

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра информационных систем управления

Аннотация к дипломной работе

**Автоматический поиск релевантных текстовых документов в
многоязычной интернет-среде**

Селяник Вадим Александрович

Научный руководитель – доктор технических наук, профессор
Совпель И. В.

Минск 2018

РЕФЕРАТ

Дипломная работа, 66 с., 4 табл., 14 рис., 27 источников.

Ключевые слова: РЕЛЕВАНТНОСТЬ, КЛЮЧЕВОЕ СЛОВО, ПОИСК, ДОКУМЕНТ, ПОИСКОВЫЙ ЗАПРОС, МУЛЬТИЯЗЫЧНОСТЬ.

Объект исследования – система автоматического поиска релевантных текстовых документов в многоязычной интернет-среде.

Цель работы – изучить существующие решения автоматического поиска релевантных текстовых документов в многоязычной интернет-среде; разработать алгоритм и на его основе реализовать программный комплекс для решения поставленной задачи.

За время работы были использованы следующие методы исследования: изучение предметной области, анализ существующих решений, анализ и разработка алгоритмов, системный подход, технологии разработки системных приложений.

В результате исследования разработан программный комплекс, позволяющей автоматически находить текстовые документы релевантные данному в многоязычной интернет-среде.

Область применения – классификация и анализ текстов, информационный поиск.

РЭФЕРАТ

Дыпломная праца, 66 с., 4 табл., 14 мал., 27 крыніц.

Ключавыя словы: РЕЛЕВАНТНАСЦЬ, КЛЮЧАВОЕ СЛОВА, ПОШУК, ДАКУМЕНТ, ПОШУКАВЫ ЗАПЫТ, МУЛЬТІМОУЕ.

Аб'ект даследавання – сістэма аўтаматычнага пошуку рэлевантных тэкставых дакументаў у шматмоўным інтэрнэт-асяроддзі.

Мэта працы – даследаваць існуючыя рашэння аўтаматычнага пошуку рэлевантных тэкставых дакументаў у шматмоўным інтэрнэт-асяроддзі; распрацаваць алгарытм і на яго аснове рэалізаваць праграмны комплекс для вырашэння пастаўленай задачы.

Падчас працы былі выкарыстаны наступныя метады: вывучэнне прадметнай вобласці, аналіз рашэнняў, аналіз і распрацоўка алгарытмаў, сістэмны падыход, тэхналогіі распрацоўкі сістэмных прыкладанняў.

У выніку даследавання распрацаваны праграмны комплекс, які дазваляе аўтаматычна знаходзіць тэкставыя дакументы рэлевантных дадзенаму ў шматмоўным інтэрнэт-асяроддзі.

Вобласць прымянення - класіфікацыя і аналіз тэкстаў, інфармацыйны пошук.

ABSTRACT

Graduate work, 66 pp., 4 tables, 14 figures, 27 sources.

Keywords: RELEVANCE, KEYWORD, SEARCH, DOCUMENT, SEARCH REQUEST, MULTILINGUALITY.

Object of research – automatic search system for relevant text documents in a multilingual Internet environment.

The work purpose – to study the existing solutions of automatic search of relevant text documents in a multilingual Internet environment; to develop an algorithm and implement a software system on its basis for solving the task.

During the course work, the following methods were used: study of the subject area, analysis of solutions, analysis and development of algorithms, system approach, technologies for developing system applications.

As a result of the research, a software package was developed that allows to automatically find text documents relevant to the given document in a multilingual Internet environment.

Scope – classification and text mining, information retrieval.