

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра информационных систем управления

Аннотация к дипломной работе

**АВТОМАТИЧЕСКОЕ ТЕМАТИЧЕСКОЕ ТЕГИРОВАНИЕ
ТЕКСТОВЫХ ДОКУМЕНТОВ**

Якубенко Анна Викторовна

Научный руководитель – к.т.н. Совпель И. В.

Минск

2017

Аннотация

Рассмотрена задача автоматического тематического тегирования англоязычных текстовых документов. Изучена предметная область задачи. Исследованы основные методы решения задач обработки текстов на естественном языке. Предложен алгоритм решения задачи автоматического тематического тегирования текстовых документов. Проведен анализ лингвистических ресурсов. Алгоритм реализован на практике. Разработано приложение для демонстрации алгоритма. Проведен экспертный анализ результатов.

Анотацыя

Разгледжана задача аўтаматычнага тэматычнага тэгіравання тэкставых дакументаў. Даследавана прадметная вобласць задачы. Разгледжаны асноўныя метады рашэння задач апрацоўвання тэкстаў на натуральнай мове. Запрапанаваны алгарытм рашэння задачы аўтаматычнага тэматычнага тэгіравання тэкставых дакументаў. Праведзены аналіз лінгвістычных рэсурсаў. Алгарытм рэалізаваны на практыцы. Распрацавана прыкладанне для дэманстрацыі алгарытму. Праведзены экспертны аналіз вынікаў.

Annotation

In this work considers automatic topic modeling of text documents problem. The subject area was studied. Methods of solving natural language problems were described. Algorithm of solving this problem was offered. Available linguistic resources were analyzed. The algorithm is implemented in practice. An application was developed to demonstrate the algorithm. An expert analysis of the results was carried out.