

**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**Факультет прикладной математики и информатики**

**Кафедра компьютерных технологий и систем**

Аннотация к дипломной работе

**«Компонент системы интеллектуального анализа текста»**

Афонова Ксения Владимировна

Научный руководитель – ассистент кафедры КТС Буяльская Ю.В., директор ООО «Релсофт» Лапатко А.В., доктор физ.-мат. наук, профессор Таранчук В.Б.

## Реферат

Дипломная работа, 36 с., 14 рис., 1 табл., 10 источников.

### ПОЛНОТЕКСТОВЫЙ ПОИСК, ИССЛЕДОВАНИЕ, ИНДЕКС, JAVA, АНАЛИЗ, ТЕКСТ, КЛЮЧЕВЫЕ СЛОВА

*Объект исследования* – печатный текст и его характеристики.

*Цель исследования* – анализируя существующие методы и системы, использующиеся для полнотекстового поиска в документах, реализовать программный модуль для реализации поиска ключевых слов.

За время работы были реализованы следующие задачи: изучены основные принципы работы систем полнотекстового поиска и построения поисковых индексов, исследованы основные существующие программные средства, предназначенные для полнотекстового поиска, проведена их сравнительная характеристика, найден и программно реализован наиболее подходящий способ интерпретации содержательной стороны неструктурированной документальной информации.

Работа имеет практическое значение (новостные порталы, каталоги сайтов, маркетинг), так как позволяет быстро получить ключевые слова по статьям, не прибегая к помощи человека. Результаты дипломной работы могут найти применение в сети Интернет и среди студентов.

Реализация проекта дает возможность с определенной долей вероятности определить ключевые слова в тексте, тем самым облегчая поиск документа.

## **Abstract**

Diploma thesis 36 p., 14 figures, 1 table, 10 sources.

FULL-TEXT SEARCH, RESEARCH, INDEX, JAVA, ANALYZE, TEXT, KEYWORDS

*Object of research* – printed text and its' characteristics.

*Purpose* – with the help of analyzing existing full-text search methods and systems implement a software module for keywords search.

Completed following tasks during the work: studied base principles of full-text search engines work and search index constructing, researched common software for full-text search and performed comparison, found and implemented the most suitable way of interpretation of unstructured data.

The work is of practical use (news blogs, site catalogues, marketing) because of ability to retrieve keywords from articles in near real-time without human help. The results of the work can be used in Internet and among students.

Project implementation helps to retrieve keywords from an article with a certain probability and makes easier articles' search.