

Лабораторная работа № 5. Разработка автоматизированной системы семантико-синтаксического анализа текста естественного языка

Цель работы: освоить принципы разработки прикладных сервисных программ для решения задачи автоматического семантико-синтаксического анализа текста естественного языка.

Задачи лабораторной работы:

- познакомиться с назначением, структурой и функциональностью, предоставляемой базовым ЛП для решения задачи автоматического семантико-синтаксического анализа ТЕЯ;
- закрепить навыки программирования при решении задач автоматической обработки ТЕЯ.

Методические указания. Требуется спроектировать и программно реализовать структуры хранения данных, алгоритмы их обработки, необходимые в рамках следующих базовых требований к разрабатываемому приложению:

- входные данные – текст заданного естественного языка;
- выходные данные – структуры, полученные при проведении автоматического семантико-синтаксического анализа предложений входного текста согласно варианту задания;
- взаимодействие с пользователем посредством графического интерфейса (интерфейс должен быть интуитивно понятным и дружелюбным пользователю);
- наличие системы средств помощи пользователю;
- обеспечение возможности построения, сохранения, просмотра, редактирования, документирования автоматически получаемого результата либо заданной его части;
- поддержка различных форматов представления входных данных (TXT, RTF, PDF, HTML, DOC, DOCX).

Рекомендуется использовать функциональность стандартной, а также специализированных библиотек языка программирования *Python* для обработки естественного языка, например *nltk*, *WordNet*, *EuroWordnet*, *ConceptNet*, *FrameNet*.

Вариант задания выбирается студентом самостоятельно (табл. 3) и согласовывается с преподавателем. Средства разработки выбираются студентом самостоятельно. Защита лабораторной работы предполагает демонстрацию работоспособности всех реализованных функций в соответствии с требованиями.

Требования к отчету. В отчете представить, в том числе графически, используя такие программные средства, как *Microsoft Visio* или *Draw.io*:

- структурно-функциональную схему разработанного приложения;
- описание структур хранения данных, алгоритмов их обработки, необходимых для реализации базовых требований к разработанной программе;
- оценку быстродействия приложения;

- выводы по работе и по перспективам использования приложения.
Отчет представить для проверки в электронном виде.
Варианты заданий для выполнения приведены в табл. 3.

Таблица 3

Варианты индивидуальных заданий

Номер индивидуального задания	Язык текста	Формат входного документа
1	Русский	TXT
2	Русский	RTF
3	Русский	PDF
4	Русский	DOC
5	Русский	DOCX
6	Английский	HTML
7	Английский	TXT
8	Английский	RTF
9	Английский	PDF
10	Английский	DOC
11	Английский	DOCX
12	Английский	HTML