

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра дифференциальных уравнений и системного анализа

Аннотация к дипломной работе

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ МАТЕМАТИЧЕСКИХ ТЕКСТОВ

Каллаур Алексей Андреевич

Научный руководитель:
ст. преп. К. Г. Атрохов

2015

В дипломной работе 43 страницы, 3 рисунка, 2 таблицы, 3 источника, 2 приложения.

СИСТЕМНЫЙ АНАЛИЗ, ПРОЕКТИРОВАНИЕ, ПРОГРАММИРОВАНИЕ, ИНФОРМАЦИОННЫЕ СИСТЕМЫ, БАЗЫ ДАННЫХ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ТЕКСТОВ.

В дипломной работе описывается подход и разработка программного обеспечения для интеллектуального анализа математических текстов.

Целью дипломной работы является разработка модуля интеллектуального анализа текстов в web-приложении для автоматизированного наполнения базы данных математическими сущностями из различных учебников и статей.

Для достижения поставленной цели использовался подход и функционал Java фреймворка GATE Embedded.

В дипломной работе получены следующие результаты:

1. Разработан модуль для анализа математических текстов.
2. Созданы словари ключевых слов, а также правила, содержащие специальные инструкции для извлечения математических сущностей из документов.
3. Приведена статистика и проанализированы результаты обработки текстов для нескольких учебников.

Дипломная работа носит практический характер. Разработанный модуль может быть использован в качестве плагина для системы GATE Developer, а также инструмента для наполнения данными математической энциклопедии.

Дипломная работа выполнена автором самостоятельно.

There are 43 pages, 3 pictures, 2 tables, 3 resources, 2 attachments in the diploma project.

SYSTEM ANALYSIS, SOFTWARE ENGINEERING, INFORMATIONAL SYSTEMS, DATABASES, TEXT MINING.

The diploma project describes the approach of software developing for text mining.

This work aims developing of text mining module for web application. This module must serve for extracting different mathematical entities from articles and books, and filling the application database with the extracted terms.

Java framework GATE Embedded was used to reach the main goal.

Diploma project results:

1. Designed text mining module.
2. Created dictionaries and JAPE rules.
3. Provided statistics and analysis of document processing.

This work is practical. Text mining module developed during the work could serve as a plugin for GATE Developer and as a tool for filling the database of the application.

This work is performed independently by the author.