Министерство образования Республики Беларусь

Учреждение образования

«Брестский государственный технический университет»

Кафедра ИИТ

Лабораторная работа №4

По дисциплине: «Естественно-языковой интерфейс ИС»

# Тема: «Разработка автоматизированной системы синтаксического анализа текста естественного языка»

Выполнил:

Студент 3 курса

Группы ИИ-21

Кирилович А. А.

Проверила:

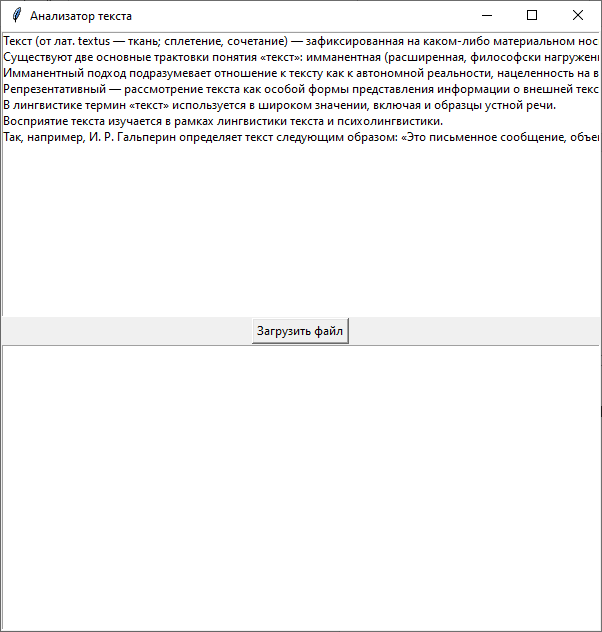
Якимук А. В.

Брест 2024

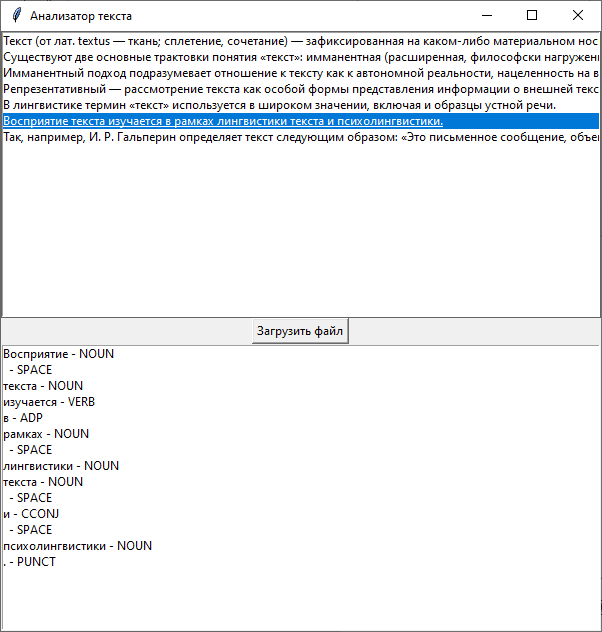
**Ход работы**

**Задание:**

1. *Входные данные* – текст заданного естественного языка;
2. *Выходные данные* – структуры, полученные при проведении автоматического синтаксического анализа предложений входного текста
3. Взаимодействие с пользователем посредствам графического интерфейса (интерфейс должен быть интуитивно-понятным и дружественным пользователю)



*Приложение с загруженным русским текстом*



*Синтаксический анализ выбранного предложения*

**Код программы:**

import tkinter as tk

from tkinter import filedialog

from docx import Document

import spacy

class TextAnalyzerApp:

def \_\_init\_\_(self, master):

self.master = master

master.title("Анализатор текста")

master.geometry("600x600")

self.sentences\_listbox = tk.Listbox(master, width=50)

self.sentences\_listbox.pack(side=tk.TOP, fill=tk.BOTH, expand=True)

self.sentences\_listbox.bind("<<ListboxSelect>>", self.show\_syntax\_analysis)

self.syntax\_listbox = tk.Listbox(master, width=50)

self.syntax\_listbox.pack(side=tk.BOTTOM, fill=tk.BOTH, expand=True)

self.syntax\_listbox.bind("<Button-1>", lambda event: "break")

self.load\_button = tk.Button(master, text="Загрузить файл", command=self.load\_file)

self.load\_button.pack(side=tk.BOTTOM)

def load\_file(self):

file\_path = filedialog.askopenfilename(filetypes=[("Word files", "\*.docx")])

if file\_path:

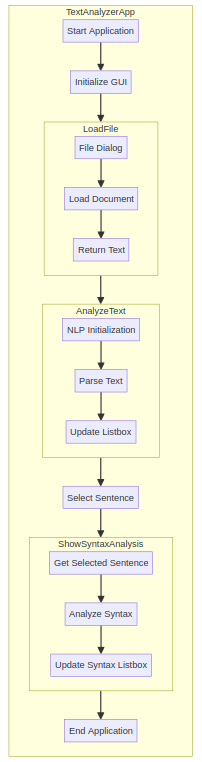
doc = Document(file\_path)

text = ""

for paragraph in doc.paragraphs:

text += paragraph.text + "\n"

self.analyze\_text(text)

def analyze\_text(self, text):

self.sentences\_listbox.delete(0, tk.END)

self.syntax\_listbox.delete(0, tk.END)

nlp = spacy.load('ru\_core\_news\_sm')

doc = nlp(text)

for sentence in doc.sents:

self.sentences\_listbox.insert(tk.END, sentence.text)

def show\_syntax\_analysis(self, event):

selected\_sentence\_index = self.sentences\_listbox.curselection()[0]

selected\_sentence = self.sentences\_listbox.get(selected\_sentence\_index)

nlp = spacy.load('ru\_core\_news\_sm')

doc = nlp(selected\_sentence)

self.syntax\_listbox.delete(0, tk.END)

for token in doc:

self.syntax\_listbox.insert(tk.END, f"{token.text} - {token.pos\_}")

def main():

root = tk.Tk()

app = TextAnalyzerApp(root)

root.mainloop()

if \_\_name\_\_ == "\_\_main\_\_":

main()

**Вывод:** в ходе выполнения лабораторной работы освоил принципы разработки прикладных сервисных программ для решения задачи анализа текста естественного языка.