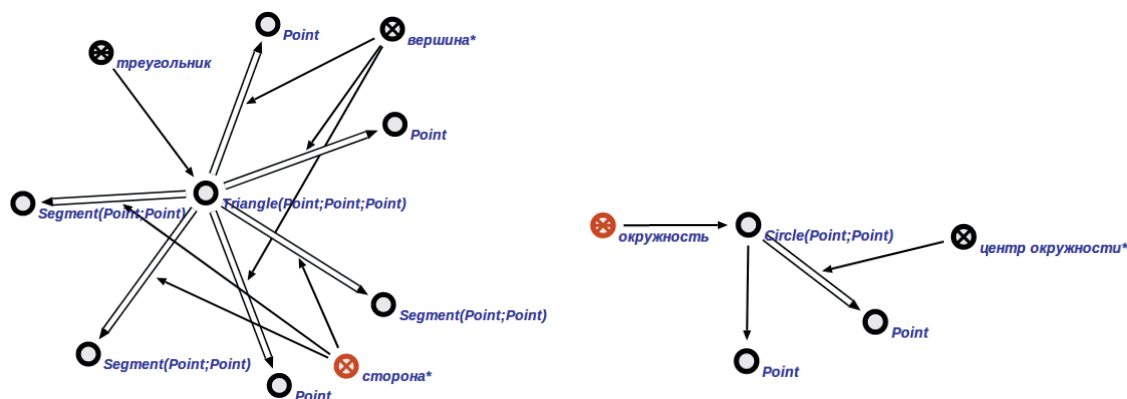


Рисунок 8 – Результат трансляции чертежа в память системы

#### 6.4. Интеллектуальные гиды

Предложенный подход к представлению и обработке информации хорошо подходит для использования в динамических ПрО, характеризующихся наличием большого числа так называемых НЕ-факторов, таких как неполнота, недостоверность, нечёткость представляемой информации. Одной из таких областей является область истории. С использованием *Технологии OSTIS* разрабатывается система-гид по городу Минску, содержащая исторические сведения о городе, его улицах, зданиях, исторических личностях и организациях. Описание учитывает динамику сущностей во времени, например, различные периоды жизни исторических деятелей, изменение названий улиц, расположения домов, организаций и т.д. Благодаря этому разрабатываемая система позволяет автоматически получать информацию об одних и тех же сущностях в различных временных срезах. На рисунке 9 показан пример описания исторической личности в БЗ системы.

Для удобства использования системы в качестве гида-экскурсовода реализован картографический интерфейс, отображающий объекты, представленные в БЗ, и их описание (рисунок 10).

**Гуттен-Чапский Кароль Эмеринович**

- ⇒ основной идентификатор\*:
  - Гуттен-Чапский Кароль Эмеринович ...
  - € Русский язык
  - Гутэн-Чапскі Кароль Эмерыкавіч ...
  - € Белорусский язык
  - Hutten-Czapsky Karol' Emierikovich ...
  - € Английский язык
- ⇒ системный идентификатор\*:
  - Hutten\_Czapsky\_Emerickovich ...
- ⇒ городской голова\*:сотрудник\*:
  - Минское городское общественное управление
- ⇒ основатель\*:
  - Минская больница для проституток
  - Минский городской ломбард
  - Минский городской приют для бездомных
- ⇒ социальная принадлежность\*:
  - титованное дворянство
  - родовое дворянство
- ⇒ место учёбы\*:
  - Здание гимназии Святой Анны
  - Здание Императорского Дерптского университета
- ⇒ супруг\*:
  - Пусловская Мария Леонтина
- ⇒ родитель\*:
  - Чапский Эмерик Август Войтех
  - Чапский Леон Богдан
  - Чапская Елизавета Карловна
- ⇒ родитель\*:
  - Гуттен-Чапская Анна Юрьевна
  - Гуттен-Чапский Эмерик Карлович
- ⇒ род\*:
  - Гуттен-Чапские
- ⇒ дата смерти\*:
  - 17 января 1904 года
- ⇒ дата рождения\*:
  - 15 мая 1860 года

⇒ дата рождения\*:
 

- 15 мая 1860 года

⇒ сотрудник\*:
 

- Министерство финансов Российской империи

⇒ второе имя\*:
 

- Кароль
- Александр

⇒ фамилия\*:
 

- Гуттен-Чапский

⇒ имя\*:
 

- Ян

⇒ владелец\*:
 

- деревня Зубаревичи
- деревня Станково
- Здание 1-ого Минского пивзавода «Оливария»
- деревня Прусиново
- деревня Негорелое

⇒ владелец\*:
 

- Дом Загорской(47)

€ ключевой sc-элемент\*:
 

- Изображение Гуттен-Чапского Кароля Эмериовича

⇒ трансляция sc-текста\*:
 

- ...
- э пример':
  - ...

€ член\*:
 

- католическая церковь

€ учащийся\*:
 

- Императорский Дерптский университет
- Гимназия Святой Анны

€ мужчина

Рисунок 9 – Описание исторической личности в БЗ

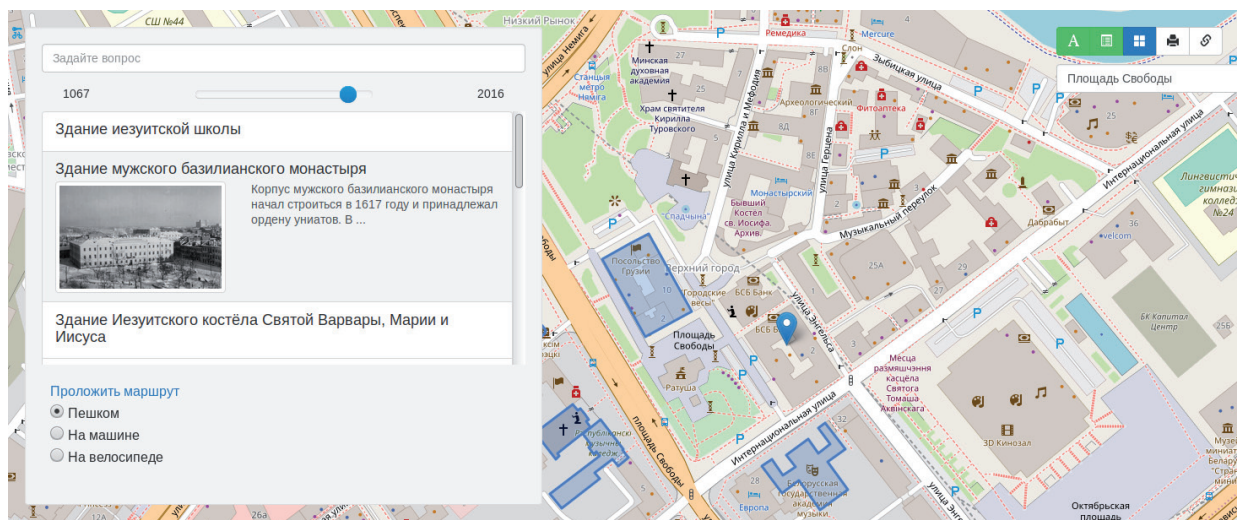


Рисунок 10 – Картографический интерфейс системы-гида



## 7 Заключение

Основными направлениями решения проблемы информационной совместимости КС являются:

- *семантическая ИТ*, в основе которой лежит смысловое представление информации в памяти КС;
- *самоорганизующаяся экосистема*, поддерживающая эволюцию и совместимость КС, построенных по семантической ИТ.

Текущий этап развития ИТ знаменует переход к *семантическим ИТ* и к соответствующей самоорганизующейся экосистеме, состоящей из *семантических КС*.

Развитию семантических ИТ и семантических КС способствуют открытые проекты, например, проект развития *Метасистемы IMS.ostis*, предоставляющий возможность каждому желающему внести свой вклад в развитие семантических ИТ.

## Благодарности

Работа выполнена при поддержке БРФФИ-РФФИ (№Ф18Р-220).

## Список источников

- [1] **Тарасов, В.Б.** От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В.Б. Тарасов. — М.: Эдиториал УРСС, 2002. — 348 с.
- [2] Нечёткие гибридные системы. Теория и практика / И.З. Батыршин [и др.]; под ред. Н.Г. Ярушкиной. — М.: Физматлит, 2007. — 208 с.
- [3] **Колесников, А.В.** Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор / А.В. Колесников, И.А. Кириков, С.В. Листопад. — М.: Ин-т проблем информатики Рос. акад. наук, 2014. — 189 с.
- [4] **Борисов, А.Н.** Построение интеллектуальных систем, основанных на знаниях, с повторным использованием компонентов / А.Н. Борисов // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2014) : материалы IV Междунар. техн.-конф., Минск, 20–22 февр. 2014г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: В. В. Голенков (отв. ред.) [и др.]. — Минск, 2014. — С. 97–102.
- [5] **Грибова, В.В.** Проект IASPaas. Комплекс для интеллектуальных систем на основе облачных вычислений / В.В. Грибова [и др.] // Искусственный интеллект и принятие решений. — 2011. — № 1. — С. 27–35.
- [6] **Гладун, А.Я.** Репозитории онтологии как средство повторного использования знаний для распознавания информационных объектов / А.Я. Гладун, Ю.В. Рогушина // Онтология проектирования. — 2013. — № 1. — С.35–50.
- [7] **Рыбина, Г.В.** Основы построения интеллектуальных систем / Г.В. Рыбина. — М.: Финансы и статистика : Инфра-М, 2010. — 430 с.
- [8] **Слободюк, А.А.** О подходе к созданию онтологий на основе системно-объектных моделей предметной области / А.А. Слободюк, С.И. Маторин, С.Н. Четвериков // Науч. ведомости Белгор. гос. ун-та. Сер.: Экономика. Информатика. — 2013. — № 22. — С. 186–193.
- [9] **Gomez-Perez, A.** Scenarios for building ontology networks within the neon methodology / A. Gomez-Perez, M.C. Suarez-Figueroa // Proceedings of the Fifth International Conference on Knowledge Capture (K-CAP 2009) / Assoc. for Computing Machinery, Spec. Interest Group on Artificial Intelligence. — New York, 2009. — P. 183–184.
- [10] **Палагин, А.В.** Проблемы трансдисциплинарности и роль информатики / А.В. Палагин // Кибернетика и системный анализ. — 2013. №5. — С. 3–13.
- [11] **Янковская, А.Е.** Анализ данных и знаний на основе конвергенции нескольких наук и научных направлений / А.Е. Янковская // Международная конференция «Интеллектуализация обработки информации» (ИОИ-8), Кипр, Пафос, 17–24 октября 2010 г. — С.196–199.
- [12] **Глоба, Л.С.** Модели и методы интеграции информационных и вычислительных ресурсов / Л.С. Глоба, Р.Л. Новогрудская // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2012) : материалы II Междунар. науч.-техн. конф., Минск, 16–18 февр. 2012 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: В. В. Голенков (отв. ред.) [и др.]. — Минск, 2012. — С. 447–452.



- [13] **Осипов, Г.С.** Приобретение знаний интеллектуальными системами: основы теории и технологии / Г.С. Осипов. — М.: Наука : Физматлит, 1997. — 112 с.
- [14] **Hazman, M.** Survey of ontology learning approaches / M. Hazman, S. El-Beltagy, A. Rafea // Intern. J. of Computer Applications. — 2011. — Vol. 22, № 9. — P. 36–43.
- [15] **Golenkov, V.** From training intelligent systems to training their development tools / V. Golenkov, N. Guliakina, N. Grakova, I. Davydenko, V. Nikulenko, A. Ereemeev, V. Tarassov // Open semantic technologies for intelligent systems, 2018. - P.81–98.
- [16] Standards of World Wide Web Consortium. — <https://www.w3.org/standards/>.
- [17] **Тузовский, А.Ф.** Интеграция информации с использованием технологий Semantic Web / А.Ф. Тузовский, В.З. Ямпольский // Проблемы информатики. — 2011. — № 2. — С.51–58.
- [18] **Мартынов, В.В.** Семиологические основы информатики / В.В. Мартынов – Минск: Наука и техника, 1974. – 192 с.
- [19] **Мельчук, И.А.** Как начиналась математическая лингвистика / И.А. Мельчук // Очерки истории информатики России, 1998 – С. 358-370.
- [20] **Мельчук, И.А.** Опыт теории лингвистических моделей Смысл-Текст, Школа «Языки русской культуры» / И.А. Мельчук. - М.: 1999 — 346 с.
- [21] **Davydenko, I.** Semantic models, method and tools of knowledge bases coordinated development based on reusable components / I. Davydenko // Open semantic technologies for intelligent systems, 2018. - P.99–118,
- [22] **Голенков, В.В.** Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 1: Принципы создания / В.В. Голенков, Н.А. Гулякина // Онтология проектирования, 2014, № 1(11). - С.42–64.
- [23] IMS.ostis metasytem. – <https://ims.ostis.net>.
- [24] **Голенков, В.В.** Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 2: Унифицированные модели проектирования / В.В. Голенков, Н.А. Гулякина // Онтология проектирования, 2014, № 4(14). - С.34–53.
- [25] M. Wooldridge, An introduction to multiagent systems. Chichester : J. Wiley, 2009.
- [26] **Максимов, Н.В.** Семантическое ядро цифровой платформы / Н.В. Максимов, О.Л. Голицина, М.Г. Ганченкова, Д.В. Санатов, А.В. Разумов // Онтология проектирования. - 2018. - № 3 (29). - С. 412-426. – DOI: 10.18287/2223-9537-2018-8-3-412-426.
- [27] **Shunkevich, D.** Agent-oriented models, method and tools of compatible problem solvers development for intelligent systems / D. Shunkevich // Open semantic technologies for intelligent systems, 2018. - P.119–132,
- [28] **Hauptert, J.** The object memory server for semantic product memories. SemProM / J. Hauptert, M. Schneider // Springer, 2013. - P.175–189.
- [29] **Grangel-González, I.** Towards a Semantic Administrative Shell for Industry 4.0 Components / I. Grangel-González, L. Halilaj, G. Coskun, S. Auer, D. Collarana, M. Hofmeister // Proc. of IEEE Tenth International Conference on Semantic Computing (ICSC). – 2016. - P. 230-237.
- [30] **Kharlamov, E.** Capturing Industrial Information Models with Ontologies and Constraints / E. Kharlamov, B. Cuenca Grau, E. Jiménez-Ruiz, S. Lamparter, G. Mehdi, M. Ringsquandl, Y. Nenov, S. Grimm, M. Roshchin, I. Horrocks // Proc. of International Semantic Web Conference. – 2016. - P. 325-343.
- [31] **Голенков, В.В.** Проектирование предприятий рецептурного производства на основе онтологий / В.В. Таберко, Д.С. Иванюк, В.В. Голенков, К.В. Русецкий, Д.В. Шункевич, И.Т. Давыденко, В.В. Захаров, В.П. Ивашенко, Д.Н. Корончик // Онтология проектирования. – 2017. – Т. 7, №2(24). - С. 123-144. – DOI: 10.18287/2223-9537-2017-7-2-123-144.
- [32] **Головко, В.А.** Интеграция искусственных нейронных сетей с базами знаний / В.А. Головко, В.В. Голенков, В.П. Ивашенко, В.В. Таберко, Д.С. Иванюк, А.А. Крощенко, М.В. Ковалев // Онтология проектирования. – 2018 – Т.8, №3(29). – С.366-386. – DOI: 10/18287/2223-9537-2018-8-3-366-386.

## ONTOLOGICAL DESIGN OF HYBRID SEMANTICALLY COMPATIBLE INTELLIGENT SYSTEMS BASED ON SENSE REPRESENTATION OF KNOWLEDGE

V. Golenkov<sup>1,a</sup>, N. Guliakina<sup>1,b</sup>, I. Davydenko<sup>1,c</sup>, D. Shunkevich<sup>1,d</sup>, A. Eremeev<sup>2,e</sup>

<sup>1</sup>Belarusian state university of informatics and radioelectronics, Minsk, Republic of Belarus

<sup>2</sup>National Research University «Moscow Power Engineering Institute», Moscow, Russian Federation

<sup>a</sup>golen@bsuir.by, <sup>b</sup>guliakina@bsuir.by, <sup>c</sup>ir.davydenko@gmail.com, <sup>d</sup>shunkevichdv@gmail.com, <sup>e</sup>YeremeevAP@mpei.ru

### Abstract

The work is devoted to the problem of ensuring the semantic compatibility of intelligent systems. It is shown that ensuring the compatibility of intelligent systems and the development of relevant standards is a key direction in the development of intelligent systems design technologies. The concept of the sense representation of information in the memory of an intelligent system has been formally clarified, which ensures the unambiguity of the presentation of information using a given set of concepts. The possibility of the automatic integration of knowledge within the the sense representation of knowledge, which is reduced to the merging of synonymous signs, is shown. The possibility of automatic integration of various models of knowledge processing is shown, if these models are teams of agents, oriented on processing of knowledge, presented in the memory of intelligent systems in a sense form, and interacting with each other also through the specified memory. The concept of OSTIS Technology, focused on the development of semantic computer systems and implementing the specified principles of information representation and processing, is proposed. The concept of the OSTIS Ecosystem is proposed, which is a group of interacting intelligent systems built using the OSTIS Technology and supporting the evolution and compatibility of these intelligent systems during their operation within this ecosystem. The examples of the use of OSTIS Technology in the development of applied intelligent systems are considered.

**Keywords:** intelligent systems, knowledge representation language, intelligent system model, component design, systems compatibility.

**Citation:** Golenkov VV, Guliakina NA, Davydenko IT, Shunkevich DV, Eremeev AP. Ontological design of hybrid semantically compatible intelligent systems based on sense representation of knowledge [In Russian]. *Ontology of designing*. 2019; 9(1): 132-151. – DOI: 10.18287/2223-9537-2019-9-1-132-151.

### Acknowledgment

This work was supported by the BRFFR and RFBR (№F18R-220).

### References

- [1] **Tarassov VB.** From multi-agent systems to intellectual organizations: philosophy, psychology, computer science [In Russian]. — Moscow: Editorial URSS, 2002. — 348 p.
- [2] Fuzzy hybrid systems. Theory and practice [In Russian]. I. Z. Batyrshyn [and other]; ed. by N. G. Yaryshkina. — Moscow: Fizmalit, 2007. — 208 p.
- [3] **Kolesnikov AV, Kirikov IA, Listopad SV.** Hybrid Intelligent Systems with Self-Organization: Coordination, Consistency, Dispute [In Russian]. — Moscow: Institute of Informatics Problems RAS, 2014. — 189 p.
- [4] **Borisov AN.** Building intelligent knowledge-based systems with reusable components [In Russian]. Open Semantic Technologies for Intelligent Systems (OSTIS-2014): Materials IV Intern. tech.-conf., Minsk, 20–22 Feb. 2014r. / Belarusian State University of Informatics and Radioelectronics; ed. by: V.V. Golenkov [and other]. — Minsk, 2014. — P.97–102.
- [5] **Gribova VV.** Project IACPaaS. The complex for intelligent systems based on cloud computing [In Russian]. Artificial intelligence and decision making. 2011; 1: 27–35.
- [6] **Gladun AYA, Rogushina YuV.** Ontology repositories as a means of knowledge reuse for the recognition of information objects [In Russian]. *Ontology of designing*. 2013; 1: 35–50.
- [7] **Rybina GV.** Basics of building intellectual systems: studies. Manual [In Russian]. - Moscow: Finance and Statistics: Infra-M, 2010. - 430 p.

- [8] **Slobodiuk AA, Matorin SI, Chetverikov SN.** On the approach to the creation of ontologies based on the system-object models of the domain [In Russian]. Scientific. statements Belgor. state un-that. Ser.: Economy. Computer science. 2013; 22: 186–193.
- [9] **Gomez-Perez A, Suarez-Figueroa MC.** Scenarios for building ontology networks within the neon methodology // Proceedings of the Fifth International Conference on Knowledge Capture (K-CAP 2009) / Assoc. for Computing Machinery, Spec. Interest Group on Artificial Intelligence. — New York, 2009. — P.183–184.
- [10] **Palagin AV.** Problems of transdisciplinarity and the role of informatics [In Russian]. Cybernetics and Systems Analysis. 2013; 5: 3–13.
- [11] **Yankovskaya AE.** Data and knowledge analysis based on the convergence of several sciences and scientific fields [In Russian]. International Conference «Intellectualization of Information Processing» (IIP-8), Cyprus, Paphos, 17–24 Oct 2010. — P.196–199.
- [12] **Globa LS, Novogrudskaya RL.** Models and methods of integration of information and computing resources [In Russian]. Open semantic technologies for designing intelligent systems (OSTIS-2012): Proceedings of the II Intern. scientific and technical Conf., Minsk, 16–18 Feb. 2012 / Belarusian. state University of Informatics and Radioelectronics; Editorial: V. V. Golenkov (ed.) [et al.]. - Minsk, 2012. - P.447–452.
- [13] **Osipov GS.** Acquisition of knowledge by intellectual systems: fundamentals of theory and technology [In Russian]. - Moscow: Science: Fizmatlit, 1997. - 112 p.
- [14] **Hazman M, El-Beltagy S, Rafea A.** Survey of ontology learning approaches // Intern. J. of Computer Applications. 2011; 22(9): 36–43.
- [15] **Golenkov V, Guliakina N, Grakova N, Davydenko I, Nikulenko V, Ereemeev A, Tarassov V.** From training intelligent systems to training their development tools // Open semantic technologies for intelligent systems, 2018: 81–98.
- [16] Standards of World Wide Web Consortium. — <https://www.w3.org/standards/>.
- [17] **Tuzovsky AF, Yampolsky VZ.** Integration of information using Semantic Web technologies [In Russian]. *Problems of Informatics*. 2011; 2: 51–58.
- [18] **Martynov VV.** Semiological foundations of computer science [In Russian]. - Minsk: Science and technology, 1974. — 192 p.
- [19] **Melchuk IA.** How did mathematical linguistics begin [In Russian]. Essays on the history of computer science in Russia, 1998. — P. 358–370.
- [20] **Melchuk IA.** Experience of the theory of linguistic models Meaning-Text, School «Languages of Russian culture» [In Russian]. - Moscow: 1999. — 346 p.
- [21] **Davydenko I.** Semantic models, method and tools of knowledge bases coordinated development based on reusable components // Open semantic technologies for intelligent systems, 2018: 99–118.
- [22] **Golenkov VV, Gulyakina NA.** Project of open semantic technology of the component design of intelligent systems. Part 1: Principles of creation [In Russian]. *Ontology of designing*. 2014; 1(11): 42–64.
- [23] IMS.ostis metasytem. — <https://ims.ostis.net>.
- [24] **Golenkov VV, Gulyakina NA.** Project of open semantic technology of the component design of intelligent systems. Part 2: Unified Design Models [In Russian]. *Ontology of designing*. 2014; 4(14): 34–53.
- [25] **Wooldridge M.** An introduction to multiagent systems. Chichester: J. Wiley, 2009.
- [26] **Maksimov NV, Golitsina OL, Ganchenkova MG, Sanatov DV, Razumov AV.** Semantic core of digital platform [In Russian]. *Ontology of designing*. 2018; 8(3): 412–426. - DOI: 10.18287/2223-9537-2018-8-3-412-426.
- [27] **Shunkevich D.** Agent-oriented models, method and tools of compatible problem solvers development for intelligent systems // Open semantic technologies for intelligent systems, 2018: 119–132.
- [28] **Haupt J, Schneider M.** The object memory server for semantic product memories. SemProM, Springer, 2013. - P.175–189.
- [29] **Grangel-González I, Halilaj L, Coskun G, Auer S, Collarana D, Hofmeister M.** Towards a Semantic Administrative Shell for Industry 4.0 Components // Proc. of IEEE Tenth International Conference on Semantic Computing (ICSC). — 2016. - P.230–237.
- [30] **Kharlamov E, Cuenca Grau B, Jiménez-Ruiz E, Lamparter S, Mehdi G, Ringsquandl M, Nenov Y, Grimm S, Roshchin M, Horrocks I.** Capturing Industrial Information Models with Ontologies and Constraints // Proc. of International Semantic Web Conference. — 2016. - P.325–343.
- [31] **Golenkov VV, Taberko VV, Ivanyuk DS, Rusetski KV, Shunkevich DV, Davydenko IT, Zakharov VV, Ivashenko VP, Koronchik DN.** Designing Batch Manufacturing Enterprises Using Ontologies [In Russian]. *Ontology of designing*. 2017; 7(2): 123–144. — DOI: 10.18287/2223-9537-2017-7-2-123-144.
- [32] **Golovko VA, Golenkov VV, Ivashenko VP, Taberko VV, Ivaniuk, DS, Kroshchanka AA, Kovalev MV.** Integration of artificial neural networks and knowledge bases [In Russian]. *Ontology of designing*. 2018; 8(3): 366–386. - DOI: 10.18287/2223-9537-2018-8-3-366-386.

## Сведения об авторах



**Голенков Владимир Васильевич**, 1949 г. рождения. В 1971 г. с отличием окончил физический факультет Белорусского государственного университета, д.т.н. (1996), профессор, заведующий кафедрой интеллектуальных информационных технологий Белорусского государственного университета информатики и радиоэлектроники БГУИР, инициатор ежегодной международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем», член Российской ассоциации искусственного интеллекта. В списке научных трудов более 100 работ в области семантических технологий.

**Vladimir Vasilevich Golenkov** was born in 1949. In 1971 he graduated from the physics department of Belarusian State University with honors diploma, Dr. of science (1996), professor. The head of the department of Intelligent information technologies of Belarusian

State University of Informatics and Radioelectronics, initiator of the annual international science and technical conference «Open Semantic technologies for Intelligent Systems» (OSTIS), member of Russian association of artificial intelligence. The list of his published works consists of more than 100 articles on the various aspects of semantic approach in the intelligent systems design.

**Гулякина Наталья Анатольевна**, 1952 г. рождения. В 1974 г. окончила факультет прикладной математики Белорусского государственного университета, к.ф.-м.н., доцент, заместитель заведующего кафедрой интеллектуальных информационных технологий БГУИР, член программного комитета конференции «Открытые семантические технологии проектирования интеллектуальных систем», автор ряда публикаций по вопросам интеллектуальных информационных технологий, автор семи учебно-методических пособий.

**Natalia Anatolievna Guliakina** was born in 1952. In 1971 she graduated from the faculty of applied mathematics of Belarusian State University, Ph.D., and assistant professor. Deputy head of the department of Intelligent information systems of Belarusian State University of Informatics and Radio-electronics, a member of Program Committee of the international conference «Open Semantic technologies for Intelligent Systems» (OSTIS), author of several publications on artificial intelligence systems design, author of seven educational toolkits.



**Давыденко Ирина Тимофеевна**, 1987 г. рождения. В 2010 г. окончила БГУИР по специальности «Искусственный интеллект», к.т.н., доцент кафедры интеллектуальных информационных технологий БГУИР. Имеет более 80 печатных работ в области семантических технологий.

**Irina Timofeevna Davydenko** was born in 1987. In 2010 she graduated from Belarusian State University of Informatics and Radioelectronics majoring in Artificial intelligence. Ph.D. Associate Professor in Belarusian State University of Informatics and Radioelectronics. Has over 80 published works on the subject of semantic technologies.

**Шункевич Даниил Вячеславович**, 1990 г. рождения. В 2012 г. с отличием окончил БГУИР по специальности «Искусственный интеллект», к.т.н., доцент кафедры интеллектуальных информационных технологий БГУИР. Имеет более 60 печатных работ в области семантических технологий.

**Daniil Vyacheslavovich Shunkevich** was born in 1990. In 2012 he graduated from Belarusian State University of Informatics and Radioelectronics with honors diploma majoring in Artificial intelligence. Ph.D. Associate Professor in Belarusian State University of Informatics and Radioelectronics. Has over 60 published works on the subject of semantic technologies.



**Еремеев Александр Павлович**, 1948 г. рождения. Д.т.н., профессор, заведующий кафедрой прикладной математики МЭИ, лауреат премии Президента РФ в области образования, действительный член Российской академии естественных наук. Сфера научных интересов: искусственный интеллект, принятие решений, системы поддержки принятия решений

**Aleksander Pavlovich Eremeev** was born in 1948. Dr. of science, professor. The head of the department of Applied Mathematics MEI, winner of the President of the Russian Federation in the field of education, full member of the Russian Academy of Natural Sciences. Research Interests: Artificial Intelligence, Decision Making, Decision Support Systems.



УДК 004.891

## ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ СЛИЯНИЯ ДАННЫХ ПРИ ДИАГНОСТИРОВАНИИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

С.М. Ковалев<sup>1</sup>, А.Е. Колоденкова<sup>2</sup>, В. Снасель<sup>3</sup>

<sup>1</sup> Ростовский государственный университет путей сообщения, Ростов-на-Дону, Россия  
ksm@rfnias.ru

<sup>2</sup> Самарский государственный технический университет, Самара, Россия  
anna82\_42@mail.ru

<sup>3</sup> Остравский технический университет, Острава, Чешская Республика  
vaclav.snasel@vsb.cz

### Аннотация

Слияние разнородных данных, полученных в реальном времени от различных датчиков, является важной задачей при диагностировании технических объектов. В статье рассмотрены вопросы терминологии слияния данных на основе обзора литературы, предложено новое определение термина «слияние данных». Обобщены и систематизированы научные взгляды на проблему слияния данных при диагностировании технических объектов в условиях множества разнотипных датчиков и разнородной информации. Приведена адаптированная классификация слияния данных с учётом различных критериев (отношения между датчиками, уровень абстракции данных, тип архитектуры), а также классификация структурных моделей слияния данных, разработанных для построения интеллектуальных систем слияния данных. Проведён сравнительный анализ моделей процесса слияния данных, представлены их структуры, выявлены достоинства и недостатки моделей. Отмечено, что для эффективного сбора исходных данных, поступающих от множества разнотипных датчиков, и их обработки можно использовать несколько моделей слияния данных или их комбинации. Все основные аспекты, касающиеся интеллектуальных технологий слияния данных, рассмотрены в столь полном объёме на русском языке впервые.

**Ключевые слова:** разнородные данные, модели слияния данных, диагностирование технических объектов, множество разнотипных датчиков.

**Цитирование:** Ковалев, С.М. Интеллектуальные технологии слияния данных при диагностировании технических объектов / С.М. Ковалев, А.Е. Колоденкова, В. Снасель // Онтология проектирования. – 2019. – Т.9, №1(31). – С.152-168. – DOI: 10.18287/2223-9537-2019-9-1-152-168.

### Введение

С развитием сложности, масштабности, многофункциональности технических объектов (ТО) первостепенное значение приобретает проблема эффективной и безопасной эксплуатации ТО. Это связано с тем, что ухудшение технического состояния на этапе эксплуатации ТО приводит к нарушению его работоспособности (отказу), выражающемуся в изменении значений параметров его работы за пределами, регламентируемыми нормативно-техническими документами, или к полной его остановке, что может привести к значительному материальному ущербу, а также нарушению экологической обстановки.

Одним из способов решения данной проблемы является автоматизация и информатизация процессов контроля и диагностирования, в ходе которых решают такие задачи, как определение вида технического состояния, обнаружение неисправностей, прогнозирование технического состояния в условиях использования различных типов датчиков. Увеличение количества различных типов датчиков, разнородных диагностических данных привело к необ-



ходимости применения технологии слияния данных. Разработка этой технологии для диагностирования ТО даёт выигрыш по сравнению с отдельной обработкой за счёт расширения объёма получаемых диагностических данных и синергетического эффекта. Достоинствами данной технологии являются: сохранение надёжности ТО в случае отказа какого-либо датчика; рост вероятности обнаружения нештатных ситуаций; уменьшение времени принятия решений при диагностировании ТО в силу получения более полной и точной информации о ТО, поступающей от множества разнородных датчиков.

Цель статьи - дать анализ существующих технологий слияния данных для решения задач технического диагностирования ТО. Авторы попытались не только обобщить научные взгляды на проблему слияния данных [1-5], но и представить свой опыт, полученный при диагностировании ТО.

## 1 Проблемы слияния данных

С увеличением количества типов датчиков данных и объёма информации возникла необходимость в обработке разнородных данных для последующего анализа. Обработку таких данных понимают как соединение, комплексирование, интеграцию, объединение, сращивание (в англоязычной литературе часто называют «слияние данных» - *data fusion*) [3-7].

В настоящей работе под *слиянием данных* понимается совместная обработка разнородной информации, полученной в реальном времени от множества различных датчиков, в информацию, легко воспринимаемую человеком, на основе которой можно контролировать ТО, проводить оценку ситуации и принимать научно-обоснованные управленческие решения [8]. Несмотря на многочисленные применения технологии слияния разнородных данных при диагностировании ТО, остаётся ряд основных проблем.

- *Разнородность данных.* Данные от датчиков могут быть представлены в виде разнородных данных и характеризоваться [9]: разнообразием шкал измерения (номинальной, числовой и др.); различными типами данных (целый, вещественный, логический и др.); различной структурой представления (статистические, темпоральные, нечёткие, экспертные данные, изображения и др.); различными типами моделей баз данных (реляционные, иерархические, сетевые и др.); различной степенью достоверности, полноты и точности данных, измеряемых в различных масштабах и единицах измерения.
- *Точность датчиков.* Часто применяются многопараметрические датчики, которые способны измерять одновременно несколько физических величин, характеризующих контролируемый объект или процесс. Объединяя данные от нескольких одинаковых датчиков, можно улучшить их качество, а также повысить их точность и достоверность.
- *Выбор метода слияния данных.* Не существует универсального метода слияния данных, который был бы оптимален при всех условиях.
- *Терминология.* Спецификации данных выполняются различными экспертами, которые могут использовать одинаковые термины в различном смысле и, наоборот, для одного и того же понятия могут использовать различные названия. Для решения данной проблемы целесообразно использовать единый словарь терминов предметной области.
- *Идентификация сущностей.* Информация об одной и той же сущности (ситуации, состоянии объекта, процессе и т.п.) представляется в распределённой форме [6, 9]. Для решения проблемы целесообразно установить однозначное соответствие между описаниями, которые находятся в различных локальных датчиках, но относятся к одному экземпляру сущности предметной области.

## 2 Классификация слияния данных

Ввиду междисциплинарности области исследований предлагается слияние данных классифицировать по нескольким критериям (рисунок 1 адаптирован по материалам [10-12]). Слияние данных по критерию отношений между датчиками представлено на рисунке 2 (адаптирован по материалам [2, 13, 14]).

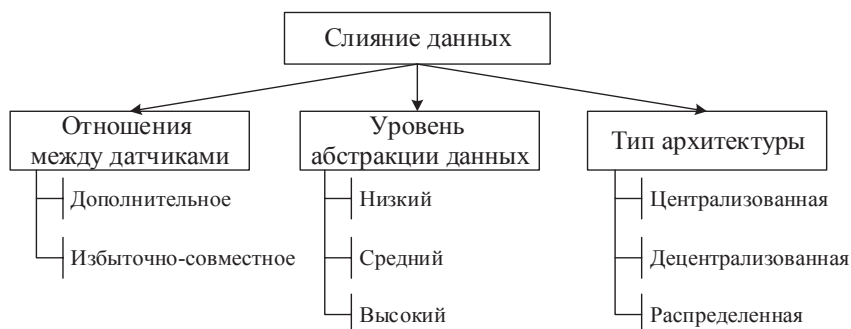


Рисунок 1 – Классификация слияния данных

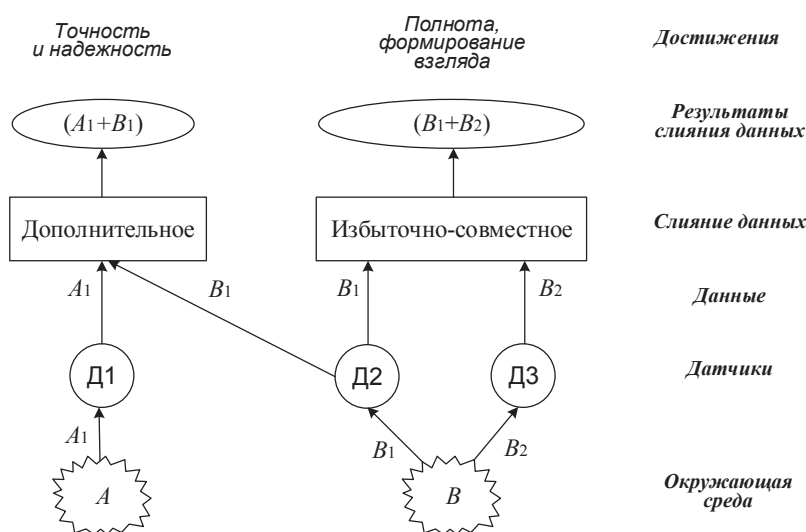


Рисунок 2 – Типы слияния данных на основе отношений между датчиками

*Дополнительное (complementary)* слияние – это слияние, при котором независимые датчики ( $D_1$  и  $D_2$ , принадлежащие различным частям пространства окружающей среды) могут быть объединены ( $A_1 + B_1$ ) для того, чтобы дать более полное представление наблюдаемому явлению (ситуации). Данные от независимых датчиков легко объединяются.

*Избыточно-совместное (competitive-cooperative)* слияние – это слияние, где каждый датчик ( $D_2$  и  $D_3$ ) поставляет информацию об одной и той же характеристике ТО, в результате которого достигается более точное представление этой информации ( $B_1 + B_2$ ). Различают две возможные конфигурации избыточно-совместного слияния данных: с разных датчиков или слияние измерений с одного датчика, сделанных в разные моменты времени. Избыточно-совместное слияние используется для отказоустойчивых и надёжных систем.

Согласно *уровню абстракции данных (level of abstraction)* слияние данных может быть классифицировано на три уровня (см. таблицу 1) [13].

- Низкий уровень слияния (*low-level fusion*), именуемый часто уровнем необработанных данных (*raw data level*), поступающих от датчиков. На данном уровне может применяться нечёткий фильтр Калмана.
- Средний уровень слияния (*medium-level fusion*), именуемый часто уровнем свойств (*feature level*). На данном уровне происходит слияние свойств (форма, рёбра, углы, линии, положение), в результате которого получают новые объекты, используемые для других задач, например, сегментации и распознавания. Также происходит фильтрация, нормализация, корреляция, классификация данных с использованием методов «мягких вычислений» и методов интеллектуального анализа данных.

- Высокий уровень слияния (*high-level fusion*), именуемый часто уровнем решения (*decision level*). Слияние на уровне решения происходит с использованием интеллектуальных методов: теории нечётких множеств, нейронных сетей, теории Демпстера-Шефера (*Dempster-Shafer*), Байесовского подхода, гибридных подходов, в результате которых получается совместное решение о диагностировании ТО с учётом разнородных данных, полученных от нескольких датчиков, а также мнений и решений нескольких экспертов.

Таблица 1 – Основные характеристики уровней слияния данных

Характеристики \ Уровни	Низкий	Средний	Высокий
Объём передаваемой информации	очень большой	средний	очень маленький
Информационные потери	без потерь	немного	средние
Математические вычисления	простые	средние	сложные

Как видно из таблицы 1, потеря информации увеличивается по мере её передачи с низких уровней на более высокие. При этом информации на низком уровне больше, чем на высоком, что делает методы низкого уровня дорогостоящими за счёт требований, предъявляемых к точности их вычислений.

На рисунках 3-5 показаны три типа архитектур слияния данных с несколькими датчиками (рисунки 3-5 адаптированы по материалам [14-16]).

- *Централизованная архитектура (centralized architecture)* (рисунок 3).

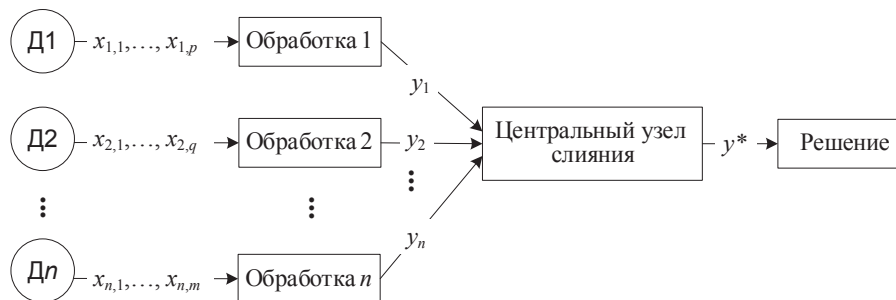


Рисунок 3 – Централизованная архитектура

Все собранные данные от датчиков ( $D_1, D_2, \dots, D_n$ ) будут обработаны и отправлены в центральный узел слияния, который способен эффективно обрабатывать большие объёмы данных. Если данные правильно выровнены и связаны, а объём передаваемой информации не ограничен, то централизованная архитектура позволяет дать теоретическое оптимальное решение для оценки состояния ТО. Однако обработка всей информации на центральном узле создаёт проблемы, такие как большая вычислительная нагрузка на центральный узел, возможность сбоя работы центрального узла, негибкость изменений в архитектуре [1, 11].

Если данные представляют собой изображения, то это приводит к низкой пропускной способности отправки необработанных данных и задержки с выводом данных.

- *Децентрализованная архитектура (decentralized architecture)* (рисунок 4).

Данная архитектура состоит из нескольких локальных процессоров, которые в соответствии со своим наблюдением вычисляют параллельно локальные оценки. Центральный центр слияния отсутствует, поэтому сначала каждый локальный процессор вычисляет локальную оценку, а затем корректирует её путём интегрирования локальных оценок других процессоров для получения точной и полной оценки. Результатом применения децентрализованной архитектуры является множество решений, из которых выбирается одно наилучшее. Увеличение числа узлов приводит к увеличению стоимости связи, что является одним из недостатков децентрализованной архитектуры.

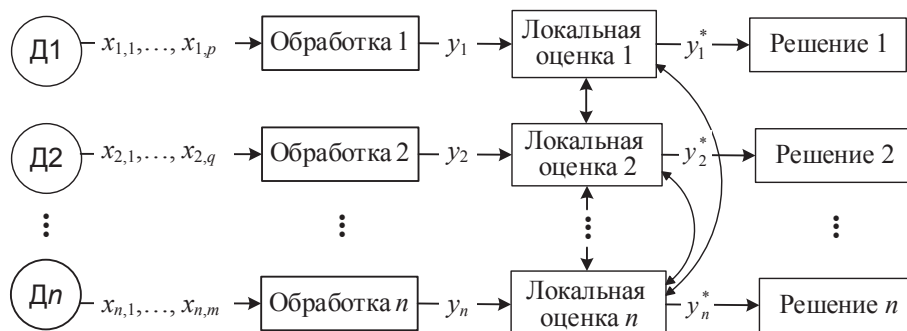


Рисунок 4 – Децентрализованная архитектура

- *Распределённая архитектура (distributed architecture) (рисунок 5).*

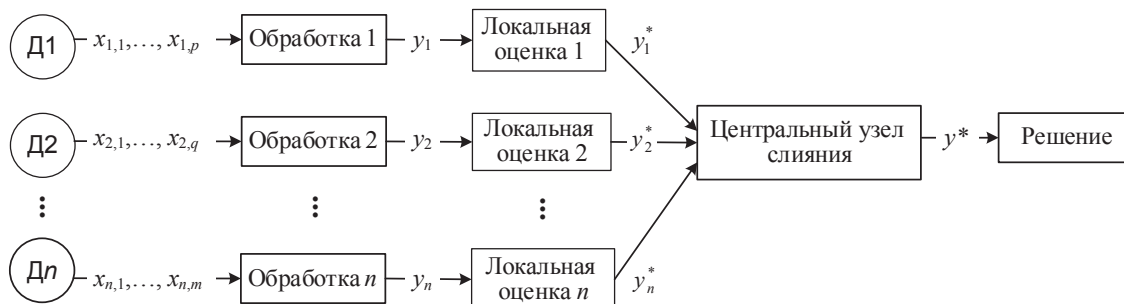


Рисунок 5 – Распределённая архитектура

Данная архитектура состоит из нескольких локальных процессоров, вычисляющих параллельно локальные оценки. Затем все собранные локальные оценки отправляются для обработки в центральный узел слияния с целью получения совместного решения. При этом локальные процессоры не сливают свои оценки с оценками, полученными от других процессоров, как в децентрализованной архитектуре, а передают их сразу в центральный узел слияния. При распределённой архитектуре за счёт распределения нагрузки достигается низкая нагрузка на каждый локальный процессор при вычислении оценки; более низкая стоимость связи, гибкость к изменениям и устойчивость к неудачам.

### 3 Структурные модели слияния данных

Получение нового качества данных, а именно полных, точных, своевременных данных, которые поддаются интерпретации, является наиболее сложным этапом преобразования данных интеллектуальных систем. Одним из путей решения данной проблемы явилась разработка моделей слияния данных. В настоящее время предложено несколько моде-

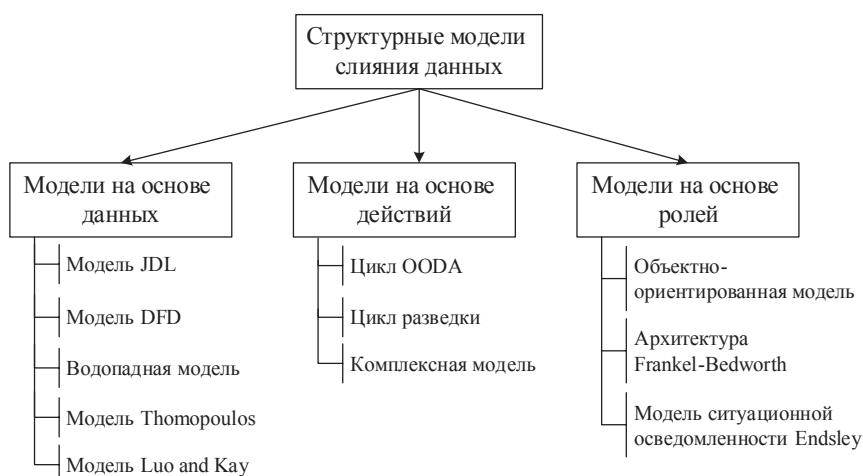


Рисунок 6 – Классификация моделей слияния данных

лей слияния данных, разработанных для выделения спецификаций, предложений и использования слияния данных в беспроводной сенсорной сети, которая представляет собой распределённую самоорганизующуюся сеть множества датчиков и исполнительных устройств, объединённых между собой посредством радиоканала. Модели слияния данных можно разделить на три основных вида: модель на основе данных, модель на основе действий и модель на основе ролей, как показано на рисунке 6 (адаптирован по материалам [17, 18]).

### 3.1 Модели на основе данных (data-based model)

#### 3.1.1 Модель JDL (*Joint Directors of Laboratories*)

Самой известной моделью в области слияния данных является модель *JDL*, предложенная исследовательской группой *Joint Directors of Laboratories* совместно с Министерством обороны США [18]. Модель рассчитана на связи между исследователями слияния данных и инженерами по внедрению [19-21].

На рисунке 7 представлена структура модели слияния *JDL*. Модель *JDL* включает следующие уровни обработки данных.



Рисунок 7 – Структура модели слияния JDL

- *Уровень 0* (предварительная обработка датчиков, *source preprocessing*) направлен на уменьшение объёма данных для последующих уровней обработки за счёт распределения данных между подходящими процессами и выбора подходящих датчиков.
- *Уровень 1* (уточнение объекта, *object refinement*) преобразует данные в согласованную структуру. На данном уровне происходит идентификация данных с применением, например, методов классификации или распознавания, а также локализация датчика; применяются алгоритмы отслеживания местонахождения объекта.
- *Уровень 2* (уточнение ситуации (*situation refinement*)) обеспечивает контекстное описание отношений между объектами и исследуемым (наблюдаемым) событием. На данном уровне используется априорная информация, знания и сведения об окружающей среде.
- *Уровень 3* (уточнение угроз (*threat refinement*)) оценивает текущую ситуацию и на её основе выявляет возможное возникновение угроз, уязвимостей и возможностей для оперативных действий.
- *Уровень 4* (уточнение процесса (*process refinement*)) отвечает за мониторинг работы системы, распределение ресурсов в соответствии с заданными целями. Могут использоваться, например, теория полезности, линейное программирование, методы, основанные на знаниях.

Модель JDL рассматривает слияние данных с точки зрения системного подхода и даёт целостное представление о структуре системы. При этом она фокусируется на данных (ввод/вывод), а не на обработке; роль человека в процессе слияния не представлена [1, 19]. Методы и алгоритмы слияния данных модели *JDL*, рассмотрены в работе [16].



### 3.1.2 Модель DFD (Data-Feature-Decision)

Предложенная В. Дасараты (*V. Dasarathy* [22]) модель DFD (рисунок 8) определяет элементы процесса слияния данных, которые задаются на основе входных (необработанных данных) и выходных данных (некоторого решения) и способствует установлению связей между данными и задачами. Подробное описание модели *DFD* дано в работах [21, 22].

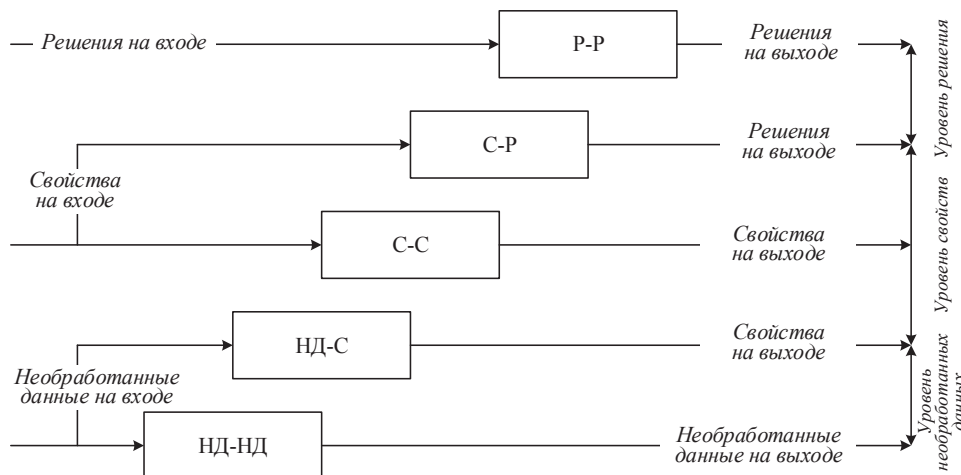


Рисунок 8 – Структура модели DFD

Модель *DFD* имеет пять уровней.

- Необработанные данные на входе - Необработанные данные на выходе (НД-НД) (*Data In-Data Out (DAI-DAO)*). Слияние осуществляется сразу после того, как данные получены с датчиков. Результатом слияния необработанных данных являются такие же данные, только более точные или надёжные.
- Необработанные данные на входе - Свойства на выходе (НД-С) (*Data In-Feature Out (DAI-FEO)*). На данном уровне для слияния используются необработанные данные, полученные от датчиков, для извлечения свойств, описывающих ТО либо ситуацию.
- Свойства на входе - Свойства на выходе (C-C) (*Feature In-Feature Out (FEI-FEO)*). На данном уровне происходит работа над набором свойств для получения нового набора свойств с целью их улучшения/уточнения или извлечения новых.
- Свойства на входе - Решения на выходе (C-P) (*Feature In-Decisions Out (FEI-DEO)*). Для получения решений используются наборы свойств ТО или ситуации.
- Решения на входе - Решения на выходе (P-P) (*Decisions In-Decisions Out (DEI-DEO)*). Данный уровень объединяет входные решения для получения новых решений.

Модель DFD полезна для разработки алгоритмов слияния в беспроводной сенсорной сети, однако не предоставляет системного представления, как модель JDL. Вместо этого модель DFD предоставляет детальный подход к определению задач слияния с помощью ожидаемых входных и выходных данных.

### 3.1.3 Водопадная модель (Waterfall model)

Модель, предложенная С. Харрисом (*C.J. Harris*) [23], представляет собой иерархическую структуру, где информация, произведённая одним модулем, используется в следующем модуле. Последний модуль (принятие решений) передаёт достаточное количество информации в модуль контроля для калибровки и настройки датчиков.

На рисунке 9 представлена структура водопадной модели, имеющая 3 уровня, каждый из которых включает два модуля и замкнутый цикл для работы в системе [24, 25].

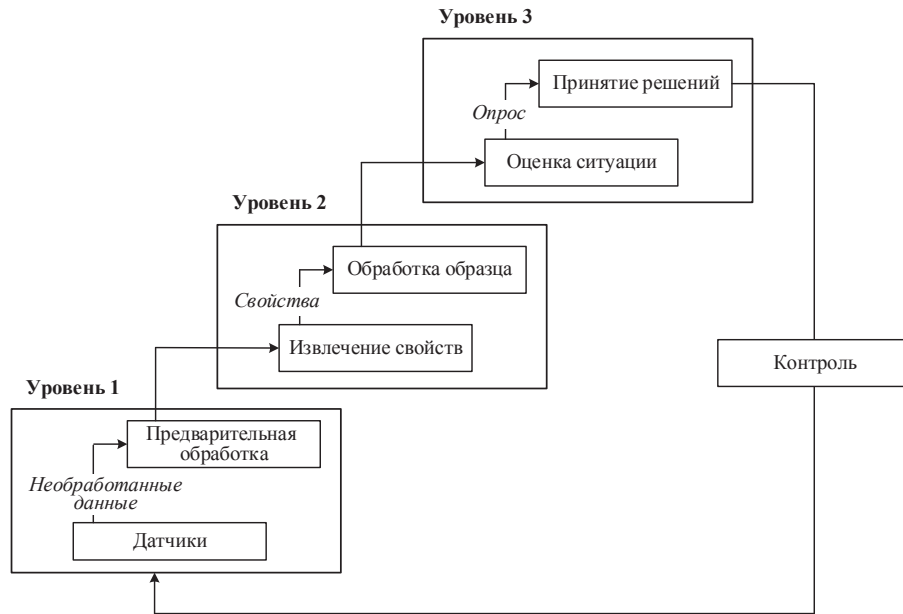


Рисунок 9 – Структура водопадной модели

- *Уровень 1.* Данные из окружающей среды собираются и обрабатываются. На следующий уровень поступают не только обработанные данные, но и информация о датчиках.
- *Уровень 2.* Извлекаются основные свойства ТО, а затем сливаются, тем самым снижая количество передаваемых данных и увеличивая их информационную насыщенность.
- *Уровень 3.* Разрабатываются различные сценарии ситуаций, а также действия для конкретных ситуаций, действий в понятной форме, осуществляются оценки ситуаций. Принимается решение на основе информации, собранной с предыдущих уровней, составляются возможные пути действий для определённой ситуации.

Водопадная модель является более точной в анализе процесса слияния данных, чем другие модели. Однако она имеет ряд недостатков: отсутствие описания обратной связи потока данных; указаний на то, должны ли датчики быть параллельными или последовательными, хотя обработка информации – последовательная [24, 26].

В работах [4, 27] предложен механизм обратной связи с тем, чтобы информация, поступающая к датчикам из модуля принятия решений, постоянно обновлялась.

### 3.1.4 Модель Томопулоса (Thomopoulos Model)

Данная модель была предложена С. Томопулосом (*S.C. Thomopoulos*, 1989) [28] и представляет трёхуровневую модель, сформированную на основе уровня сигнала (*signal level*), уровня доказательств (*evidence level*) и уровня динамики (*dynamics level*) (рисунок 10).

На каждом уровне собранные данные сталкиваются с ранее обработанными и сохранёнными данными, сохраняя при этом заданный порядок, что означает необходимость решения проблем с задержкой или ошибками в передаче данных. В зависимости от применения эти уровни слияния могут быть реализованы последовательно или поочередно [4, 25].

После того, как датчики проследили за ТО, они выводят свои измерения. Затем

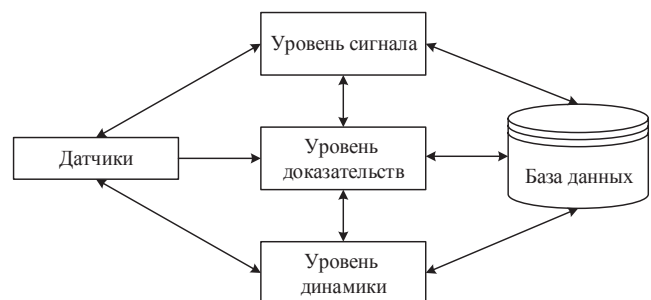


Рисунок 10 – Структура модели Thomopoulos

уровень сигнала обрабатывает эти данные, выполняя корреляции из-за отсутствия математической модели. Поэтому в процессе обучения собранные данные соотносятся с информацией, ранее сохранённой в базе данных. На уровне доказательств данные объединяются на различных уровнях вывода на основе статистической модели и оценки, требуемой пользователем (например, принятие решений или тестирование гипотез). Недостатком данной модели является отсутствие математической модели, описывающей процесс сбора данных.

### 3.1.5 Модель Ло и Кей (Luo and Kay)

Данная модель представлена Ло и Кей (Luo & Kay, 1998) [29] как общая структура слияния данных, основанная на иерархической модели, но отличающаяся от водопадной модели (рисунок 11). В данной модели данные от датчиков попадают на различные центры слияния, тем самым повышая уровень представления от необработанных данных к решениям.

Из рисунка 11 видно, что в системе присутствуют три функции, которые обычно используются в процессе слияния: выбор датчика, модель ситуации, преобразование данных. Линии от системы к каждому узлу слияния представляют собой любой из возможных сигналов [4, 24, 27]. Данная модель предполагает параллельный ввод и обработку данных, которые могут входить в систему на разных уровнях слияния. Модель основана на децентрализованной архитектуре и не предполагает контроля обратной связи [24].

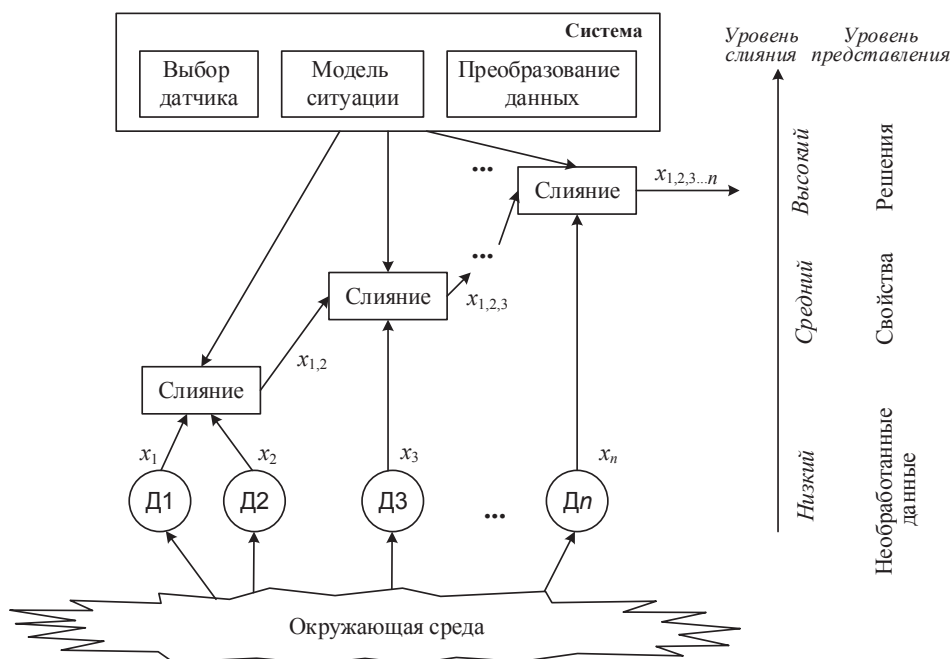


Рисунок 11 – Структура модели Luo и Kay

## 3.2 Модели на основе действий (activity-based model)

Данные модели определяются на основе определённой последовательности действий, которые указываются и должны выполняться интеллектуальной системой слияния данных.

### 3.2.1 Цикл OODA

Цикл OODA (O – observe, O – orient, D – decide, A – act), предложенный Дж. Бойдом (J.R. Boyd) [30], рассматривается в качестве единой типовой модели цикла принятия решений для систем командования и управления.

На рисунке 12 представлен цикл *OODA*, предполагающий многократное повторение четырёх последовательных взаимосвязанных действий [17, 20, 21].

- Наблюдать (*observe*). Осуществляется сбор данных из доступных различных датчиков.
- Ориентироваться (*orientate*). Осуществляется слияние собранных данных с тем, чтобы получить интерпретацию текущей ситуации.
- Решать (*decide*). Определяется план действий в ответ на понимание ситуации.
- Действовать (*act*). Выполняется план.

Данный цикл имеет обратную связь, позволяет проводить полный обзор и разделять системные задачи. Однако из-за своей структуры не обеспечивает надлежащее представление слияния данных [20], а также не показывает результат этапа «действовать» на других этапах.

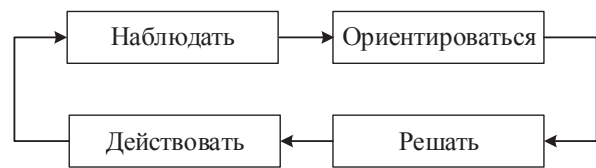


Рисунок 12 – Цикл OODA

### 3.2.2 Цикл разведки (*intelligence cycle*)

Цикл разведки описывает процесс от сбора необработанных данных до выдачи готовых обработанных данных, которые могут использоваться при принятии решений [31]. На рисунке 13 представлены четыре этапа цикла разведки.

- Сбор (*collection*). Собираются необработанные данные от датчиков.
- Сопоставление (*collation*). Собранные данные анализируются, сравниваются между собой, проводится корреляция. Неважные и ненадёжные данные отбрасываются.
- Оценка (*evaluation*). Собранные данные обрабатываются, а затем анализируются.
- Распространение (*dissemination*). Результаты слияния предоставляются пользователям для выработки решений и действий в ответ на выявленные ситуации.

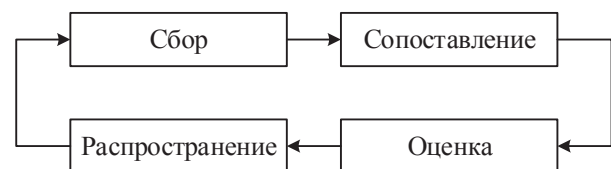


Рисунок 13 – Этапы цикла разведки

В отличие от цикла *OODA*, цикл разведки не делает явными этапы решения и исполнения, которые включены в этапы оценки и распространения. Цикл разведки является общим и может быть использован в любой области применения. Однако он не выполняет конкретные задачи, касающиеся слияния данных.

### 3.2.3 Комплексная модель (*Omnibus Model*)

Комплексная модель, предложенная М. Бедворсом (*M.D. Bedworth*) и Дж. Брайном (*J.C. Brien*) [32], организует этапы слияния данных точно так же как цикл разведки и *OODA*. Подробное описание модели представлено в работах [17, 21, 25, 33]. На рисунке 14 представлена структура комплексной модели, состоящая из четырёх этапов.

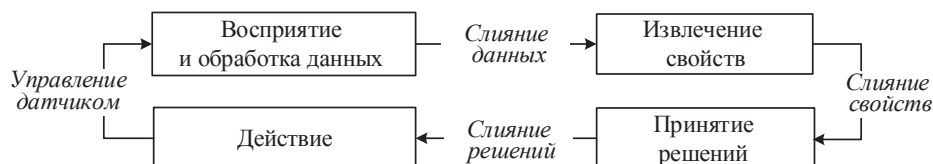


Рисунок 14 – Структура комплексной модели

- Восприятие и обработка данных, поступающих от различных датчиков.
- Извлечение свойств из собранных данных, которые затем сливаются.

- Принятие решений о наиболее подходящем плане действий.
- Действие по выполнению выбранного плана.

Комплексная модель является циклической по структуре, при этом может быть использована несколько раз для одного и того же приложения. Однако она не поддерживает декомпозицию задач на модули, поэтому её необходимо реализовывать и использовать отдельно для различных применений.

### 3.3 Модели на основе ролей (role-based model)

Данные модели определяются на основе ролей слияния и взаимосвязей между ними, тем самым обеспечивая более детальную модель для интеллектуальных систем слияния данных и системный подход к слиянию данных, однако при этом не указываются задачи слияния. Вместо этого для слияния данных они предоставляют набор функций и определяют отношения между ними [21].

#### 3.3.1 Объектно-ориентированная модель (object-oriented model)

Для слияния данных М. Кокар (М. Kokar) предложил воспользоваться объектно-ориентированной моделью, состоящей из четырёх ролей [21, 34] (рисунок 15).

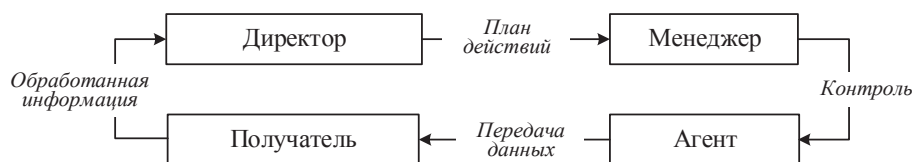


Рисунок 15 – Структура объектно-ориентированной модели

- Агент (*actor*) взаимодействует с окружающей средой, собирает данные.
- Получатель (*perceiver*) проводит анализ, оценивает полученные данные после сбора данных, а затем передаёт директору.
- Директор (*director*) строит план действий с указанием целей системы на основе анализа, предоставленного получателем.
- Менеджер (*manager*) контролирует агентов для выполнения планов, разработанных директором.

Данная модель описывает различные роли системы, однако при этом не рассматривает системные задачи.

#### 3.3.2 Архитектура Франкеля-Бедворта (Frankel-Bedworth)

Б. Франкель предложил архитектуру, состоящую из двух саморегулирующихся процессов, у которых разные цели и роли [17, 35] (рисунок 16):

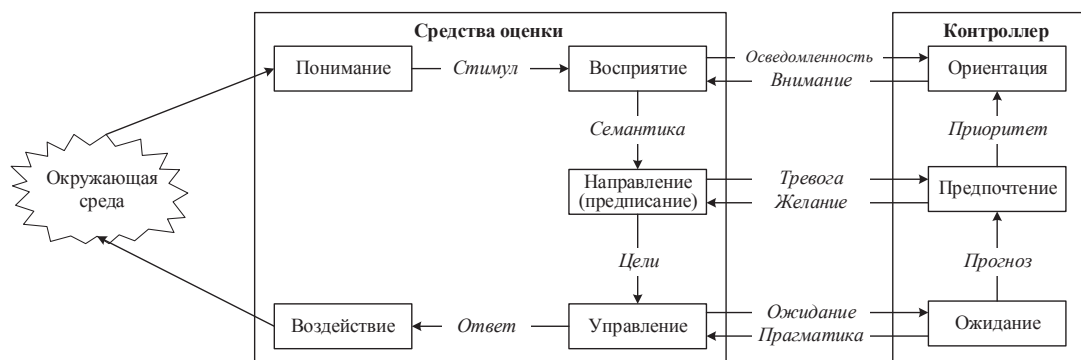


Рисунок 16 – Архитектура Frankel-Bedworth



- локальный процесс пытается достичь целей, предусмотренных глобальным процессом, и обладает ролью средства оценки, которая подобна предыдущим моделям слияния данных;
- глобальный процесс обновляет цели в соответствии с обратной связью, предоставленной локальным процессом.

### 3.3.3 Модель ситуационной осведомлённости (Situation Awareness Endsley)

Модель Эндсли (M.R. Endsley) [36] является наиболее известной моделью оценки обстановки. Она разрабатывалась для анализа деятельности операторов сложных динамических систем (самолетов, АЭС, заводов и т.п.), включая сравнительную оценку эффективности их работы, поиск причин ошибок и выработку рекомендаций по обучению для автоматизации процесса управления [20, 37].

*Ситуационная осведомлённость* означает возможность получения достаточно полной и точной информации о ситуациях в реальном режиме времени, необходимой для принятия решения и прогнозирования состояний. Согласно модели Эндсли состояние ситуационной осведомлённости является результатом процесса анализа и оценки ситуации и включает три уровня [38] (рисунок 17).

- Уровень 1 – восприятие данных и обстановки. Быстро и эффективно осуществляется сбор необходимых данных с целью получения знаний того, что происходит вокруг;
- Уровень 2 – понимание значения и важности ситуации. Осуществляется процесс слияния данных от разнородных источников с целью понимания значения собственных действий и действий других участников ситуации.
- Уровень 3 – прогнозирование будущих состояний и событий. Представление собранных и проанализированных данных в необходимой форме с целью представления сценария развития ситуации.

Модель Эндсли может применяться в различных областях, которые характеризуются высокой динамикой, большими объёмами и структурной сложностью обрабатываемых данных, необходимостью одновременного решения нескольких задач в условиях ограниченных субъективных возможностей (например, рабочей памяти), высокой степенью риска, т.е. где неудачное решение может привести к тяжелым последствиям.



Рисунок 17 – Структура модели ситуационной осведомлённости

Главные требования, которые предъявляются к системам ситуационного осведомления о состоянии сети, это: масштабируемость и возможность интеграции с существующей инфраструктурой безопасности; возможность централизованного сбора, корреляции, анализа и отображения разнородных данных от всех компонентов инфраструктуры; возможность взаимодействия специалистов различных служб; предоставление всесторонних комплексных отчётов в режиме реального времени, уведомления об угрозах, высокий уровень визуализации; возможность получения данных из внешних источников [38].

Отметим, что достаточно сложно определить универсальную модель слияния данных для конкретного использования датчика. Проектирование и разработка интеллектуальных и сенсорных систем и сбор данных с нескольких разнотипных датчиков является сложной зада-

чей. По этой причине для эффективного сбора данных и их обработки можно использовать несколько моделей слияния или их комбинации. Типичным примером является Комплексная модель, которая представляет собой сочетание трёх моделей слияния данных: цикл *OODA*, водопадная модель и *DFD*.

Проведённый анализ показал достаточно большое количество моделей слияния данных, которые могут найти широкое применение при диагностировании ТО любых отраслей промышленности и транспорта. Однако при выборе той или иной модели необходимо учитывать ряд аспектов, например: какие собраны будут данные и на каком этапе; какая преследуется цель использования той или иной модели с точки зрения требований и будущего использования; на каком этапе будут приниматься решения.

В связи с этим предлагаются следующие рекомендации: например, *JDL* модель и водопадную модель использовать для работы на высоком уровне, однако не исключено их использование и на низком уровне; модель *OODA* целесообразно применять для деятельности отдельных лиц и организаций в условиях конкурентной среды (соперничества); модель ситуационной осведомлённости использовать в случае, когда специалист, обеспечивающий эффективное управление безопасностью организации, глубоко владеет текущей обстановкой (ситуационной осведомленностью).

## **Заключение**

Слияние данных играет важную роль при диагностировании ТО, поскольку это увеличивает вероятность обнаружения неисправностей ТО в условиях множества разнотипных датчиков, а также уменьшает время реакции экспертов на различные ситуации.

Исследованы различные структурные модели слияния данных и их принципы работы. Проведено сравнение рассмотренных моделей между собой, выявлены их преимущества и недостатки. Отмечено, что для эффективного сбора исходных данных, поступающих от множества разнотипных датчиков, и их обработки можно использовать несколько моделей слияния данных либо их комбинации.

## **Благодарности**

Работа выполнена при поддержке РФФИ, проекты № 19-07-00263, 19-07-00195, 19-08-00152.

## **Список источников**

- [1] **Bahador Khaleghi** Multisensor data fusion: A review of the state-of-the-art / Bahador Khaleghi, Alaa Khamis, Fakhreddine O. Karray, Saiedeh N.R. // Information Fusion. 2013. - Vol.41. - No.1. - P. 28-44.
- [2] **Fouad, M.** Data mining and fusion techniques for WSNs as a source of the big data / M.M. Fouad, N.E. Oweis, T. Gaber, M. Ahmed, V. Snasel // Procedia Computer Science. - Elsevier, 2015. - Vol.65. - P. 778-786.
- [3] **Valet, L.A.** A statistical overview of Recent Literature in Information Fusion / L.A. Valet, G. Mauris, P. Bolon // 3rd International Conference on Information Fusion. France. 2001. - P.532-536.
- [4] **Esteban, J.** A review of data fusion models and architectures: Towards engineering guidelines / J. Esteban, A. Willetts, R. Hannah, P. Bryanston-Cross // Neural Computing & Applications. 2005. - Vol.14. - No.4. - P.273-281.
- [5] **Bleiholder, J.** Data Fusion and Conflict Resolution in Integrated Information Systems / J. Bleiholder. – Potsdam: Hasso-Plattner-Institut, 2010. - 184 p.
- [6] **Goodman, I.R.** Mathematics of Data Fusion / I.R. Goodman, R.P. Mahler, H.T. Nguyen // Kluwer Academic Publishers, 1997. - 508 p.
- [7] **Вовченко, А.Е.** Методы разрешения сущностей и слияния данных в ETL-процессе и их реализация в среде Нодоор / А.Е. Вовченко, Л.А. Калиниченко, Д.Ю. Ковалев // Информатика и ее применение. 2014. - Т.8. - Вып. 4. - С.94-109.

- [8] **Долгий, А.И.** Проблемы и методы слияния разнородных данных в гибридных интеллектуальных системах / А.И. Долгий, А.Е. Колоденкова, С.М. Ковалев // Гибридные и синергетические интеллектуальные системы: материалы IV Всероссийской Поспеловской конференции с международным участием. – Калининград: БФУ им. И. Канта, 2018. – С.181-187.
- [9] **Городецкий, В.И.** Многоагентная технология принятия решений в задачах объединения данных / В.И. Городецкий, О.В. Карсаев, В.В. Самойлов // Труды СПИИРАН. – СПб.: СПИИРАН. 2002. – Вып.1. – Т.2. – С.12-37.
- [10] **Kushwaha, M.** Feature-Level Information Fusion Methods for Urban Surveillance Using Heterogeneous Sensor Networks / M. Kushwaha // Vanderbilt University, Nashville, TN, USA, 2010. – 185 p.
- [11] **Castanedo, F.** A review of data fusion techniques / F. Castanedo // The Scientific World Journal, 2013. – 19 p.
- [12] **Afnan, Alofi** A Review of Data Fusion Techniques / Afnan Alofi, Anwaar Alghamdi, Razan Alahmadi, Najla Aljuaid, M. Hemalatha // International Journal of Computer Applications. 2017. – Vol.167. – No.7. – P.37-41.
- [13] **Afef Ben, Brahim** Solving data fusion problems using ensemble approaches. – Tunisia: University of Tunisia, 2010. – 104 p.
- [14] **Elmenreich, A.W.** Sensor fusion in time-triggered systems / A.W. Elmenreich // Technische Universität Wien, Wien 2002. – 173 p.
- [15] **Xinhan, H.** Multi-sensor data fusion structures in autonomous systems: A Review / H. Xinhan, M. Wang // International symposium on intelligent control, Houston, Texas, 2003. – P.817-821.
- [16] **Hall, D.L.** An Introduction to Multisensor Data Fusion / D.L. Hall, J. Llinas // Proceedings of the IEEE. 1998. – Vol.85. – No.1. – P.6-23.
- [17] **Nakamura, E.F.** Information fusion for wireless sensor networks: methods, models, and classifications / E.F. Nakamura, A.A. Loureiro, A.C. Frery // ACM Comput. Surv. 2007. – Vol.39. – No.3. – 55 p.
- [18] **White, F.E.** A Model for Data Fusion / F.E. White // 1st National Symposium on Sensor Fusion: process. 1988. – Vol.2. – P.20-26.
- [19] **Llinas, J.** Handbook of Multisensor Data Fusion: Theory and Practice / J. Llinas, D.L. Hall, M.E. Liggins. – London: CRC Press, 2009. – 872 p.
- [20] **Foo, P.H.** High-level Information Fusion: An Overview / P.H. Foo, G.W. Ng // J. Adv. Inf. Fusion. 2013. – Vol.8. – No.1. – P.33-72.
- [21] **Abdelgawad, A.** Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks / A. Abdelgawad, M. Bayoumi // Lecture Notes in Electrical Engineering. 2012. – Vol.118. – 119 p.
- [22] **Dasarathy, B.V.** Sensor fusion potential exploitation-innovative architectures and illustrative applications / B.V. Dasarathy // Proceedings of the IEEE. 1997. – Vol.85. – No.1. – P.24-38.
- [23] **Harris, C.J.** Multi-sensor Data Fusion in Defence and Aerospace / C.J. Harris, A. Bailey, T.J. Dodd // Aeronaut J. 1998. – Vol.102. – No.1015. – P.229-244.
- [24] **Veloso, M.** Multi-Sensor Data Fusion on Intelligent Transport Systems / M. Veloso, C. Bentos, F. Camara Pereira // MIT Portugal Transportation Systems Working Paper Series, 2009. – 18 p.
- [25] **Rawa, Adla.** Multi Sensor Data Fusion, Methods and Problems / Rawa Adla, Youssef Bazzi, Nizar Al-Holou // Proceedings of the 2013 International conference on parallel and distributed processing techniques and applications. – Las Vegas Nevada, USA. 2013. – Vol.1. – P.106-111.
- [26] **Bedworth, M.** The Omnibus Model: A New Model of Data Fusion / M. Bedworth, J. O'Brien // IEEE Aerospace and Electronic Systems Magazine. – 2000. – Vol. 15. – P. 30-36.
- [27] **Zegras, C.** Data Fusion for Travel Demand Management: State of the Practice & Prospects / C. Zegras, F. Pereira, A. Amey, M. Veloso, L. Liu, C. Bento, A. Biderman // 4th International Symposium on Travel Demand Management, 2008. – 98 p.
- [28] **Thomopoulos, S.C.** Sensor Integration and Data Fusion / S.C. Thomopoulos // Proceedings of SPIE 1198, Sensor fusion II: Human and machine strategies. – 1989. – P.178-191.
- [29] **Luo, R.C.** Multisensor Integration and Fusion: Issues and Approaches / R.C. Luo, M.G. Kay // Proceedings of 1988 Orlando Technical Symposium. 1989. – P.42-49.
- [30] **Boyd, J.R.** A discourse on winning and losing. Unpublished set of briefing slides available / J.R. Boyd // Air University Library: Maxwell AFB, Alabama. 1987. – <http://dnipogo.org/john-r-boyd/>.
- [31] **Shulsky, A.N.** Silent warfare: understanding the world of intelligence / A.N. Shulsky, G.J. Schmitt // Inc., New York, NY, 2002. – 248 p.
- [32] **Bedworth, M.D.** The omnibus model: A new model for data fusion / M.D. Bedworth, J.C. O'Brien // Proceedings of the 2nd International Conference on Information Fusion. ISIF, Sunnyvale. 1999. – P.437-444.
- [33] **Elmenreich, W.** An Introduction to Sensor Fusion / W. Elmenreich // Vienna University of Technology, Austria, 2002. – Vol.502. – 29 p.
- [34] **Kokar, M.M.** Reference Model for data fusion Systems / M.M. Kokar, M.D. Bedworth, C.B. Frankel // International Society for Optics and Photonics: In AeroSense. 2000. – P.191-202.

- [35] **Frankel, B.** Control, estimation and abstraction in fusion architectures: Lessons from human data processing / B. Frankel, M.D. Bedworth // *Proceeding of the 3rd International Conference on Data Fusion, Paris, France. 2000.* - Vol.1. - P.3-10.
- [36] **Endsley, M.R.** Toward a theory of situation awareness in dynamic systems / M.R. Endsley // *Human Factors: The J. of the Human Factors and Ergonomics Society.* 1995. - Vol.37. - No.1. - P.32-64.
- [37] **Лебедев, С.В.** Концептуальная модель подсистемы оценки обстановки интеллектуального агента реального времени / С.В. Лебедев, М.Г. Пантелеев // *Известия СПбГЭТУ «ЛЭТИ».* 2015. - №7. - С.41-46.
- [38] **Федулов, М.** Ситуационная осведомленность. Новый подход к проблемам комплексной безопасности / М. Федулов // *Безопасность: Информационное обозрение.* - 09-08-2011. - <http://www.securityinfowatch.ru/view.php?section=articles&item=4>.

## INTELLECTUAL TECHNOLOGIES OF DATA FUSION FOR DIAGNOSTICS OF TECHNICAL OBJECTS

**S.M. Kovalev<sup>1</sup>, A.E. Kolodenkova<sup>2</sup>, V. Snasel<sup>3</sup>**

<sup>1</sup>*Rostov State Transport University, Rostov-on-Don, Russia*  
*ksm@rfniias.ru*

<sup>2</sup>*Samara State Technical University, Samara, Russia*  
*anna82\_42@mail.ru*

<sup>3</sup>*VSB-Technical University of Ostrava, Ostrava, Czech Republic*  
*vaclav.snasel@vsb.cz*

### Abstract

Fusion of heterogeneous data in real time is an important task in the diagnosis of technical objects. This is due to the need of taking into account not only the data coming from the sensors, but also external factors affecting the technical object. The article addresses problems of data fusion terminology on the basis of a review of the national and foreign literature. New definition of the term "data fusion" is proposed. Different scientific views of domestic and foreign experts on a problem of data fusion for diagnosing of technical objects in presence of different types of sensors and heterogeneous information are generalized and systematized. The adapted classification of data fusion, taking into account various criteria (the relations between sensors, the level of abstraction of data, architecture type), is presented. Classification of structural models of data fusion, developed for creation of intellectual systems of data fusion is given. Models of process of data fusion are investigated, their structures are presented, model merits and demerits are revealed. It is noted that for effective collecting of the basic data, arriving from of different types of sensors, and its processing it is possible to use several models of data fusion or their combination. It will allow making scientifically based management decisions when diagnosing difficult technical objects. While a lot of works of foreign researchers are devoted to separate sections of data fusion technologies, the paper presents the first full research in Russian where all the main aspects that belong to the intellectual technologies of data fusion are considered.

**Key words:** *heterogeneous data, models of data fusion, diagnostics of technical objects, many different types of sensors.*

**Citation:** *Kovalev SM, Kolodenkova AE, Snasel V. Intellectual technologies of data fusion for diagnostics of technical objects [In Russian]. Ontology of designing. 2019; 9(1): 152-168. – DOI: 10.18287/2223-9537-2019-9-1-152-168.*

### Acknowledgment

The work was supported by RFBR grants No. 19-07-00263, 19-07-00195, 19-08-00152.

### References

- [1] **Bahador Khaleghi, Alaa Khamis, Fakhreddine OK, Saiedeh NR.** Multisensor data fusion: A review of the state-of-the-art. *Information Fusion.* 2013; 41(1): 28-44.

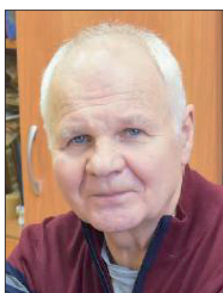


- [2] **Fouad M, Oweis NE, Gaber T, Ahmed M, Snasel V.** Data mining and fusion techniques for WSNs as a source of the big data. *Procedia Computer Science*: Elsevier. 2015; 65: 778-786.
- [3] **Valet LA, Mauris G, Bolon P.** A statistical overview of Recent Literature in Information Fusion. 3rd International Conference on Information Fusion; 2001: P. 532-536.
- [4] **Esteban J, Willetts A, Hannah R, Bryanston-Cross P.** A review of data fusion models and architectures: Towards engineering guidelines. *Neural Computing & Applications*. 2005; 14(4): 273-281.
- [5] **Bleiholder J.** Data Fusion and Conflict Resolution in Integrated Information Systems. – Potsdam: Hasso-Plattner-Institut; 2010.
- [6] **Goodman IR, Mahler RP, Nguyen HT.** Mathematics of Data Fusion. Kluwer Academic Publishers; 1997.
- [7] **Vovchenko AE, Kalinichenko LA, Kovalev DU.** Methods of entity resolution and data fusion in the ETL-process and their implementation in the Hadoop environment [In Russian]. *Informatics and her application*. 2014; 4: 94-109.
- [8] **Dolgy AI, Kolodenkova AE, Kovalev SM.** Problems and methods of heterogeneous data fusion in hybrid intelligent systems [In Russian]. Hybrid and synergetic intelligent systems: Proceedings of the IV Russian Pospelov conference with international participation. – Kaliningrad: IK BFU; 2018: 181-187.
- [9] **Gorodetski VI, Karsayev OV, Samoilov VV.** Multi-agent data fusion technology [In Russian]. SPIIRAS Proceedings. – SPb: SPIIRAS 2002; 1: 12-37.
- [10] **Kushwaha M.** Feature-Level Information Fusion Methods for Urban Surveillance Using Heterogeneous Sensor Networks. Vanderbilt University, Nashville, TN, USA; 2010.
- [11] **Castanedo F.** A review of data fusion techniques. *The Scientific World Journal*; 2013.
- [12] **Afnan Alofi, Anwaar Alghamdi, Razan Alahmadi, Najla Aljuaid, Hemalatha M.** A Review of Data Fusion Techniques. *International Journal of Computer Applications*; 2017: 167 (7): 37-41.
- [13] **Afef Ben B.** Solving data fusion problems using ensemble approaches. Tunisia: University of Tunisia; 2010: 104.
- [14] **Elmenreich AW.** Sensor fusion in time-triggered systems. Technische Universität Wien, Wien; 2002: 173.
- [15] **Xinhan H, Wang M.** Multi-sensor data fusion structures in autonomous systems. International symposium on intelligent control. Houston, Texas; 2003: 817-821.
- [16] **Hall DL, Llinas J.** An Introduction to Multisensor Data Fusion. Proceedings of the IEEE. 1998; 85(1): 6-23.
- [17] **Nakamura EF, Loureiro AA, Frery AC.** Information fusion for wireless sensor networks: methods, models, and classifications. *ACM Comput. Surv.* 2007; 39(3): 1-55.
- [18] **White FE.** A Model for Data Fusion. 1st National Symposium on Sensor Fusion: process 1988; 2: 20-26.
- [19] **Llinas J, Hall DL, Liggins ME.** Handbook of Multisensor Data Fusion: Theory and Practice. – London: CRC Press; 2009.
- [20] **Foo PH, Ng GW.** High-level Information Fusion: An Overview. *Inf. Fusion* 2013; 8 (1): 33-72.
- [21] **Abdelgawad A, Bayoumi M.** Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks. Lecture Notes in Electrical Engineering 2012; 118: 1-119.
- [22] **Dasarathy BV.** Sensor fusion potential exploitation-innovative architectures and illustrative applications. Proceedings of the IEEE 1997; 85 (1): 24-38.
- [23] **Harris CJ, Bailey A, Dodd TJ.** Multi-sensor Data Fusion in Defence and Aerospace. *Aeronaut J.* 1998; 102(1015): 229-244.
- [24] **Veloso M, Bentos C, Camara Pereira F.** Multi-Sensor Data Fusion on Intelligent Transport Systems. MIT Portugal Transportation Systems Working Paper Series; 2009.
- [25] **Rawa Adla, Youssef Bazzi, Nizar Al-Holou.** Multi Sensor Data Fusion, Methods and Problems. Proceedings of the 2013 International conference on parallel and distributed processing techniques and applications. Las Vegas Nevada, USA 2013; 1.1:106-111.
- [26] **Bedworth M, O'Brien J.** The Omnibus Model: A New Model of Data Fusion. *IEEE Aerospace and Electronic Systems Magazine* 2000; 15: 30-36.
- [27] **Zegras C, Pereira F, Amey A, Veloso M, Liu L, Bento C, Biderman A.** Data Fusion for Travel Demand Management: State of the Practice & Prospects. 4th International Symposium on Travel Demand Management; 2008.
- [28] **Thomopoulos SC.** Sensor Integration and Data Fusion. Proceedings of SPIE 1198, Sensor fusion II: Human and machine strategies; 1989: 178-191.
- [29] **Luo RC, Kay MG.** Multisensor Integration and Fusion: Issues and Approaches. Proceedings of 1988 Orlando Technical Symposium; 1989: 42-49.
- [30] **Boyd JR.** A discourse on winning and losing. Unpublished set of briefing slides available. Air University Library: Maxwell AFB, Alabama 1987. - <http://dnipogo.org/john-r-boyd/>.
- [31] **Shulsky AN, Schmitt GJ.** Silent warfare: understanding the world of intelligence. Inc., New York, NY; 2002.
- [32] **Bedworth MD, O'Brien JC.** The omnibus model: A new model for data fusion. Proceedings of the 2nd International Conference on Information Fusion. ISIF, Sunnyvale; 1999: 437-444.
- [33] **Wilfried E.** An Introduction to Sensor Fusion. Vienna University of Technology, Austria 2002; 502: 29.



- [34] **Kokar MM, Bedworth MD, Frankel CB.** Reference Model for data fusion systems. International Society for Optics and Photonics: In AeroSense; 2000: 191-202.
- [35] **Frankel B, Bedworth MD.** Control, estimation and abstraction in fusion architectures: Lessons from human data processing. Proceeding of the 3rd International Conference on Data Fusion, Paris, France 2000; 1: 3-10.
- [36] **Endsley MR.** Toward a theory of situation awareness in dynamic systems. Human Factors: The J. of the Human Factors and Ergonomics Society. 1995; 37(1): 32-64.
- [37] **Lebedev SV, Panteleev MG.** A conceptual model of the subsystem of assessment of the situation an intelligent agent real-time [In Russian]. Proceedings of Saint Petersburg Electrotechnical University «LETI». 2015; 7: 41-46.
- [38] **Fedulov M.** Situational awareness. A new approach to integrated security [In Russian]. Security: informational review. 09-08-2011. - <http://www.securityinfowatch.ru/view.php?section=articles&item=4>.

## Сведения об авторах



**Ковалев Сергей Михайлович**, 1954 г. рождения. Окончил Таганрогский радиотехнический институт в 1976 г., д.т.н. (2002). Профессор кафедры «Автоматика и телемеханика на железнодорожном транспорте» Ростовского государственного университета путей сообщения. Член Российской ассоциации искусственного интеллекта. В списке научных трудов более 200 работ в области интеллектуальных и нечётких систем, нечётко-темпоральных баз знаний, мягких вычислений и экспертной поддержки принятия решений.

**Sergey Mikhailovich Kovalev** (b. 1954) graduated from the Taganrog radio engineering Institute (Taganrog-city) in 1976, D. Sc. Eng. (2002). He is Professor at Rostov State Transport University (Department of Automation and Remote Control at the Railway on Transport). He is member of Russian Association of Artificial Intelligence. He is co-author of about 200 scientific articles

and abstracts in the field of intelligent and fuzzy systems, fuzzy temporal knowledge bases, soft computing, expert and decision support.



**Колоденкова Анна Евгеньевна**, 1982 г. рождения. Окончила Уфимский государственный авиационный технический университет в 2004 г., д.т.н. (2017). Заведующая кафедрой «Информационные технологии» Самарского государственного технического университета. Член Российской ассоциации искусственного интеллекта, заместитель председателя научного совета Российской ассоциации нечётких систем и мягких вычислений. В списке научных трудов более 150 работ в области атомной энергетики, программной инженерии, системного анализа, интеллектуальных систем, мягких вычислений и экспертной поддержки принятия решений.

**Anna Evgenyevna Kolodenkova** (b. 1982) graduated from the Ufa State Aviation Technical University (Ufa-city) in 2004, D. Sc. Eng. (2017). She is Associate Professor, Head of «Information technologies» Department at Samara State Technical University. She is member of Russian Association of Artificial Intelligence and Deputy Chairman of the scientific Council of the Russian Association of fuzzy systems and soft computing. She is co-author of about 150 scientific articles and abstracts in the field of nuclear energy, software engineering, system analysis, intelligent systems, soft computing and expert decision support.



**Вацлав Снасель**, 1957 г. рождения. Профессор (2006), получил докторскую степень в области алгебры и геометрии в Масариковом университете в Брно Чешской Республики. Ректор Остравского технического университета. Является членом IEEE, ACM, AMS и SIAM. В списке научных трудов более 400 работ, посвященных искусственному интеллекту, управлению знаниями, поиску информации, интеллектуальному анализу данных, которые применяются к различным проблемам реального мира.

**Vaclav Snasel** (b. 1957). Professor (2006), he received Ph.D. degree in Algebra and Geometry from Masaryk University, Brno, Czech Republic. Rector of VSB-Technical University of Ostrava. He is a member of IEEE, ACM, AMS and SIAM. He is co-author of about 400 scientific articles. He works in a multi-disciplinary environment involving artificial intelligence, knowledge management, information retrieval, data mining and applied to various real-world problems.

## СОДЕРЖАНИЕ

<b>От редакции</b>	<b>5-6</b>
<b>ВЫХОД ЗА ГРАНИЦЫ ИНДУСТРИИ</b>	
<b>Г.Б. Евгеньев</b> (Москва, Россия)	<b>7-23</b>
ИНДУСТРИЯ 5.0 КАК ИНТЕГРАЦИЯ ИНТЕРНЕТА ЗНАНИЙ И ИНТЕРНЕТА ВЕЩЕЙ	
<b>Ю.П. Похабов</b> (Железнодорожск, Россия)	<b>24-35</b>
ПРОЕКТИРОВАНИЕ СЛОЖНЫХ ИЗДЕЛИЙ С МАЛОЙ ВЕРОЯТНОСТЬЮ ОТКАЗОВ В УСЛОВИЯХ ИНДУСТРИИ 4.0	
<b>А.А. Шарипбай, Ж.С. Сауханова, Г.Б. Шахметова, М.С. Сауханова</b> (Астана, Казахстан)	<b>36-49</b>
ОНТОЛОГИЯ КОНЕЧНО-АВТОМАТНОЙ КРИПТОГРАФИИ	
<b>А.А. Муромский, Н.П. Тучкова</b> (Москва, Россия)	<b>50-69</b>
ПРЕДСТАВЛЕНИЕ МАТЕМАТИЧЕСКИХ ПОНЯТИЙ В ОНТОЛОГИИ НАУЧНЫХ ЗНАНИЙ	
<b>Ю.В. Рогущина</b> (Киев, Украина)	<b>70-84</b>
ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДЛЯ СОЗДАНИЯ СОВРЕМЕННЫХ ЭНЦИКЛОПЕДИЧЕСКИХ ПОРТАЛОВ	
<b>Н.М. Боргест, А.С. Галахарь, М.В. Овсянников, Р.О. Самсонов</b> (Москва, Россия; Самара, Россия)	<b>85-100</b>
ПРЕДПРОЕКТНЫЙ АНАЛИЗ ИНТЕЛЛЕКТУАЛЬНОГО ЖИЛОГО ДОМА ДЛЯ УСЛОВИЙ АРКТИКИ	
<b>С.В. Лебедев</b> (Санкт-Петербург, Россия)	<b>101-116</b>
МОДЕЛЬ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ПОСТРОЕНИЮ СВЯЗАННЫХ ДАННЫХ НА ОСНОВЕ РАЗНОРОДНЫХ ИСТОЧНИКОВ	
<b>О.Н. Тушканова, В.В. Самойлов</b> (Санкт-Петербург, Россия)	<b>117-131</b>
KNOWLEDGE NET: МОДЕЛЬ И СИСТЕМА НАКОПЛЕНИЯ, ПРЕДСТАВЛЕНИЯ И ИСПОЛЬЗОВАНИЯ ЗНАНИЙ И ДАННЫХ	
<b>В.В. Голенков, Н.А. Гулякина, И.Т. Давыденко, Д.В. Шункевич, А.П. Еремеев</b> (Минск, Республика Беларусь; Москва, Россия)	<b>132-151</b>
ОНТОЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ ГИБРИДНЫХ СЕМАНТИЧЕСКИ СОВМЕСТИМЫХ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ НА ОСНОВЕ СМЫСЛОВОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ	
<b>С.М. Ковалев, А.Е. Колоденкова, В. Снасель</b> (Ростов-на-Дону, Россия; Самара, Россия; Острава, Чешская Республика)	<b>152-168</b>
ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ СЛИЯНИЯ ДАННЫХ ПРИ ДИАГНОСТИРОВАНИИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ	

**Информационный партнер AI CONFERENCE**

**9 апреля 2019, Москва**

**«ТЕХНОПОЛИС «МОСКВА»**

<https://aiconference.ru/ru> © AI Conference



*Ontologists and designers of all countries and subject areas, join us!*



**Издательство «Новая техника» - Publisher «New Engineering» Ltd.**

Россия, 443010, Самара, ул.Фрунзе 145 - 145, Frunze Str., Samara, 443010, Russia