



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства):
2024612896

Дата регистрации: 06.02.2024

Номер и дата поступления заявки:
2024611527 30.01.2024

Дата публикации и номер бюллетеня:
06.02.2024 Бюл. № 2

Контактные реквизиты:
нет

Автор(ы):

Кашин Илья Владимирович (RU),
Шатилов Макар Романович (RU),
Данилов Евгений Максимович (RU),
Заложный Виталий Евгеньевич (RU)

Правообладатель(и):

Федеральное государственное автономное
образовательное учреждение высшего
образования «Уральский федеральный
университет имени первого Президента России
Б.Н.Ельцина» (RU)

Название программы для ЭВМ:

Ансамблевый метод исправления фактологических ошибок в тексте, генерируемом большими лингвистическими моделями

Реферат:

Программа представляет собой высокопроизводительный модуль для осуществления автономной коррекции отклика большой лингвистической модели «LLaMa 2» на основе внешнего открытого бесплатного источника – Википедии. В рамках модуля осуществляется первичная подготовка текста путём решения задач «coreference resolution» и «semantic analysis», направленных на определения тематики текста и разбиения сложных предложений на более простые. Затем для каждого сегмента текста производится автоматический поиск наиболее релевантного материала в Википедии. Сверка данных происходит двумя независимыми методами: путём решения задачи «textual entailment» моделью «roberta-large-mnli», а также средствами самой «LLaMa 2». Текст с выявленными корректировками генерируется «LLaMa 2» и возвращается пользователю. Тип ЭВМ: IBM PC-совмест. ПК. ОС: Windows XP и выше.

Язык программирования: Python 3.10

Объем программы для ЭВМ: 200 КБ