

AI: обзор происходящего

Алексей Козятинский

ak239spb@gmail.com

https://twitter.com/the_kozy

План

- История
- Примеры
- Инструменты
- Интересные ссылки
- Будущее

Традиционные языковые модели

- Плохо понимали контекст...
- ... особенно отдаленный
- Очень ресурсозатратные
- Специализированные

2017: "Внимание — это все, что вам нужно"

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin

- Трансформеры
- Значительно более эффективная архитектура
- Более 100000 цитирований (4 место)
- Скачок в понимании естественных языков

2018: "GPT-1"

OpenAI

- Generative Pre-trained Transformer
- 117 миллионов параметров
- Предложения, похожие на настоящие

2019: "GPT-2"

OpenAI

- 1.5 миллиарда параметров
- Лучшее понимание контекста
- Более человеческие ответы

2020: "GPT-3"

OpenAI

- 175 миллиардов параметров
- ChatGPT (2022)
- Дополнение текста
- Перевод
- Резюмирование больших текстов
- Alignment problem с помощью реальных людей

2023: "GPT-4"

OpenAI

- ? триллионов параметров
- Sparks of the Artificial General Intelligence
<https://www.youtube.com/watch?v=qblk7-JPB2c>
- Картинки/голос и прочие способы спросить

Помощь в написании кода

- Github Copilot
- Модель предсказывает продолжение

Поиск ошибок

- <https://joel.tools/codegen/>
- Предсказание против нашего кода

Немного задачек...

— Сколько сестер?

<https://chat.openai.com/share/e5e4b417-5c5d-467a-ad82-5e633d1bf482>

— Перемещение предметов

<https://chat.openai.com/share/f3e4641c-1983-4c5a-9d1e-045d28cdaa31>

— Перемещение кота

<https://chat.openai.com/share/b8cb26c2-0964-4dd7-883b-807a8d9c2807>

— Капибара на HTML/CSS

<https://chat.openai.com/share/91f8caa9-89ce-4631-9ea4-06d757145f0f>

WTF?

— Сломанный C++

<https://chat.openai.com/share/6f0e4930-254b-4370-ab93-e1d0bd82cf18>

— Сломанная формула из таблички

<https://chat.openai.com/c/b7d9a48a-e07c-44e9-be82-8cfa22bb9179>

Код

— Bash

<https://chat.openai.com/share/1025d70e-3b89-424e-a50f-3aa6248cc2a8>

— “Перевод”

<https://chat.openai.com/share/c4293fd9-b6cb-4c36-a5b0-b5f1ccd5beeb2>

Странное — Wireshark

<https://chat.openai.com/c/b413afdf-c259-4d79-b5a0-25a5fd3b8cd7>

— Собеседования/проверка домашних

Жизненное

- Поговорить с PDF контракта
- Переписываться с арендодателем
- Проверить рукописное задание...
- ... сгенерировать новых заданий
- Поговорить о кино

- Попробовать** — ChatGPT (OpenAI)
<https://chat.openai.com/>
- Github Copilot
<https://github.com/features/copilot>
- Claude (Anthropic)
<https://claude.ai/>
- perplexity.ai
<https://www.perplexity.ai/>
- Llama 3 (Meta)
<https://llama.meta.com/llama3/>

Ссылки — Sparks of AGI

<https://www.youtube.com/watch?v=qblk7-JPB2c>

— A Hackers' Guide to Language Models

<https://www.youtube.com/watch?v=jkrNMKz9pWU>

— Practical Deep Learning

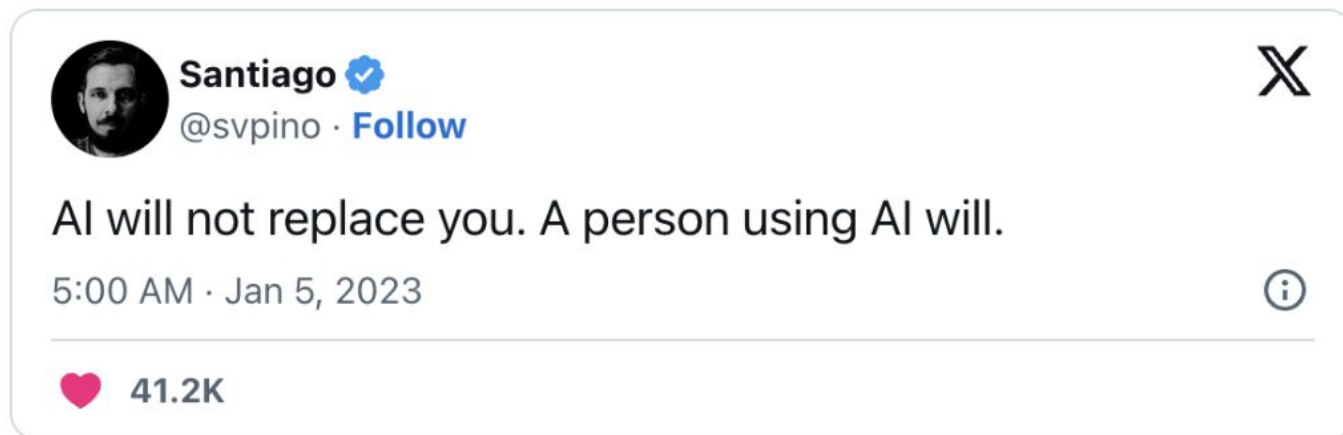
<https://course.fast.ai/>

— What Is an AI Anyway?

https://www.youtube.com/watch?v=KKNCiRWd_j0

- Будущее** — Никакого фронтенда - все чат бот?
- Никакого бэкенда - все на устройстве?
 - Никаких собеседований - все ChatGPT?
 - Никаких разработчиков - AI пишет код?
 - Никакого электричества - все на GPU?

Спасибо!



Бонус

— из чата

<https://chat.openai.com/share/9057fd9c-ad77-4d1c-83d3-ef2f227bdba2>

— преодоление реки

<https://chat.openai.com/share/705227bd-9694-47b9-82c0-af514d8b4377>

<https://chat.openai.com/share/c087d560-5c4d-41fe-88bc-ccd4798288bd>