

ПРОЕКТИРОВАНИЕ И АВТОМАТИЧЕСКОЕ ПОСТРОЕНИЕ ГРАФА ЗНАНИЙ «МАТЕМАТИЧЕСКИЕ УРАВНЕНИЯ»

Е. К. Липачев¹ [0000-0001-7789-2332], Б. Р. Мурадымов² [0009-0004-1187-8158]

^{1,2}Казанский (Приволжский) федеральный университет, г. Казань, Россия

¹Университет Иннополис, г. Иннополис, Россия

¹elipachev@gmail.com, ²muradymov.bulat@mail.ru

Аннотация

Предложен подход к проектированию и реализации графа знаний для представления и хранения знаний о математических уравнениях. В сформированном прототипе графа представлены знания об основных типах алгебраических уравнений, обыкновенных дифференциальных уравнений, уравнениях в частных производных и интегральных уравнениях. Граф знаний проектировался как математический артефакт экосистемы цифровой математической библиотеки Lobachevskii-DML, поэтому учитывались общие для экосистемы требования совместимости. Разработаны программные инструменты извлечения и обработки информации об уравнениях, представленной в цифровых библиотеках и электронных научных ресурсах. Прототип графа знаний сформирован на основе онтологии профессиональной математики OntoMath^{PRO} и таксономии уравнений, построенной на основе информации, извлеченной с веб-страниц научно-образовательного портала EqWorld «Мир математических уравнений». Онтология OntoMath^{PRO} расширена новыми классами уравнений и новыми отношениями для согласования с иерархией типов уравнений, представленной на портале EqWorld. Реализован комплекс программных модулей, обеспечивающих полный цикл формирования графа знаний: автоматическое извлечение сущностей из источников, связывание сущностей с концептами онтологии OntoMath^{PRO}, преобразование полученных знаний в RDF-представление с последующим сохранением в хранилище данных с возможностью выполнения SPARQL-запросов.

Ключевые слова: *граф знаний, извлечение знаний, математическое уравнение, математическая онтология, представление математического знания.*

ВВЕДЕНИЕ

Как известно, математические уравнения являются важнейшей составляющей математических документов. Алгебраические уравнения, обыкновенные дифференциальные уравнения, уравнения в частных производных, нелинейные уравнения, интегральные уравнения – это неполный список основных типов математических уравнений. Для каждого уравнения факты о нем в виде теорем, свойств, методов точного и приближенного решений представлены в многочисленных, не связанных между собой источниках, включая справочники, монографии и статьи. Значительная часть источников оцифрована и доступна в электронном варианте для чтения и обработки (см., например, [1–5]). Для формирования пространства математического знания как составляющей единого пространства научных знаний необходимо, в частности, решить задачи структурирования и семантического представления знаний (см., например, [6]). Создание математических онтологий направлено на решение этих задач (см., например, [7–9]). В работах [7, 10] проведены исследования по применению онтологий в прикладных задачах, представлены примеры использования математических онтологий в качестве компонент рекомендательных и поисковых систем.

Эффективным подходом к интеграции разнородных данных в единое пространство научных знаний являются графы знаний, в которых знания представлены в виде ориентированного графа сущностей и отношений между ними. Эффективность этого подхода подтверждается развитием глобального облака открытых связанных данных (Linked Open Data Cloud, <https://lod-cloud.net/>).

Термин «граф знаний» был введен в официальном блоге Google для обозначения использования в веб-поиске знаний, полученных с помощью связей между объектами [11]. Минимальный набор характеристик, позволяющий отличать графы знаний от других наборов знаний, описан в [12].

Для представления графов знаний используется стандарт консорциума W3C для моделирования и обмена данными Resource Description Framework (RDF, <https://www.w3.org/RDF/>), а также графы свойств (Property Graphs) (см., например, [13]).

Отметим ряд исследований, связанных с проектированием графов знаний в математической области. Подход к формированию графа знаний на основе онтологического описания научных предметных областей представлен в [14]. В качестве источника структурированных данных использована семантическая библиотека LibMeta (<https://libmeta.ru/>). Онтология содержимого семантической библиотеки применена в качестве средства формализации.

Технология построения графов знаний для необработанных научных текстов представлена в [15]. Предложена семантическая модель предметной области междисциплинарного научного журнала на основе онтологии библиотеки LibMeta. В этой работе показано, как можно перейти от неструктурированных текстов к тематическому анализу и встраиванию в граф знаний LibMeta.

В [16] предложен метод обогащения графа знаний и расширения предметной онтологии с помощью больших языковых моделей. В [17] представлены исследования процесса адаптации LLM к предметной области обыкновенных дифференциальных уравнений, представленной в виде онтологии русскоязычных ресурсов. Предложен интеграционный подход, рассматривающий ограничения больших языковых моделей в рамках библиотеки LibMeta, а также оценку релевантности ответа для различных LLM.

Настоящая работа является продолжением исследований, представленных в [18]. Предложены архитектура графа знаний «Математические уравнения» и алгоритм его реализации, основанный на извлечении знаний из научно-образовательного портала EqWorld «Мир математических уравнений» [2] и онтологии профессиональной математики OntoMath^{PRO} [7]. Использовались также сведения об уравнениях, представленных в справочниках по обыкновенным дифференциальным уравнениям и дифференциальным уравнениям в частных производных [19, 20], справочники по уравнениям математической физики [21, 22], справочник по интегральным уравнениям [23], а также переводы указанных книг на английский язык [24–28]. Для пополнения онтологии, которое также выполнено в работе, использовались сведения, представленные в Математической энциклопедии под редакцией И. М. Виноградова в пяти томах [29], а также перевод этой энциклопедии на английский язык [5, 30].

1. ПОСТАНОВКА ЗАДАЧИ

Настоящая работа посвящена разработке метода автоматического построения графа знаний «Математические уравнения» для представления и хранения знаний об основных типах математических уравнений.

Определение (см., например, [31]). *Граф знаний – это кортеж $G = (E, R, T, D)$, где E – множество вершин, представляющих сущности предметной области, R – множество отношений, связывающих сущности, T – множество RDF-триплетов $(s, p, o) \in E \times R \times E$, s – субъект, p – предикат, o – объект, D – множество описаний сущностей и отношений.*

Проектирование графа знаний «Математические уравнения» (далее «граф знаний») предполагает решение задач по нескольким взаимосвязанным направлениям:

- определение узлов графа знаний;
- определение отношений между узлами графа знаний;
- разработка программных инструментов для извлечения и импорта именованных сущностей и фактов из внешних источников;
- построение системы RML-правил отображения структурированных данных в RDF-формат;
- автоматическая генерация RDF-триплетов и формирование на их основе графа знаний;
- организация хранилища графа знаний с поддержкой SPARQL-запросов.

2. МЕТОД ФОРМИРОВАНИЯ ГРАФА ЗНАНИЙ

2.1. Архитектура системы построения графа знаний

На рис. 1 представлена архитектура системы автоматического построения графа знаний «Математические уравнения». Указаны основные источники знаний, а также основные модули, участвующие в формировании графа знаний.

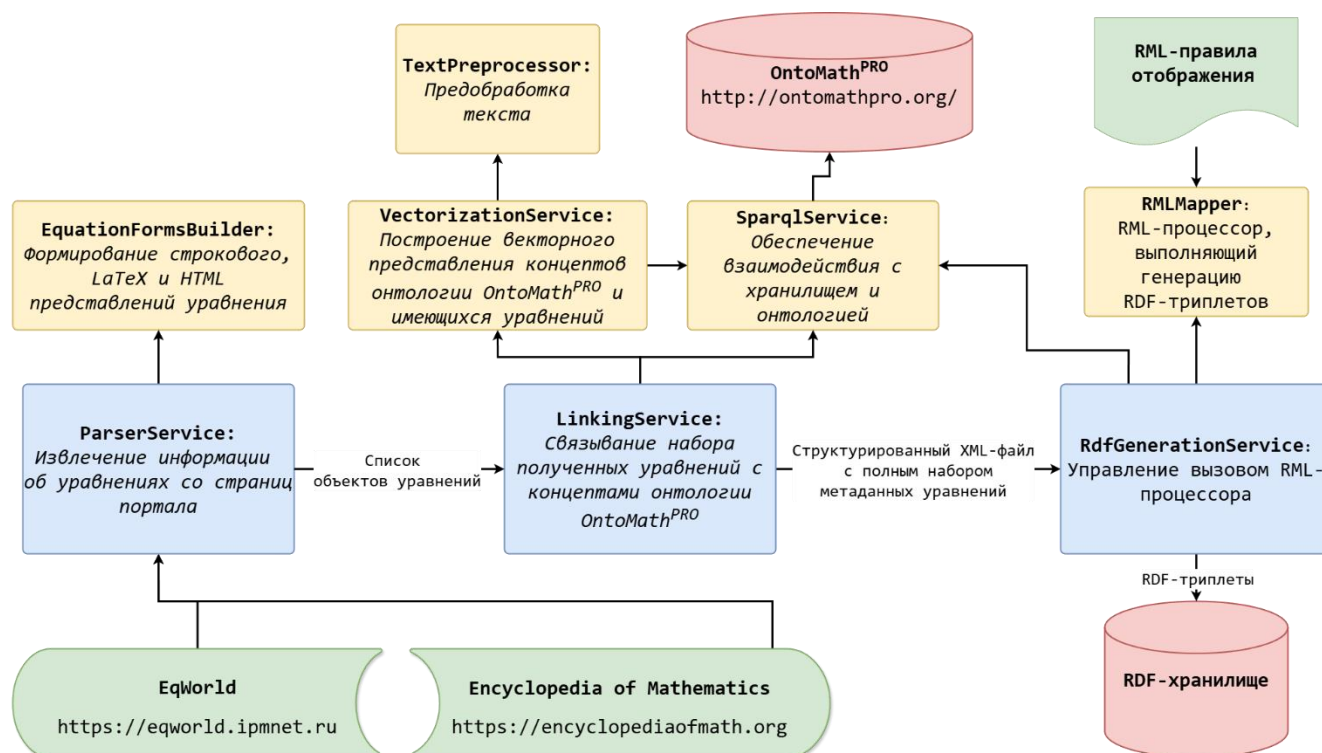


Рис. 1. Архитектура системы построения графа знаний «Математические уравнения».

2.2. Алгоритм извлечения знаний об уравнениях с портала EqWorld

Международный научно-образовательный портал «Мир математических уравнений» EqWorld содержит информацию о примерно 800 уравнениях, включая сведения о методах их решения и ссылки на внешние источники [2]. Представлена иерархия уравнений, распределенных по типам в 38 разделах и 60 подразделах.

Для автоматического сбора данных об уравнениях разработан алгоритм последовательного обхода 2000 веб-страниц портала, извлечения сведений об уравнениях и преобразования полученных данных в T_EX-представление с последующим формированием объекта уравнения.

Алгоритм извлечения информации об уравнениях с портала EqWorld представлен в виде блок-схемы на рис. 2. Производится последовательный обход веб-страниц портала EqWorld с автоматическим поиском и извлечением данных о математических уравнениях, представленных на этих страницах. Далее орга-

низуется цикл, в котором обрабатывается информация о каждом уравнении. Извлекаются типы и подтипы уравнения, их названия, ссылки на PDF-документы. Производится преобразование извлеченной информации в $L^A T_E X$ - и HTML- представления. В результате формируется список объектов уравнений.

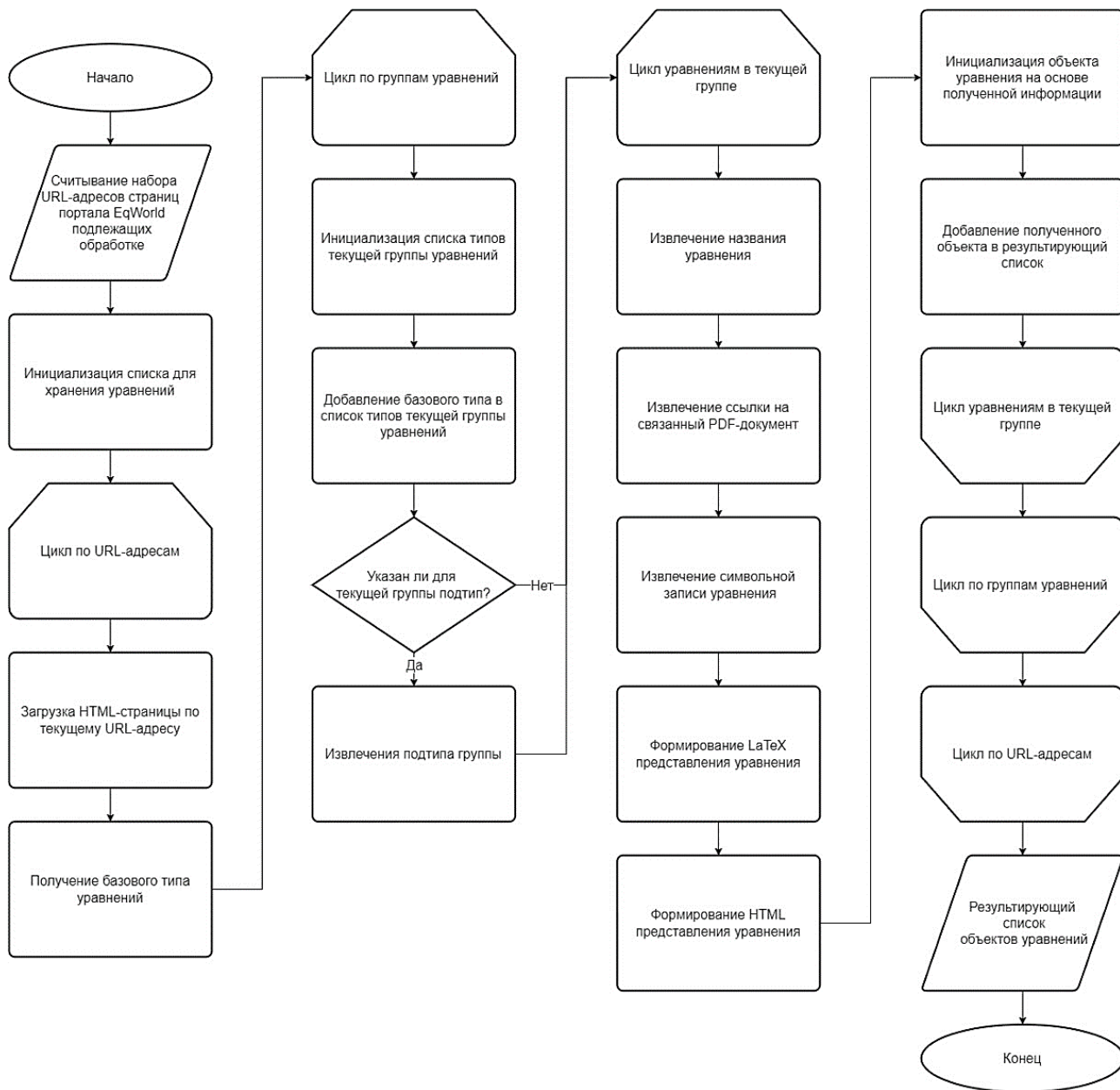


Рис. 2. Блок-схема алгоритма извлечения информации об уравнениях с веб-страниц портала EqWorld.

Далее представлен псевдокод алгоритма извлечения информации об уравнениях с веб-страниц портала EqWorld (алгоритм 1).

Алгоритм 1. Алгоритм извлечения информации об уравнениях.

Входные данные: набор URL-адресов страниц, подлежащих обработке

Выходные данные: Список уравнений

Чтение набора url-адресов из текстового файла

```
1  urls = GetUrls()
   # Инициализация результирующего списка
2  result = []
3  for url in urls:
   # Загрузка html-страницы по текущему url-адресу
4  htmlDoc = Load(url)
   # Извлечение базового типа уравнений (например, Обыкновенные дифференциальные уравнения первого порядка)
5  baseType = GetBaseType(htmlDoc)
   # Извлечение групп уравнений
6  eqGroups = GetEquationsGroups(htmlDoc)
7  for eqGroup in eqGroups:
   # Формирование списка типов группы уравнений. Например, {Линейные обыкновенные дифференциальные уравнения второго порядка, Дифференциальные уравнения, содержащие степенные функции}
8  types = GetGroupTypes(baseType, eqGroup)
   # Получение списка уравнений группы
9  groupEquations = GetGroupEquations(eqGroup)
10 for equationContainer in groupEquations:
   # Извлечение ссылки на связанный pdf-документ
11 pdfLink = GetPdfLink(equationContainer)
   # Формирование строкового, LaTeX и html представлений уравнений
12 stringForm, latexForm, htmlForm = EquationFormsBuilder.GetEquationForms(equationContainer)
   # Извлечение названий уравнения
13 label = GetLabel(equationContainer)
   # Создание объекта уравнения, на основе извлеченной информации
14 equationObject = new Equation(
    types,
    pdfLink,
    label,
```

```
        stringForm,  
        latexForm,  
        htmlForm)  
    # Добавление объекта уравнений в результирующий список  
15     result.Add(equationObject)  
16     end for  
17     end for  
18 end for  
19 return result
```

2.3. Алгоритм извлечения знаний из онтологии OntoMath^{PRO}

Онтология математического знания OntoMath^{PRO} разработана с целью классификации и систематизации основных понятий профессиональной математики [7, 32]. В этой онтологии представлены основные разделы математики, включая, в частности, таксономии «Уравнение», «Элемент теории дифференциальных уравнений», «Уравнение математического анализа», «Уравнение численного анализа», используемые в качестве источников фактов об уравнениях в процессе формирования графа знаний [33].

Для соответствия типов уравнений, представленных на портале EqWorld, онтология OntoMath^{PRO} была пополнена новыми классами уравнений. Созданы новые отношения в онтологии: `omp2:ParticularSolution`, `omp2:ExactSolution`. Эти связи позволили связать каждое уравнение с методами его решения и точным решением.

Разработан алгоритм связывания информации о классах уравнений, представленной в онтологии OntoMath^{PRO}, с информацией об уравнениях, извлеченной с портала EqWorld. Под связыванием понимается поиск в онтологии концепта, подходящего для конкретного уравнения, присваивание объекту уравнения URI найденного концепта и присоединение к объекту уравнения сведений, полученных из онтологии.

Алгоритм состоит из двух этапов. На первом этапе производится поиск подходящего концепта онтологии на основе точного соответствия названия или названий типов связываемого уравнения (см. Алгоритм 2).

Алгоритм 2. Метод прямого поиска.

Входные данные: Объект уравнения

Выходные данные: Логическое значение (*true*\false)

Если у уравнения есть название, но изначально осуществляем поиск по нему

```
1  if equation.Label != 'Без названия':
    # Запрос к онтологии на поиск концепта по переданному названию
2  result = SparqlService.GetConceptByLabel(equation.Label)
    # Если концепт найден
3  if result != null:
    # Дополнение объекта уравнения информацией, имеющейся в онто-
    # логии
4    CompleteEquation(equation, result)
5    return true
    # Если поиск по названию не удался, то последовательно выполняется
    # поиск по названиям типов уравнения
6  for eqType in equation.Types:
7    result = SparqlService.GetConceptByLabel(eqType)
8    if result != null:
9        CompleteEquation(equation, result)
10       return true
    # Поиск не дал результатов
11 return false
```

Если поиск не дает результата, например, из-за различий в названии типа уравнения, на втором этапе выполняется процедура семантического сопоставления названий. Концепты онтологии предварительно векторизуются по схеме TF-IDF. Затем из названия и типов уравнения формируется текстовый документ, который проходит аналогичную векторизацию, после это наиболее близкий концепт определяется по косинусной мере сходства.

Блок-схема алгоритма представлена на рис. 3. По этому алгоритму производится последовательная обработка входного списка уравнений с целью сопоставления каждого уравнения с соответствующим концептом онтологии. Для каждого уравнения сначала предпринимается попытка прямого поиска кон-

цепта по названию уравнения, а затем выполняется поиск по его типам и подтипам, перебираемым от частного к общему. В случае успешного нахождения концепта уравнению присваиваются его URI и дополнительные сведения из онтологии. Если же подходящий концепт не был найден, применяется метод семантического сопоставления: из названия и типов уравнения формируется текстовый документ, к которому применяются предобработка и векторизация на основе схемы TF-IDF, далее вычисляется значение меры сходства и определяется наиболее близкий концепт онтологии. В результате формируется XML-файл, содержащий список уравнений.

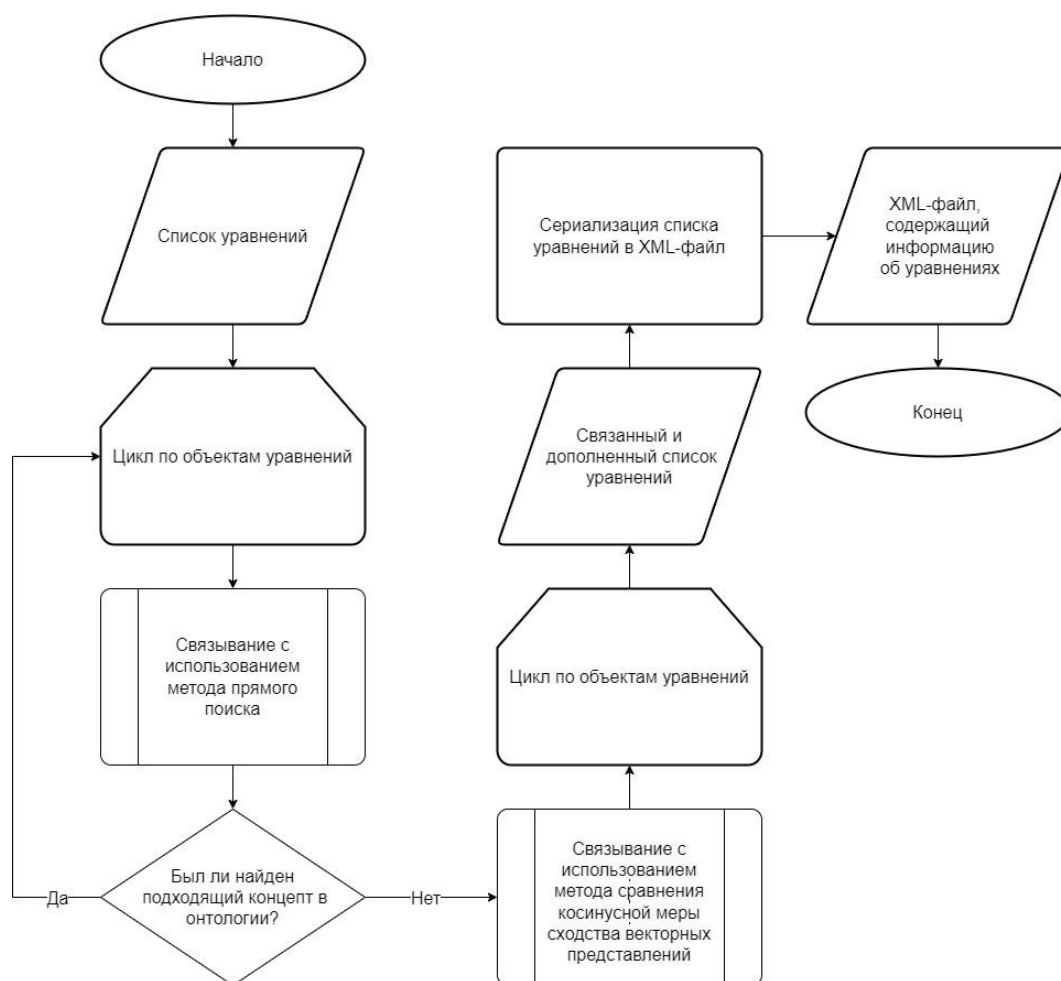


Рис. 3. Блок-схема алгоритма связывания объектов уравнений с концептами онтологии OntoMath^{PRO}.

Ниже представлен псевдокод алгоритма связывания объектов уравнений с концептами онтологии (Алгоритм 3).

Алгоритм 3. Метод связывания объектов уравнений с концептами онтологии

Входные данные: Список объектов уравнений

Выходные данные: XML-файл, содержащий информацию об уравнениях, размеченных в терминах концептов онтологии

```
1 for equation in equations:
    # Изначально происходит попытка прямого поиска
2 if TryDirectLinking(equation):
3     continue
    # Получение наиболее близкого концепта, на основе оценки косинус-
    # ной меры сходства векторных представлений концептов онтологии и
    # объекта уравнения
4 closestConcept = Vectorizer.GetClosestConcept(equation)
    # Извлечение информации по найденному концепту
5 conceptInfo = SparqlService.GetConceptInfo(closestConcept)
    # Дополнение объекта уравнения найденной информацией
6 CompleteEquation(equation, conceptInfo)
    # Сохранение результирующего списка уравнений в виде XML-файла
7 SaveAsXML(equations)
```

2.4. Алгоритм генерации RDF-триплетов

В результате выполнения операций, представленных в п. 2.2 и 2.3, формируется XML-файл, содержащий метаданные уравнений. Для каждого объекта уравнений в XML-файле, в частности, зафиксированы следующие метаданные:

- Названия уравнения на русском и английском языках (xml-тэги LabelText и LangTag);
- Тип уравнения (xml-тэг Types);
- Строковое, HTML и L^AT_EX представления формульной записи уравнения (тэги StringFrom, HtmlForm, LatexForm);
- URI родительского и рассматриваемого концептов онтологии (SubClassOf, LinkedOmpConcept);
- Описание свойств уравнения (тэг Comments).

Фрагмент XML-файла приведен на рис. 4.

Преобразование XML-данных в RDF-представление выполняется на основании разработанной системы правил, записанной на языке RML (RDF Mapping

Language) [34]. Система RML-правил задает карты троек, однозначно определяющие субъект, предикат и объект каждого генерируемого триплета (фрагмент карты представлен на рис. 5).

Набор правил отображения и XML-файл уравнений передаются RML-процессору, который генерирует триплеты и записывает их файл в формате Turtle (фрагмент файла представлен на рис. 6).

```

<Equation>
  <URI>9ef5f5b0-e176-462d-9d70-5a54d54bc895</URI>
  <Labels>
    <Label>
      <LabelText>Уравнение Абеля второго рода</LabelText>
      <LangTag>ru</LangTag>
    </Label>
    <Label>
      <LabelText>Abel's differential equation of the second kind</LabelText>
      <LangTag>en</LangTag>
    </Label>
    <Label>
      <LabelText>Дифференциальное уравнение Абеля 2-го рода</LabelText>
      <LangTag>ru</LangTag>
    </Label>
    <Label>
      <LabelText>Абеля дифференциальное уравнение 2-го рода</LabelText>
      <LangTag>ru</LangTag>
    </Label>
  </Labels>
  <Types>
    <string>Обыкновенные дифференциальные уравнения первого порядка</string>
  </Types>
  <SubClassOf>http://ontomathpro.org/omp2#AbelDifferentialEquation</SubClassOf>
  <LinkedOmpConcept>http://ontomathpro.org/omp2#Abel%27sDifferentialEquationOfTheSecondKind<
  <StringForm>yy'=f(x)y^(2)+g(x)y+h(x).</StringForm>
  <Reference>https://eqworld.ipmnet.ru/en/solutions/ode/ode0126.pdf</Reference>
  <Comments>
    <string>Abel's differential equation of the first kind is an equation of the form
  $$
  y' = f_0(x) + f_1(x)y + f_2(x)y^2 + f_3(x)y^3.
  $$

```

Abel's differential equations of the first kind represent a natural generalization of the Ricc

Рис. 4. Фрагмент XML-файла с информацией об уравнении Абеля 2-го рода.

```

<#EquationMapping> a rr:TriplesMap;
  rml:logicalSource [
    rml:source "equations.xml" ;
    rml:iterator "/ArrayOfEquation/Equation";
    rml:referenceFormulation ql:XPath;
  ];

  rr:subjectMap [
    rr:template "http://www.eqgraph.ru/{URI}";
  ];

  rr:predicateObjectMap [
    rr:predicate rdf:type;
    rr:objectMap [
      rml:reference "LinkedOmpConcept";
      rr:termType rr:IRI;
    ]
  ];

```

Рис. 5. Фрагмент карты троек, в которой определены логический источник, карта субъекта и карта предиката-объекта.

```

<http://www.eqgraph.ru/9ef5f5b0-e176-462d-9d70-5a54d54bc895> a <http://ontomathpro.org/omp2#Abel%27sDifferentialEquationOfTheSecondKind>;
eq:htmlForm "<i>y</i><span>&prime;</span> = <i>f</i><span><i>x</i></span><i>y</i><sup>2</sup> + <i>g</i><span><i>y</i></span><i>y</i></span>";
eq:latexForm "\\left[y\\frac{dy}{dx}=f(x)y^2+g(x)y+h(x)\\right]";
eq:reference "https://eqworld.ipmnet.ru/en/solutions/ode/ode0126.pdf";
eq:stringForm "yy'=f(x)y^2+g(x)y+h(x).";
rdfs:comment ""Abel's differential equation of the second kind is an equation of the form
$$
\\left(g_0(x) + g_1(x)y \\right)y' = f_0(x) + f_1(x)y + f_2(x)y^2 + f_3(x)y^3.
$$
If $g_0, g_1 \\in C^1(a,b)$ and $g_1(x) \\neq 0$, $g_0(x) + g_1(x)y \\neq 0$, Abel's differential equation of the second kind can be reduced
Abel's differential equations of the first and second kinds, as well as their further generalizations"";
rdfs:label "Abel's differential equation of the second kind"@ru, "Абеля дифференциальное уравнение 2-го рода"@ru,
"Дифференциальное уравнение Абеля 2-го рода"@ru, "Уравнение Абеля второго рода"@ru;

```

Рис. 6. Фрагмент файла, содержащего сгенерированные RDF-триплеты для сущности «Уравнение Абеля второго рода».

Блок-схема алгоритма генерации RDF-триплетов представлена на рис. 7. Производится преобразование структурированных данных об уравнениях в формат RDF с последующей записью в граф знаний. Для этого формируются RML-правила отображения, на основе которых настраивается RML-процессор, получающий на вход XML-файл с описанием уравнений. После вызова процессора сгенерированные RDF-триплеты сохраняются в хранилище Open Link Virtuoso в виде графа знаний. Созданное хранилище предоставляет возможность выполнения SPARQL-запросов.

На Рис. 8 приведен фрагмент графа знаний, представляющий узлы и отношения для сущности «Уравнение Абеля второго рода».

уравнениях. Для извлечения данных с веб-страниц (веб-скрейпинг) используются функции библиотеки HTMLAgilityPack (<https://html-agility-pack.net/>). Каждое обнаруженное уравнение представляется экземпляром класса Equation.

Модуль LinkingService обеспечивает дополнение данных с помощью сопоставления уравнений с концептами онтологии OntoMath^{PRO}. Для векторизации и вычисления семантической близости задействована библиотека машинного обучения Accord.Net (<https://accord-framework.net/>). Взаимодействие с онтологией осуществляется через систему сформированных SPARQL-запросов.

Модуль RdfGenerationService координирует работу RML-процессора: передает ему XML-файл с описаниями уравнений и набор правил отображения, после чего сохраняет сгенерированное множество RDF-триплетов в хранилище графа знаний.

Вспомогательный уровень состоит из следующих программных сервисов:

- SparqlService – формирование, отправка запросов к онтологии и обработка полученных ответов;
- VectorizationService – построение TF-IDF-векторных представлений концептов онтологии;
- TextPreprocessor – предобработка текстовых данных (токенизация, лемматизация и пр.);
- EquationFormsBuilder – формирование строкового, HTML- и L^AT_EX-представлений уравнений.

ЗАКЛЮЧЕНИЕ

Предложен метод и реализован алгоритм построения графа знаний «Математические уравнения», объединяющий факты об уравнениях, представленные на научно-образовательном портале EqWorld «Мир математических уравнений» и онтологии профессиональной математики OntoMath^{PRO}. Сформированный граф включен в цифровую экосистему OntoMath [35] цифровой математической библиотеки Lobachevskii-DML [36, 37] в качестве математического артефакта, что определяет его практическую значимость.

Дальнейшее развитие работы предполагает значительное пополнение графа знаний, включая добавление новых классов уравнений, введение допол-

нительных типов отношений между ними, а также установление связей с другими объектами знаний математического пространства.

Благодарности

Выражаем благодарность Наталии Павловне Тучковой, Александру Михайловичу Елизарову и Ольге Авенировне Невзоровой за проявленный интерес к исследованию, значимые замечания и советы при оформлении статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Общероссийский портал Math-Net.Ru. URL: <https://www.mathnet.ru/> (дата обращения 22.03.2026).
 2. EqWorld. «Мир математических уравнений». URL: <https://eqworld.ipmnet.ru/> (дата обращения 22.03.2026).
 3. Цифровая математическая библиотека Lobachevskii-DML. URL: <https://lobachevskii-dml.ru/> (дата обращения 22.03.2026).
 4. Numdam, the French digital mathematics library. URL: <https://www.numdam.org/> (дата обращения 22.03.2026).
 5. Encyclopedia of Mathematics. URL: <https://encyclopediaofmath.org/> (дата обращения 22.03.2026).
 6. *Elizarov A.M., Lipachev E.K.* Lobachevskii Digital Library in the Scientific Space of Mathematical Knowledge // Scientific and Technical Information Processing. 2023. Vol. 50, No. 1. P. 35–39. <https://doi.org/10.3103/s0147688223010021>
 7. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* OntoMathPRO: an Ontology of Mathematical Knowledge // Doklady Mathematics. 2022. Vol. 106, No. 3. P. 429–435. <https://doi.org/10.1134/S1064562422700016>
 8. *Муромский А.А., Тучкова Н.П.* Представление математических понятий в онтологии научных знаний // Онтология проектирования. 2019. Т. 9, №1 (31). С. 50–69. <https://doi.org/10.18287/2223-9537-2019-9-1-50-69>
 9. *Nevzorova O.A., Falileeva M.V., Kirillovich A.V. et al.* OntoMathEdu Educational Ontology: Problems of Ontological Engineering // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications. 2023. Vol. 33, No. 3. P. 460–466. <https://doi.org/10.1134/S1054661823030367>
 10. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* Ontological Approach:
-

Knowledge Representation and Knowledge Extraction // *Lobachevskii J. Math.* 2020. Vol. 41 (10). P. 1938–1948. <https://doi.org/10.1134/S1995080220100030>

11. *Singhal A.* Introducing the Knowledge Graph: things, not strings // Google Official Blog, 2012.

URL: <https://blog.google/products/search/introducing-knowledge-graph-things-not/> (дата обращения 22.03.2026).

12. *Paulheim H.* Knowledge graph refinement: A survey of approaches and evaluation methods // *Semantic Web.* 2017. Vol. 8. P. 489–508.

<https://doi.org/10.3233/SW-160218>

13. *Hogan A., Gutierrez C., Cochez M. et al.* Data Graphs // In: *Knowledge Graphs. Synthesis Lectures on Data, Semantics, and Knowledge.* Springer, Cham, 2022. P. 5–23. https://doi.org/10.1007/978-3-031-01918-0_2

14. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* Ontological Approach to a Knowledge Graph Construction in a Semantic Library // *Lobachevskii J. Math.* 2023. Vol. 44, No. 6. P. 2229–2239. <https://doi.org/10.1134/S1995080223060471>

15. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* From Texts to Knowledge Graph in the Semantic Library LibMeta // *Lobachevskii J. Math.* 2024. Vol. 45, No. 5. P. 2211–2219. <https://doi.org/10.1134/S1995080224602625>

16. *Халов А.П., Атаева О.М.* Автоматические и полуавтоматические методы построения графа знаний предметной области и расширения онтологии // *Электронные библиотеки.* 2025. Т. 28. №. 6. С. 1481–1519.

<https://doi.org/10.26907/1562-5419-2025-28-6-1481-1519>

17. *Ataeva O.M., Tuchkova N.P.* Navigation with Large Language Models in Subject Domain of Ordinary Differential Equation // *Lobachevskii J. Math.* 2025. Vol. 46, No. 6. P. 2723–2735. <https://doi.org/10.1134/S1995080225608227>

18. *Липачев Е.К., Мурадымов Б.Р.* На пути к построению графа знаний математических уравнений // *Системы высокой доступности.* 2026. Т. 22, № 1. С. 41–46. <https://doi.org/10.18127/j20729472-202601-08>

19. *Зайцев В.Ф., Полянин А.Д.* Справочник по обыкновенным дифференциальным уравнениям. М.: Физматлит, 2001. 576 с.

20. *Зайцев В.Ф., Полянин А.Д.* Справочник по дифференциальным уравнениям с частными производными первого порядка. М.: Физматлит, 2003. 416 с.

21. Полянин А.Д. Справочник по линейным уравнениям математической физики. М.: Физматлит, 2001. 575 с.

22. Полянин А.Д., Зайцев В.Ф. Нелинейные уравнения математической физики. М.: Юрайт, 2017. 432 с.

23. Полянин А.Д., Манжиров А.В. Справочник по интегральным уравнениям. М.: Физматлит, 2003. 369 с.

24. Polyanin A.D., Zaitsev V.F. Handbook of Ordinary Differential Equations: Exact Solutions, Methods, and Problems. CRC Press/Chapman and Hall, 2017. 1496 p. <https://doi.org/10.1201/9781315117638>

25. Polyanin A.D., Zaitsev V.F. Handbook of Exact Solutions for Ordinary Differential Equations, 2nd Edition (Updated and Extended). CRC Press, Boca Raton–New York, 2003. 816 p.

26. Polyanin A.D., Zaitsev V.F., Moussiaux A. Handbook of First Order Partial Differential Equations. CRC Press, 2001. 520 p. <https://doi.org/10.1201/b16828>

27. Polyanin A.D., Zaitsev V.F. Handbook of Nonlinear Partial Differential Equations, Second edition. CRC Press, 2012. 1912 p. <https://doi.org/10.1201/b11412>

28. Polyanin A.D., Manzhirrov A.V. Handbook of Integral Equations, 2nd Edition. Chapman and Hall/CRC Press, Boca Raton–London, 2008. 1144 p. <https://doi.org/10.1201/9781420010558>

29. Виноградов И.М. (Ред.) Математическая энциклопедия (в 5 томах) М.: Советская энциклопедия (1977–1985).

30. Hazewinkel M. (Ed.) Encyclopaedia of Mathematics. An updated and annotated translation of the Soviet 'Mathematical Encyclopaedia'. Vol. 1–10. Springer Dordrecht, 1988. <https://doi.org/10.1007/978-94-009-6000-8>

31. Ji S., Pan S., Cambria E., Marttinen P., Yu P.S. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications // IEEE Transactions on Neural Networks and Learning Systems. 2022. Vol. 33, No. 2. P. 494–514. <https://doi.org/10.1109/TNNLS.2021.3070843>

32. Kirillovich A.V., Nevzorova O.A., Lipachev E.K. OntoMathPRO 2.0 Ontology: Up-dates of Formal Model // Lobachevskii J. of Math. 2022. Vol. 43, No. 12. P. 3504–3514. <https://doi.org/10.1134/S1995080222150136>

33. Елизаров А.М., Кириллович А.В., Липачев Е.К., Невзорова О.А. Новые

компоненты онтологии OntoMathPRO представления математического знания // Научный сервис в сети Интернет. 2023. № 25. С. 141–151.

<https://doi.org/10.20948/abrau-2023-32>

34. RDF Mapping Language (RML). Unofficial Draft, 20 June 2024.

URL: <https://rml.io/specs/rml/> (дата обращения 22.03.2026)

35. *Елизаров А.М., Кириллович А.В., Липачев Е.К., Невзорова О.А.* Цифровая экосистема OntoMath как подход к построению пространства математических знаний // Электронные библиотеки. 2023. Т. 26. № 2. С. 154–202.

<https://doi.org/10.26907/1562-5419-2023-26-2-154-202>

36. *Elizarov A.M., Lipachev E.K.* Lobachevskii Digital Library in the Scientific Space of Mathematical Knowledge // Scientific and Technical Information Processing. 2023. Vol. 50, No. 1. P. 35–39. <https://doi.org/10.3103/s0147688223010021>

37. *Elizarov A., Lipachev E.* Big math methods in Lobachevskii-DML // CEUR Workshop Proc. 2019. Vol. 2523. P. 59–72.

<https://ceur-ws.org/Vol-2523/invited08.pdf>, last accessed 2026/04/04

ENGINEERING AND AUTOMATIC CONSTRUCTION OF A KNOWLEDGE GRAPH “MATHEMATICAL EQUATIONS”

E. K. Lipachev¹ [0000-0001-7789-2332], **B. R. Muradymov**² [0009-0004-1187-8158]

^{1, 2} *Kazan (Volga region) Federal University, Kazan, Russia*

¹ *Innopolis University, Innopolis, Russia*

¹elipachev@gmail.com, ²muradymov.bulat@mail.ru

Abstract

We propose an approach to engineering and implementing a knowledge graph for representing and storing knowledge about mathematical equations. We have developed a knowledge graph prototype that represents knowledge about the main types of mathematical equations, including algebraic equations, ordinary differential equations, partial differential equations, and integral equations. We designed the knowledge graph of mathematical equations as a mathematical artifact. We are inte-

grating this artifact into the digital ecosystem of the Lobachevskii Digital Mathematical Library, therefore, we took into account the ecosystem's general compatibility requirements during the design. We have developed software tools for extracting and processing information about equations presented in digital libraries and electronic scientific resources. The current version of the knowledge graph prototype is based on the OntoMathPRO ontology of professional mathematics and a taxonomy of equations, built on information extracted from the web pages of the portal EqWorld "The World of Mathematical Equations." We expanded the OntoMathPRO ontology with new equation classes and new relationships to align with the equation type hierarchy presented on the EqWorld portal. We implemented a set of software modules that support the full cycle of knowledge graph generation, including a module for automatically extracting entities from external sources, a module for linking entities to OntoMathPRO ontology concepts, and a module for converting the acquired knowledge into an RDF representation and then storing it in a data warehouse. The knowledge graph supports SPARQL queries.

Keywords: *knowledge graph, knowledge extraction, mathematical equation, mathematical ontology, representation of mathematical knowledge.*

REFERENCES

1. All-Russian Portal Math-Net.Ru. URL: <https://www.mathnet.ru/> (Accessed: 22.03.2026).
2. EqWorld. The World of Mathematical Equations. URL: <https://eqworld.ipmnet.ru/> (Accessed: 22.03.2026).
3. Lobachevskii Digital Mathematical Library. URL: <https://lobachevskii-dml.ru/> (Accessed: 22.03.2026).
4. Numdam, the French digital mathematics library. URL: <https://www.numdam.org/> (Accessed: 22.03.2026).
5. Encyclopedia of Mathematics. URL: <https://encyclopediaofmath.org/> (Accessed: 22.03.2026).
6. *Elizarov A.M., Lipachev E.K.* Lobachevskii Digital Library in the Scientific Space of Mathematical Knowledge // Scientific and Technical Information Processing. 2023. Vol. 50, No. 1. P. 35–39. <https://doi.org/10.3103/s0147688223010021>
7. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* OntoMathPRO:

an Ontology of Mathematical Knowledge // *Doklady Mathematics*. 2022. Vol. 106, No. 3. P. 429–435. <https://doi.org/10.1134/S1064562422700016>

8. *Muromskiy A.A., Tuchkova N.P.* Representation of Mathematical Concepts in the Ontology of Scientific Knowledge // *Ontology of Designing*. 2019. Vol. 9, No. 1 (31). P. 50–69. <https://doi.org/10.18287/2223-9537-2019-9-1-50-69>

9. *Nevzorova O.A., Falileeva M.V., Kirillovich A.V. et al.* OntoMathEdu Educational Ontology: Problems of Ontological Engineering // *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications*. 2023. Vol. 33, No. 3. P. 460–466. <https://doi.org/10.1134/S1054661823030367>

10. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* Ontological Approach: Knowledge Representation and Knowledge Extraction // *Lobachevskii J. Math.* 2020. Vol. 41 (10). P. 1938–1948. <https://doi.org/10.1134/S1995080220100030>

11. *Singhal A.* Introducing the Knowledge Graph: things, not strings // *Google Official Blog*, 2012. URL: <https://blog.google/products/search/introducing-knowledge-graph-things-not/> (Accessed: 22.03.2026)

12. *Paulheim H.* Knowledge graph refinement: A survey of approaches and evaluation methods // *Semantic Web*. 2017. Vol. 8. P. 489–508. <https://doi.org/10.3233/SW-160218>

13. *Hogan A., Gutierrez C., Cochez M. et al.* Data Graphs // In: *Knowledge Graphs. Synthesis Lectures on Data, Semantics, and Knowledge*. Springer, Cham, 2022. P. 5–23. https://doi.org/10.1007/978-3-031-01918-0_2

14. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* Ontological Approach to a Knowledge Graph Construction in a Semantic Library // *Lobachevskii J. Math.* 2023. Vol. 44, No. 6. P. 2229–2239. <https://doi.org/10.1134/S1995080223060471>

15. *Ataeva O.M., Serebryakov V.A., Tuchkova N.P.* From Texts to Knowledge Graph in the Semantic Library LibMeta // *Lobachevskii J. Math.* 2024. Vol. 45, No. 5. P. 2211–2219. <https://doi.org/10.1134/S1995080224602625>

16. *Khalov A.P., Ataeva O.M.* Automatic and Semi-automatic Methods for Domain Knowledge-Graph Construction and Ontology Expansion // *Russian Digital Libraries Journal*. 2025. Vol. 28, No. 6. P. 1481–1519. <https://doi.org/10.26907/1562-5419-2025-28-6-1481-1519>

17. *Ataeva O.M., Tuchkova N.P.* Navigation with Large Language Models in Subject Domain of Ordinary Differential Equation // *Lobachevskii J. Math.* 2025. Vol. 46, No. 6. P. 2723–2735. <https://doi.org/10.1134/S1995080225608227>
 18. *Lipachev E.K., Muradymov B.R.* Towards Building the Knowledge Graph of Mathematical Equations // *Highly Available Systems.* 2026. Vol. 22, No. 1. P. 41–46. <https://doi.org/10.18127/j20729472-202601-08>
 19. *Zajcev V.F., Polyanin A.D.* *Spravochnik po obyknovennym differenci-al'nym uravneniyam.* M.: Fizmatlit, 2001. 576 s.
 20. *Zajcev V.F., Polyanin A.D.* *Spravochnik po differencial'nym uravneniyam s chastnymi proizvodnymi pervogo poryadka.* M.: Fizmatlit, 2003. 416 s.
 21. *Polyanin A.D.* *Spravochnik po linejnym uravneniyam matematicheskoy fiziki.* M.: Fizmatlit, 2001. 575 s.
 22. *Polyanin A.D., Zajcev V.F.* *Nelinejnye uravneniya matematicheskoy fiziki.* M.: Yurajt, 2017. 432 s.
 23. *Polyanin A.D., Manzhirov A.V.* *Spravochnik po integral'nym uravneniyam.* M.: Fizmatlit, 2003. 369 s.
 24. *Polyanin A.D., Zaitsev V.F.* *Handbook of Ordinary Differential Equations: Exact Solutions, Methods, and Problems.* CRC Press/Chapman and Hall, 2017. 1496 p. <https://doi.org/10.1201/9781315117638>
 25. *Polyanin A.D., Zaitsev V.F.* *Handbook of Exact Solutions for Ordinary Differential Equations, 2nd Edition (Updated and Extended).* CRC Press, Boca Raton–New York, 2003. 816 p.
 26. *Polyanin A.D., Zaitsev V.F., Moussiaux A.* *Handbook of First Order Partial Differential Equations.* CRC Press, 2001. 520 p. <https://doi.org/10.1201/b16828>
 27. *Polyanin A.D., Zaitsev V.F.* *Handbook of Nonlinear Partial Differential Equations, Second edition.* CRC Press, 2012. 1912 p. <https://doi.org/10.1201/b11412>
 28. *Polyanin A.D., Manzhirov A.V.* *Handbook of Integral Equations, 2nd Edition.* Chapman & Hall/CRC Press, Boca Raton–London, 2008. 1144 p. <https://doi.org/10.1201/9781420010558>
 29. *Vinogradov I.M. (Red.)* *Matematicheskaya ehnciklopediya (v 5 tomah) M.: Sovetskaya ehnciklopediya (1977–1985).*
-

30. *Hazewinkel M. (Ed.)* Encyclopaedia of Mathematics. An updated and annotated translation of the Soviet 'Mathematical Encyclopaedia'. Vol. 1–10. Springer Dordrecht, 1988. <https://doi.org/10.1007/978-94-009-6000-8>

31. *Ji S., Pan S., Cambria E., Marttinen P., Yu P.S.* A Survey on Knowledge Graphs: Representation, Acquisition, and Applications // IEEE Transactions on Neural Networks and Learning Systems. 2022. Vol. 33, No. 2. P. 494–514. <https://doi.org/10.1109/TNNLS.2021.3070843>

32. *Kirillovich A.V., Nevzorova O.A., Lipachev E.K.* OntoMathPRO 2.0 Ontology: Up-dates of Formal Model // Lobachevskii J. of Math. 2022. Vol. 43, No. 12. P. 3504–3514. <https://doi.org/10.1134/S1995080222150136>

33. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* New components of the OntoMathPRO ontology for representing math knowledge // Nauchnyj servis v seti Internet. 2023. № 25. S. 141–151. <https://doi.org/10.20948/abrau-2023-32>

34. RDF Mapping Language (RML). Unofficial Draft, 20 June 2024. URL: <https://rml.io/specs/rml/> (Accessed: 22.03.2026)

35. *Elizarov A.M., Kirillovich A.V., Lipachev E.K., Nevzorova O.A.* Digital Ecosystem OntoMath as an Approach to Building the Space of Mathematical Knowledge // Russian Digital Libraries Journal. 2023. Vol. 26, No. 2. P. 154–202. <https://doi.org/10.26907/1562-5419-2023-26-2-154-202>

36. *Elizarov A.M., Lipachev E.K.* Lobachevskii Digital Library in the Scientific Space of Mathematical Knowledge // Scientific and Technical Information Processing. 2023. Vol. 50, No. 1. P. 35–39. <https://doi.org/10.3103/s0147688223010021>

37. *Elizarov A., Lipachev E.* Big math methods in Lobachevskii-DML // CEUR Workshop Proc. 2019. Vol. 2523. P. 59–72. <https://ceur-ws.org/Vol-2523/invited08.pdf> (Accessed: 22.03.2026).

СВЕДЕНИЯ ОБ АВТОРАХ



ЛИПАЧЕВ Евгений Константинович – кандидат физико-математических наук, доцент кафедры цифровой аналитики и технологий искусственного интеллекта Института информационных технологий и интеллектуальных систем Казанского федерального университета, доцент Университета Иннополис. Научные интересы: цифровые библиотеки, интеллектуальный анализ данных, рекомендательные системы, технологии извлечения знаний.

Evgeny Konstantinovich LIPACHEV – Candidate of Physics and Mathematics, Associate Professor, Kazan Federal University, Innopolis University. Research interests: digital libraries, data mining, recommender systems, knowledge extraction technologies.

email: elipachev@gmail.com;

ORCID: 0000-0001-7789-2332



МУРАДЫМОВ Булат Русланович – магистрант Института информационных технологий и интеллектуальных систем Казанского федерального университета. Научные интересы: интеллектуальный анализ данных, рекомендательные системы, технологии извлечения знаний, графы знаний.

Bulat MURADYMOV – student at the Institute of Information Technologies and Intelligent Systems, Kazan Federal University. Research interests: data mining, recommender systems, knowledge extraction technologies, knowledge graphs.

email: muradyimov.bulat@mail.ru;

ORCID: 0009-0004-1187-8158

Материал поступил в редакцию 23 марта 2026 года