

MONOGRAFIA
POKONFERENCYJNA

SCIENCE,
RESEARCH, DEVELOPMENT #13

TECHNICS AND TECHNOLOGY.

Berlin (Берлин)

30.01.2019 - 31.01.2019

U.D.C. 330+339.138+658+657+336.71+339+082

B.B.C. 94

Z 40

Zbiór artykułów naukowych recenzowanych.

(1) Z 40 Zbiór artykułów naukowych z Konferencji Międzynarodowej Naukowo-Praktycznej (on-line) zorganizowanej dla pracowników naukowych uczelni, jednostek naukowo-badawczych oraz badawczych z państw obszaru byłego Związku Radzieckiego oraz byłej Jugosławii.

(30.01.2019) - Warszawa, 2019. - 208 str.

ISBN: 978-83-66030-79-4

Wydawca: Sp. z o.o. «Diamond trading tour»

Adres wydawcy i redakcji: 00-728 Warszawa, ul. S. Kierbedzia, 4 lok.103

e-mail: info@conferenc.pl

Wszelkie prawa autorskie zastrzeżone. Powielanie i kopiowanie materiałów bez zgody autora jest zakazane. Wszelkie prawa do artykułów z konferencji należą do ich autorów.

W artykułach naukowych zachowano oryginalną pisownię.

Wszystkie artykuły naukowe są recenzowane przez dwóch członków Komitetu Naukowego.

Wszelkie prawa, w tym do rozpowszechniania i powielania materiałów opublikowanych w formie elektronicznej w monografii należą Sp. z o.o. «Diamond trading tour».

W przypadku cytowań obowiązkowe jest odniesienie się do monografii.

Nakład: 80 egz.

«Diamond trading tour» ©

Warszawa 2019

ISBN: 978-83-66030-79-4

АВТОМАТИЧЕСКАЯ СИСТЕМА АНАЛИЗА

Абжалова Манзура Абдурашатовна

Навоийский государственный горный институт Соискатель Ташкентского государственного университета узбекского языка и литературы имени Алишера Навои

***Аннотация.** В данной статье говорится о важности разработки модулей автоматического анализа текста.*

***Annotation.** In this article talks about the importance of developing automatic text analysis modules.*

Ключевые слова: лингвистический модуль, автоматическая система анализа, морфологический анализ, основная форма слова, морфологическая характеристика.

Key words: linguistic module, automatic analysis system, morphological analysis, basic form of the word, morphological characteristic.

Автоматическая система анализа (АСА) – это процедура выполнения программы для анализа текста.

При создании ССТ нужно сделать следующие:

1. Создание лингвистического модуля, включающего все лингвистические принципы.
2. Создание программного алгоритма с использованием модуля.
3. Создание программы с использованием алгоритма.

Когда все это сделано успешно, человек может очень быстро проанализировать текст на компьютере с помощью голоса или письменных инструкций.

Как известно, тексты разные: художественный текст, публичный текст, научный текст и так далее. Эти тексты разные по стилю. При создании программы автоматического анализа текстов на узбекском языке большое значение имеют созданные языковедами

лингвистические данные. При создании лингвистической базы обязательно учитываются функциональные методы национального языка. или будет создана программа автоматического анализа для определенного стиля.

В русском энциклопедическом словаре дается полное описание автоматического анализа текста: Автоматический анализ текста (АА), операция, которая заключается в том, что из данного текста на естественном языке извлекается содержащаяся в этом тексте грамматическая и семантическая информация, выполняемая по некоторому алгоритму в соответствии с заранее разработанным описанием данного языка¹. Этап морфологический анализ имеет важное значение в автоматическом анализе. Автоматический морфологический анализ – обеспечивает определение нормальной формы, от

¹ <https://slovar.cc/enc/bse/1970174.html>
Автоматический Анализ

которой была образована данная словоформа и набора параметров, приписанных данной словоформе. Это делается для того, чтобы ориентироваться в дальнейшем только на нормальную форму, а не на все словоформы, использовать параметры, например, для проверки согласования слов².

Цель морфологического анализа, это определить:

- 1) основную форму слова;
- 2) полную морфологическую характеристику слова (морфологические параметры слова).

Автоматический морфологический анализ помогает избежать лингвисту, обрабатывать все словоформу одного слова.

Узбекский язык является агглютинативным. В нем часть информации об употреблении слова добавляется в конец слова в виде аффиксов, то есть словоформа образуется от основы, которому аффиксы связываются по порядку: основа+словообразующий аффикс+формаобразующий аффикс+словоизменяющий аффикс. В основном, основа остается так она есть. Например, *иш+чи+лар+имиз* «наше работники».

Морфологические параметры – это пара: имя параметра, значения параметра³. Именем параметра может служить число, падеж, время, склонение,

наклонения и другие признаки слов, принятые в данном языке. Значение параметра – это конкретное значение, которое может принимать данный элемент языка. Например, время может быть прошедшим, настоящим и будущим; степени сравнения прилагательных может быть сравнительная степень, превосходная степень; число – единственным и множественным и т.д.

Программа автоматизированного анализа может использоваться в официальных офисах, научно-технических учреждениях, в основном для анализа официальных и научных текстов.

Роль лингвистических модулей очень важна при реализации программах как автоматический перевод и автоматический анализ. Создание лингвистического модуля является проблемной. Потому что модуль требует от лингвиста глубокого понимания семантики и структуры языка, а также для поиска конкретного решения многих теоретических и практических вопросов языковых явлений.

Система автоматического анализа не только работает с текстом, но и выполняет другие важные задачи. В результате создания интеллектуального компьютерного программного обеспечения сегодня наиболее совершенные программы выполняют следующие задачи:

- орфографическая, грамматическая и стилистическая экспертиза текстов;
- автоматическое реферирование – создания краткую аннотацию существенной информации с помощью ключевых слов в тексте;

2 Большакова Е.И. и др. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика (учеб. пособие) — М., 2011. — С.106

3 Большакова Е.И. и др. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика (учеб. пособие) — М., 2011. — стр.110.

- определить абзацы, символы, точные слова в тексте (показать статистику);

- транслитерация – перевод одной графической системы алфавита в другую (то есть передача букв одной письменности буквами другой). Например, транслитерации кирилловского письма латинским алфавитом.

- учитывать семантику слов;
- морфологический анализ ключевых слов;

- ответить на вопрос;
- определить значимые слова и фразы в документе;

- проведение поиска и исследования пользовательских ключевых слов и их синонимов;

- формулировка счета в соответствии с частотой слов в документе;

- расчет результата поиска, актуальности или даты составления списка документов;

- служба тезауруса – определение вариантов синонимии, антонимии, омонимии или другие лексические значения конкретного слова в тексте;

Самым простым решением проблемы создания морфологического словаря является таблица, в которой в первой колонке будет записана словоформа, во второй – нормальная форма, а в третьей – набор параметров. Для морфологического анализа в таком слова-

ре необходимо просто найти все соответствующие словоформы и выдать найденные результаты.

Муртазо Махмудян из Лозаннского университета (Университет Лувра) и парижской Высшей школы практики подчеркнул, что «язык должен охватывать все аспекты человеческого опыта» В этом отношении можно сказать, проблема разработать база данных для программы компьютерного анализа текста на узбекском языке ожидает своего решение.

Список литературы:

1. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Большакова Е.И. я доктор - М.: МИЭМ, 2011.
2. Болховитянов А.В., Гусев А.В., Чеповский А.М. Морфологические модели вычислительной лингвистики: учеб. пособие - М. МГУП, 2010.
3. Иомдин Л. Л. Автоматическая обработка текста на естественном языке: Модель согласования. – М.: «Наука», 1990
4. Мильчин А. Е. Методология редактирования текста. М.: Логос, 2005.
5. Новое в зарубежной лингвистике. Вып. XXIV. - компьютерная лингвистика. - М.: Прогресс, 1989.
6. Пулатов А.К., Джураева Н. Разработка формальной модели грамматики Узбекскоязычный // Узбекский математический журнал. –2002, № 1
7. <http://en.wikipedia.org>. Компьютерная лингвистика.
8. <http://www.redaktoram.ru>