

Учреждение образования
“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ ”

Кафедра интеллектуальных информационных технологий

Отчет по лабораторной работе №4
по курсу: “Естественно-языковой интерфейс интеллектуальных систем”

Выполнили студенты группы 121701:

Пашин Н.А.
Маевский В.Г.

Проверил:

Крапивин Ю.Б.

Минск

2022

Цель работы: освоить на практике основные принципы машинного перевода документов.

Вариант 9:

Направление перевода: Англо-немецкий.

Предметная область: Научные статьи по медицине, Критика предметов изобразительного искусства.

Требования к разрабатываемой системе:

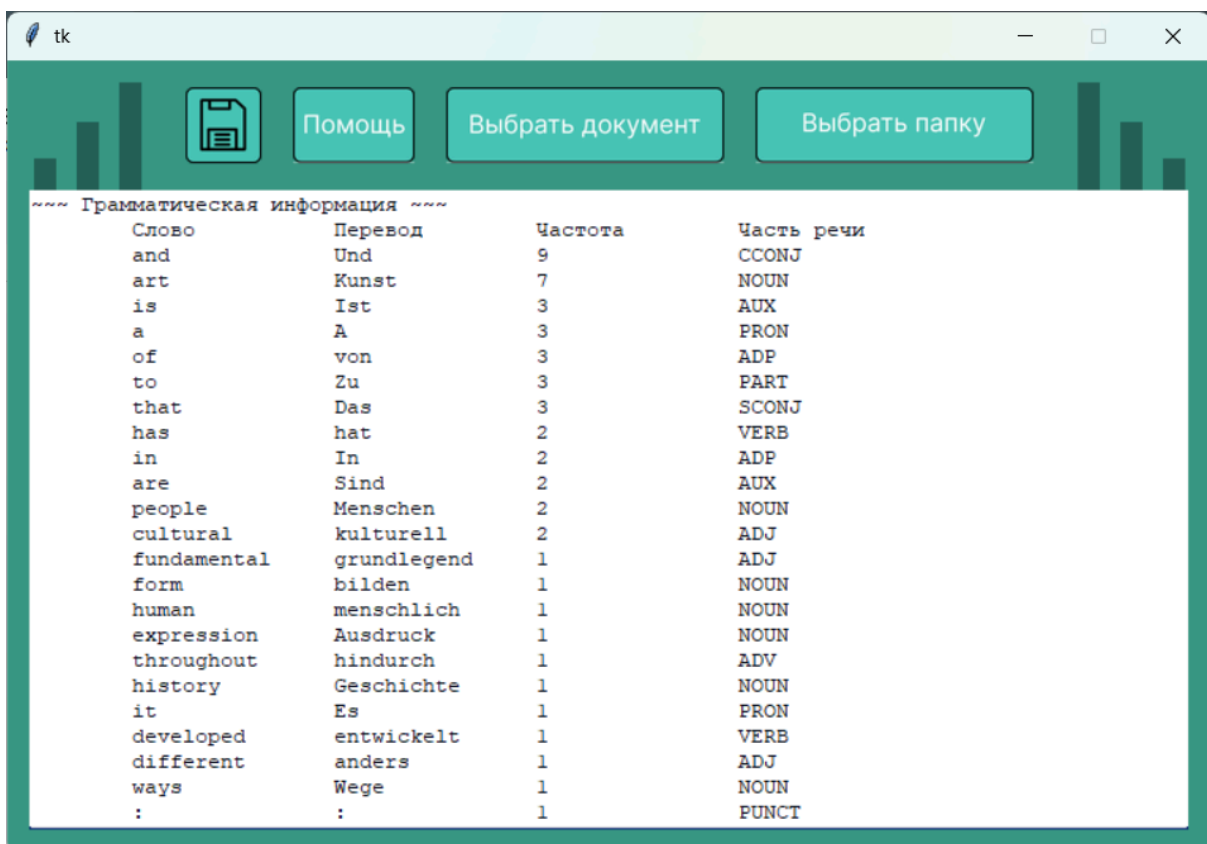
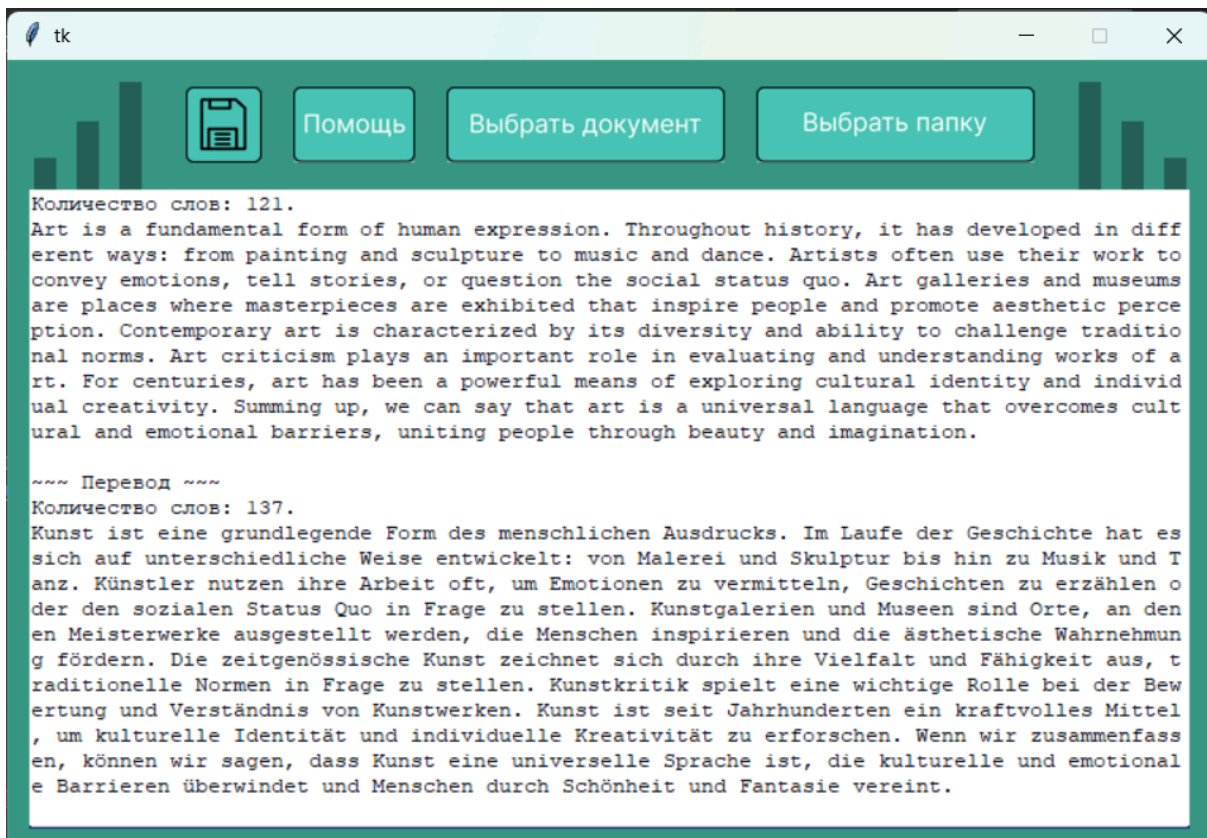
- на входе — естественно-языковой текст на входном языке, подлежащий процедуре машинного перевода;
- подсчитать количество слов во входном тексте, количество переведенных слов, определить грамматическую информацию (теги частей речи и их расшифровка, для этого следует использовать функциональность лр. №3 весенний семестр).
- на выходе: — перевод входного текста на выходной язык;
- — упорядоченный по частоте встречаемости в тексте список слов и их переводов на выходной язык с грамматической информацией (вкладка 1) (можно использовать функциональность лр. №1 весеннего семестра);
- — построенное дерево синтаксического разбора выбранного предложения (вкладка 2).
- обеспечить наличие утилиты автоматического пополнения/корректировки полученного словаря (таблицы БД).
- обеспечить сохранение и распечатку результатов перевода и упорядоченных по частоте встречаемости в тексте списков слов и их переводов на выходной язык с грамматической информацией в файл формата txt кодировки Unicode.
- интерфейс системы должен быть простым и доступным для пользователей любого уровня, содержать понятный набор инструментов.

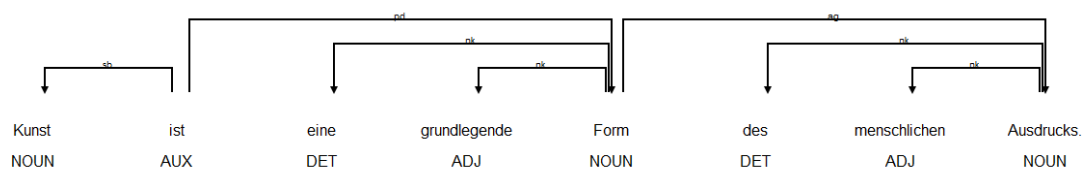
Структура разработанной системы

Система реализована на языке Python версия 3.10 . Графический интерфейс системы создан с помощью библиотеки tkinter. Система может принимать тексты путем ввода абсолютного пути файла. Выполняет англо-немецкий перевод текста, считает слова в обоих текстах, создает список слов с их переводом, частотой и грамматической информацией, выполняет синтаксический разбор предложений. Все результаты сохраняются в текстовый файл.

Перевод осуществляется путем отправки запроса с исходным текстом на <http://translate.google>. Грамматическая информация определяется из предобученной модели английского языка "en_core_web_sm". И при помощи модели "de_core_news_sm" выполняет синтаксический разбор переведенного текста и при помощи displacy.render визуализирует граф зависимостей.

Тестирование:





Вывод:

Разработанная система позволяет переводить текст и собирать некоторую информацию о нем: частоту встречаемости слов и синтаксические деревья. Система позволяет осуществлять англо-немецкий перевод.