





TextEaseBot

TextEaseBot

Умный бот для упрощения и перевода текстов

Создано с использованием:

 Transformer-моделей  Telegram API  Google Colab  NLP

*«Не все могут понять сложный текст.
Я помогаю сделать знания доступными.»*

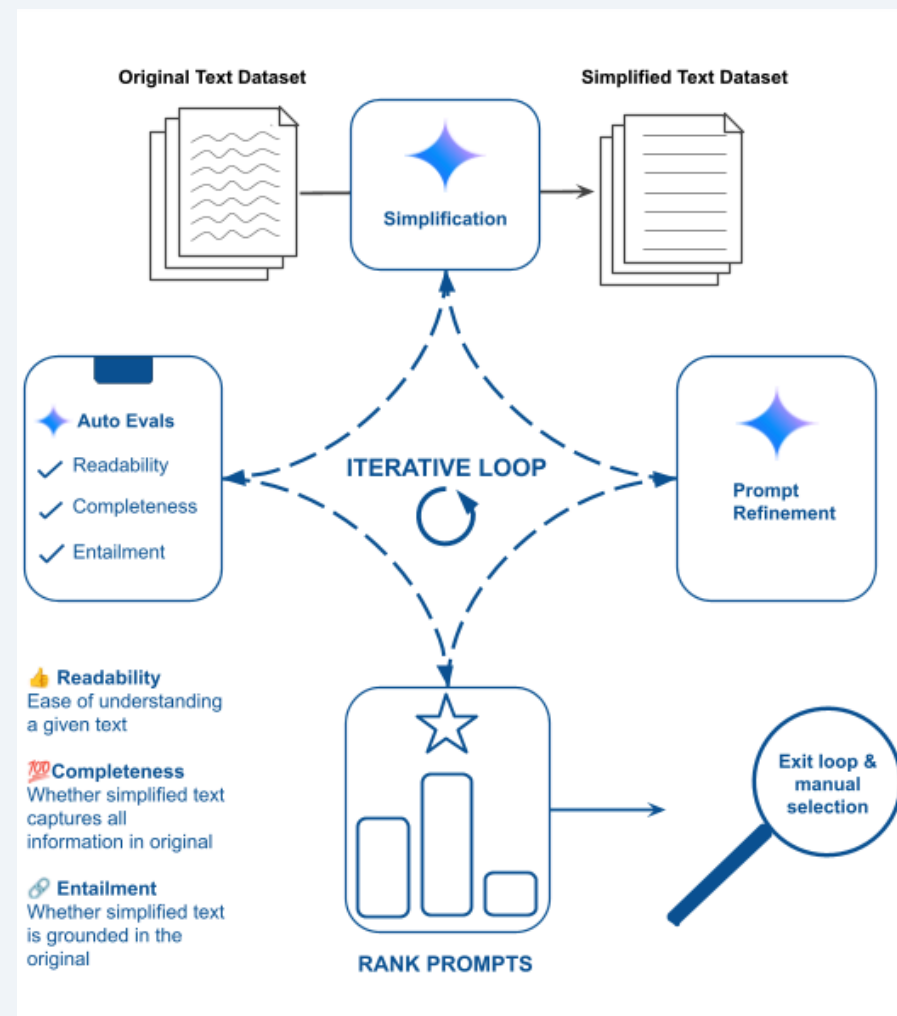
Проблема, которую решает бот

Сложные тексты отпугивают

- 📄 Научные статьи, юридические документы, медицинская информация — часто непонятны даже образованным людям
- 🔤 Языковой барьер мешает доступу к знаниям
- 🔧 Нет удобного инструмента, который **одновременно упрощает и переводит**

Цель проекта:

Создать бота, который делает **любой текст понятным и доступным** — без потери смысла.



Функционал и возможности TextEaseBot

TextEaseBot — это ИИ-помощник в Telegram, который:

T↑ Упрощает тексты на 2 уровнях:

 Средний — баланс простоты и смысла

 Сильный — максимально просто

 Переводит на английский

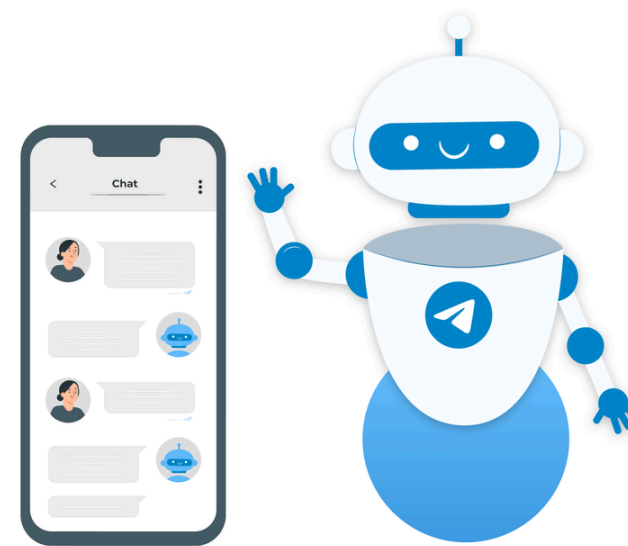
 Работает с файлами: **.txt**, **.docx**

 Поддерживает длинные тексты

 Показывает оценку качества упрощения (BLEU, длина)

 Включает режим "**Фактчекинг**" — разбивка на утверждения

Пользователь присылает текст — бот делает его ясным.



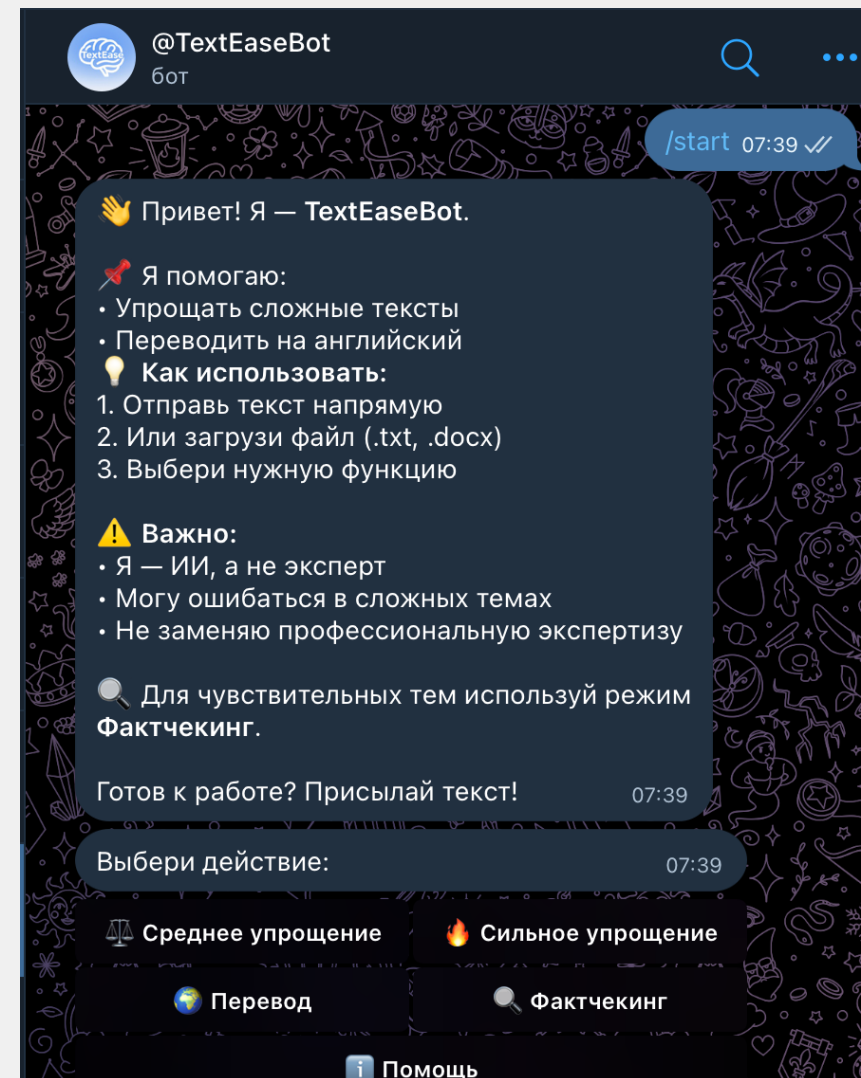
Технологии и архитектура решения

Стек технологий

-  **RuT5** — русскоязычная T5-модель для упрощения
-  **BERT эмбединги** — для гибкой обработки текстов
-  **Helsinki-NLP** — перевод с русского на английский
-  **Telegram Bot API** — интерфейс и взаимодействие
-  **NLTK** — разбивка на предложения
-  **sacrebleu** — оценка качества
-  **Google Colab / VS Code** — запуск и демонстрация

Основной пайплайн

- 1 Пользователь присылает текст
- 2 Бот разбивает на части (если текст длинный)
- 3 Упрощает с помощью RuT5
- 4 Оценивает качество с помощью BLEU
- 5 Предлагает перевод и результаты анализа



Особенности и инновации



Режим "Фактчекинг"

Разбивает текст на отдельные утверждения — помогает критически оценивать информацию, особенно в спорных или политических темах.



Оценка качества упрощения

- Длина до/после упрощения
- BLEU — насколько сохранён оригинальный смысл
- Визуальная индикация: 🟢 Отлично 🟡 Средне

🔴 Требуется проверки



Перевод на английский

Автоматический перевод упрощённых текстов на английский язык с сохранением контекста и смысла сообщения.



Пост-обработка

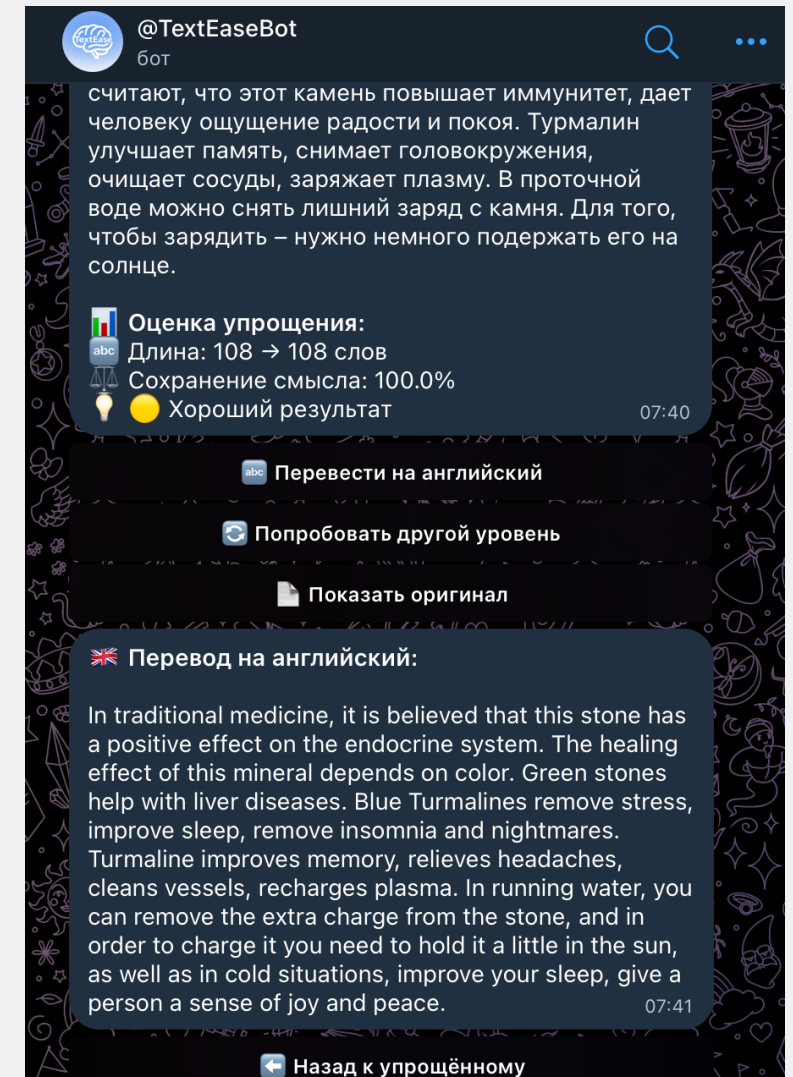
Исправляет типичные ошибки модели:

- "вы берёшь" → "вы берёте"
- "глупая" → "глубокая"
- Семантические корректировки и сохранение смысла



Этичный ИИ

Бот чётко обозначает свои ограничения и призывает к критическому мышлению



Примеры работы с текстом

Оригинал (сложный):


"Считают, что этот камень повышает иммунитет, дает человеку ощущение радости и покоя. Турмалин улучшает память, снимает головокружения, очищает сосуды, заряжает плазму. В проточной воде можно снять лишний заряд с камня. Для того, чтобы зарядить – нужно немного подержать его на солнце."

Упрощённый (средний):

"Считают, что турмалин укрепляет здоровье и улучшает настроение. Он помогает памяти, убирает головокружение и чистит сосуды. Камень можно промыть в воде и зарядить на солнце."

Перевод:

"It is believed that this stone has a positive effect on the endocrine system. Turmaline improves memory, relieves headaches, cleans vessels, recharges plasma. In running water, you can remove the extra charge from the stone, and to charge it you need to hold it a little in the sun."

 **Смысл сохранён, язык — доступный.**

Original*

The complex pathology of this condition involves **emphysematous destruction of lung parenchyma, diffuse interstitial fibrosis**, changes in the composition of lung immune cells, increased production of immunomodulatory factors and the prominent **remodeling of pulmonary vasculature**

Simplified

This complex condition involves damage to the lung tissue from **emphysema, a disease that damages the air sacs in the lungs**, and **widespread scarring of the lung tissue, called fibrosis**. The immune cells in the lungs change, and the body makes more immunomodulatory factors, substances that control the immune system. The **blood vessels in the lungs** also change a lot.

Ограничения и этические аспекты

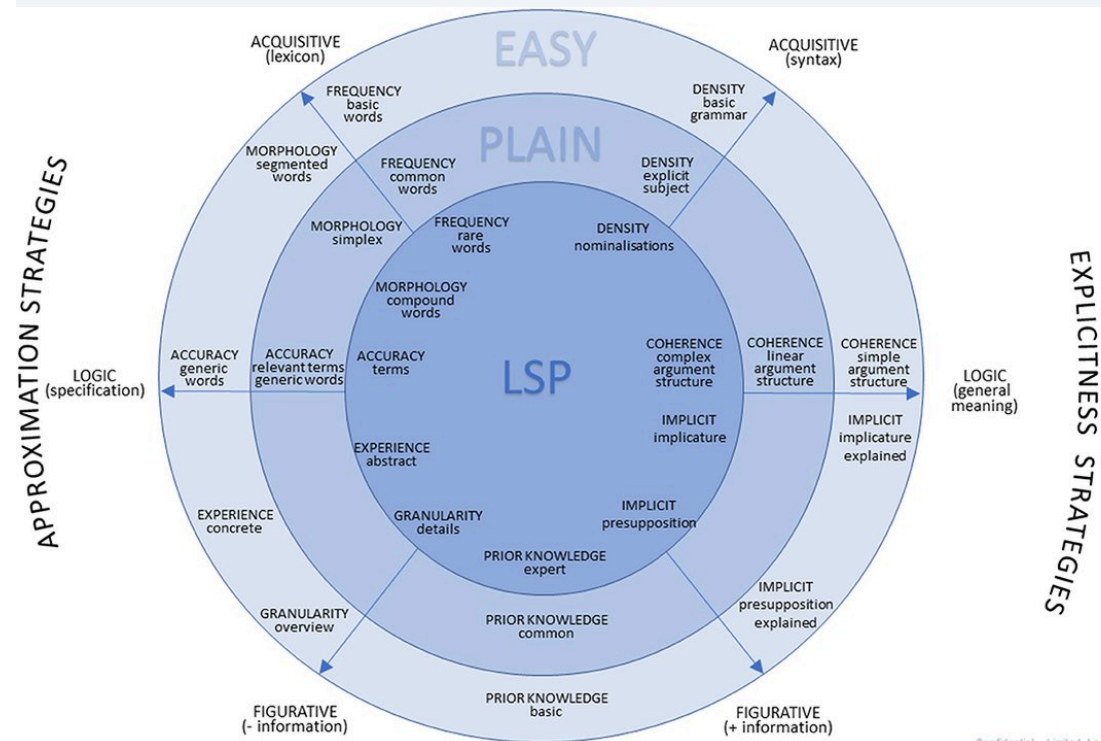
Что бот НЕ делает:

- ✗ Не гарантирует 100% точность
- ✗ Не предоставляет специализированные знания
- ✗ Может исказить смысл в поэтических или политических текстах

Как мы с этим работаем:

- ⚠ Предупреждение при старте: «Я — ИИ, а не эксперт»
- 🔍 Режим "Фактчекинг" — только разбивка на утверждения и их упрощение
- </> Пост-обработка для критических исправлений

Мы не скрываем ограничения — мы их объясняем.



Confidential – Limited Access and Use

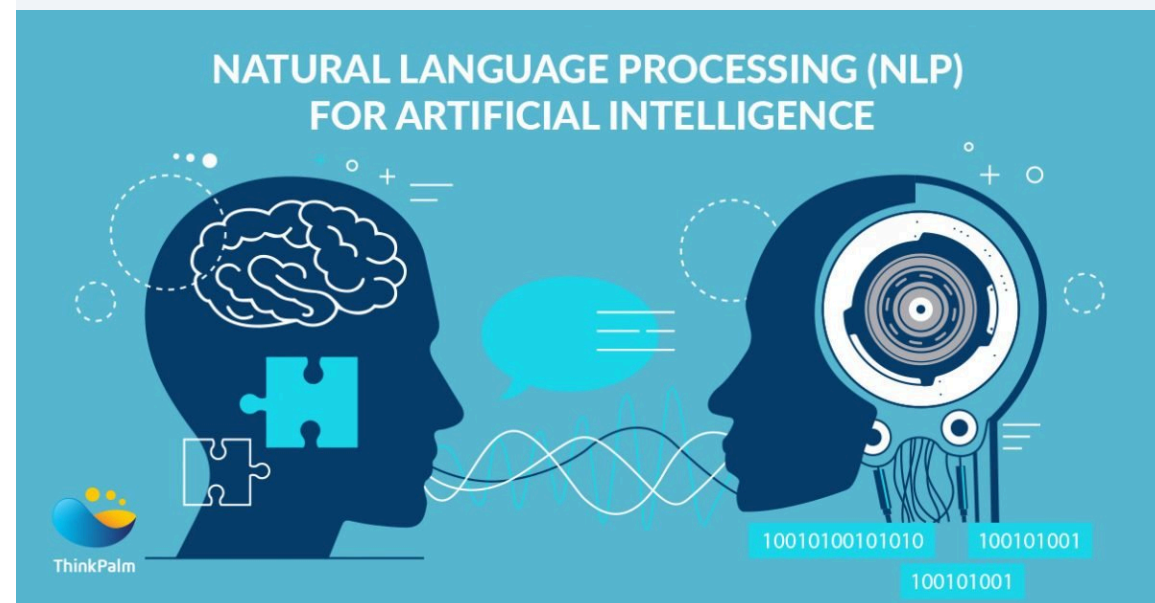
Итоги и перспективы развития

TextEaseBot — это:

- 💡 Инструмент доступности знаний
- 🔧 Практическое применение NLP
- 🎓 Отличный пример проекта для портфолио

Что можно улучшить:

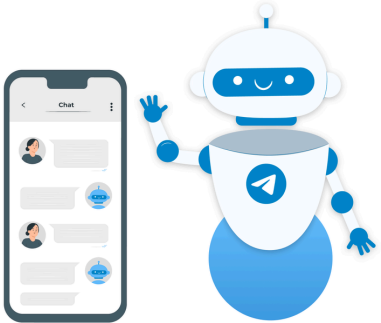
- 🔄 Дообучить модель на парах "сложный → простой" текст
- 🌐 Добавить больше языков
- 📄 Поддержка PDF
- 📤 Экспорт в DOCX/TXT
- 🗄️ Интеграция с фактчекинговыми базами



Демонстрация и контакты

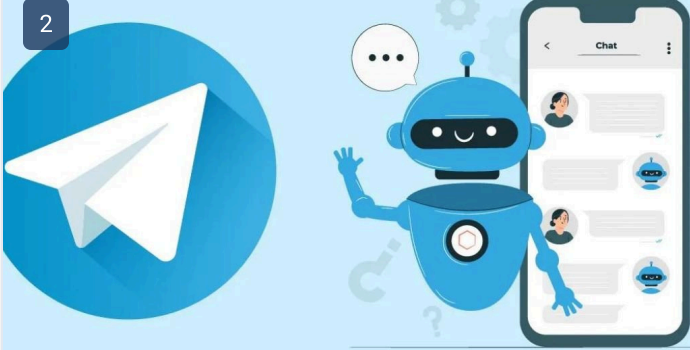
Скриншоты работы в Telegram:

1



Приветствие и функции

2



Оценка упрощения текста

3

Original*

The complex pathology of this condition involves **emphysematous destruction of lung parenchyma**, **diffuse interstitial fibrosis**, changes in the composition of lung immune cells, increased production of immunomodulatory factors and the prominent **remodeling of pulmonary vasculature**

Simplified

This complex condition involves damage to the lung tissue from **emphysema**, a **disease that damages the air sacs in the lungs**, and **widespread scarring of the lung tissue, called fibrosis**. The immune cells in the lungs change, and the body makes more immunomodulatory factors, substances that control the immune system. The **blood vessels in the lungs** also change a lot.

Перевод на английский

4



Режим "Фактчекинг"

Информация о проекте

- Проект доступен на GitHub:
github.com/gratati/TextEaseBot
- Модель: RuT5, обучена на упрощении текстов
- Email: gratati49@gmail.com
- Telegram: [@TextEaseBot](https://t.me/TextEaseBot)

- Версия: 1.2.1 (август 2025)
- Хостинг: Google Cloud Platform