

**Ю. Ю. Белых**

*(ГрГУ им. Я. Купалы, Гродно)*

## **РАЗРАБОТКА ПРОГРАММНОГО ИНТЕРФЕЙСА ДЛЯ ОБРАБОТКИ ЕСТЕСТВЕННЫХ ЯЗЫКОВ**

Обработка естественного языка (Natural Language Processing, далее NLP) это набор средств для перевода человеческого, естественного языка в структуры данных позволяющие дальнейший компьютерный анализ. (либо наоборот в некоторых случаях). Основными задачами, которые решает NLP являются: анализ текста, синтез текста, вопросно-ответные системы, извлечение информации, упрощение текста, распознавание речи, информационный поиск, машинный перевод. Несмотря на безграничные возможности приложения, построение таких систем с нуля дело чрезвычайно трудное, так как требует большого количества подготовительных работ (реализация инструментария), что, как правило, намного более длительный и сложный процесс чем реализация бизнес логики самого приложения. Существующие же наборы инструментов либо не предоставляют необходимой функциональности, либо имеют проблемы с доступностью (нет хорошего удаленного API), либо не отвечают промышленным стандартам надёжности и безопасности данных. Таким образом построение такого программного интерфейса является крайне актуальной на сегодняшний день задачей.

Цель работы – проектирование и реализация полноценного набора NLP инструментов высокого уровня, позволяющего на своей основе строить конкретные бизнес-приложения. Для достижения цели решаются следующие задачи:

- 1) анализ и классификация существующих алгоритмов обработки естественных языков;
- 2) построение веб и программного API покрывающего максимально полный спектр задач NLP;
- 3) обеспечение высокой производительности и безопасности системы промышленного уровня;
- 4) реализация нескольких референс-приложений для демонстрации возможностей продукта;
- 5) полное покрытие документацией для конечного пользователя.

В работе использованы методы построения систем искусственного интеллекта и математической лингвистики. Непосредственно реализация выполнена используя промышленные инструменты разработки программного обеспечения.

На данный момент не существует полноценного программного интерфейса для обработки естественных языков промышленного уровня. Результаты данной работы будут внедрены в научно-производственный процесс при реализации конкретных бизнес-приложений.