

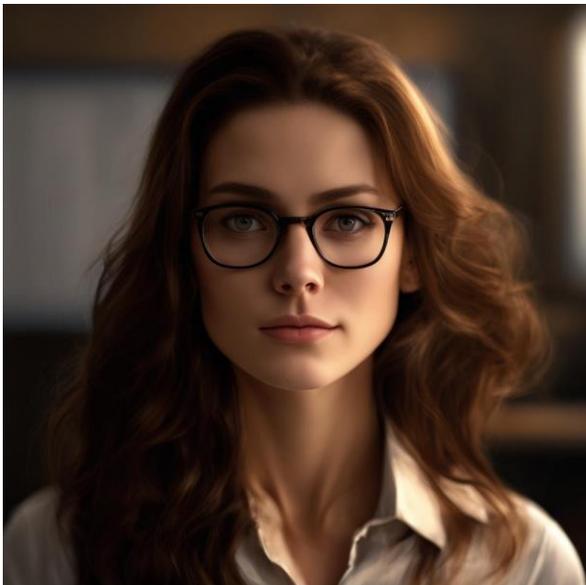
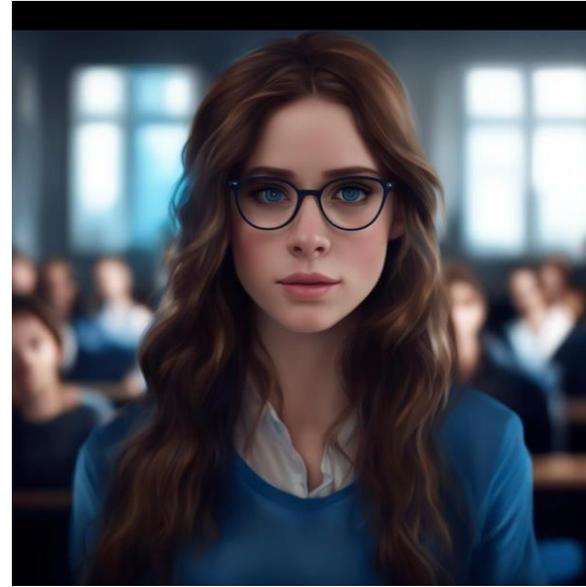
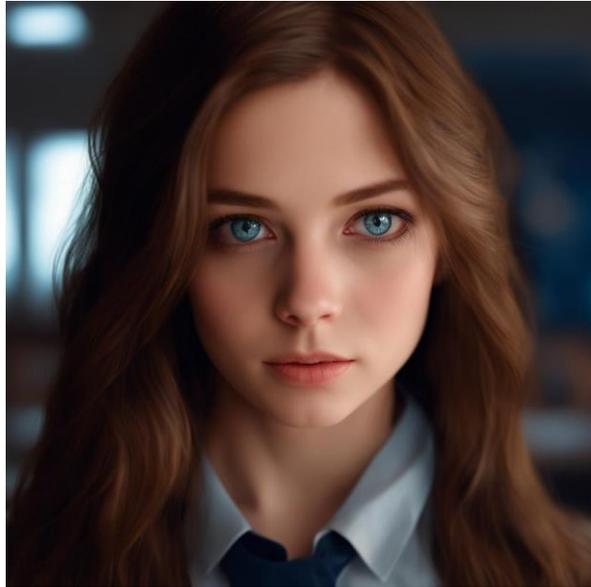
# Нейросети и ИИ: чему учиться сегодня, чтобы быть востребованным завтра

Горденко М.К.

международная премия  
**МЫВМЕСТЕ**



# Познакомимся



# Познакомимся

## Горденко Мария

- Заместитель генерального директора по информационным технологиям Завода Промышленного Машиностроения
- Академический руководитель магистерской программы «Анализ данных в девелопменте»
- Старший преподаватель НИУ ВШЭ, НИТУ МИСИС
- Лектор Российского общества «Знание»



# Познакомимся

 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

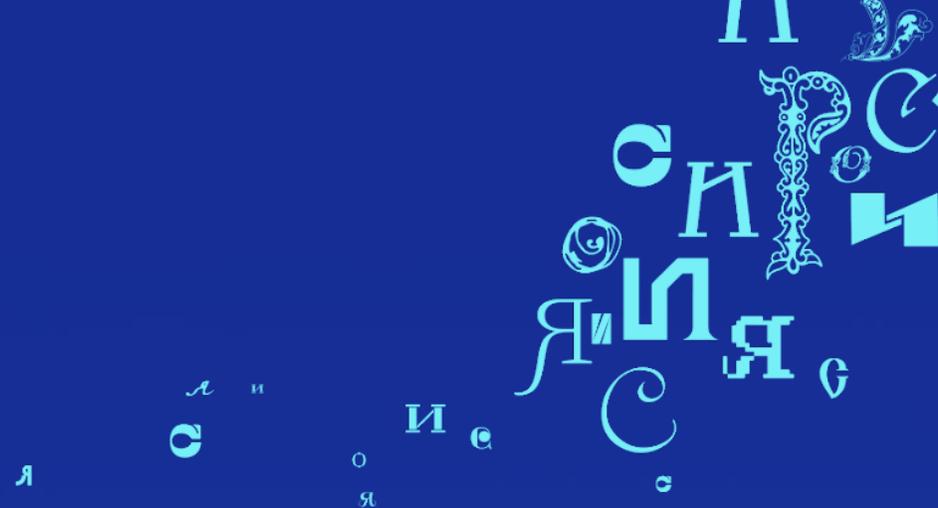
 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

 Не удастся отобразить рисунок.

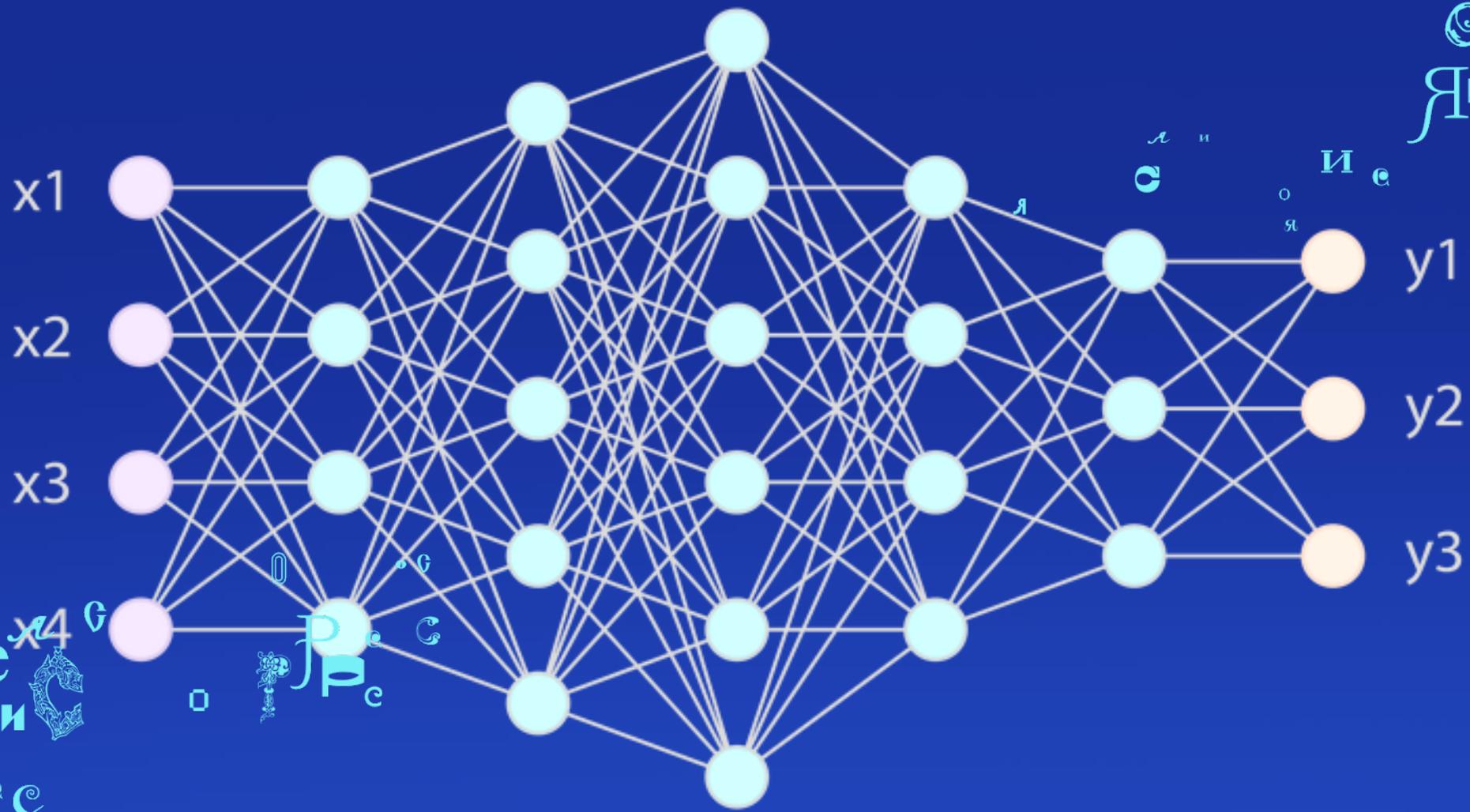
ошибка

404





А нейросети?



И  
я  
С  
я  
е  
р  
Д  
и  
я  
е  
е

И  
С  
Н  
я  
и  
я  
е

Когда появились нейросети?

Когда появились нейросети?  
1958 год, Перцептрон

В чем разница?



# Искусственный интеллект

и

о  
я

и

е

с  
я  
с



Искусственный  
интеллект



Машинное  
обучение





Искусственный  
интеллект



Машинное  
обучение



Нейросети



Искусственный  
интеллект



Машинное  
обучение



Нейросети



Глубокое  
обучение



# Слово года – 2023?

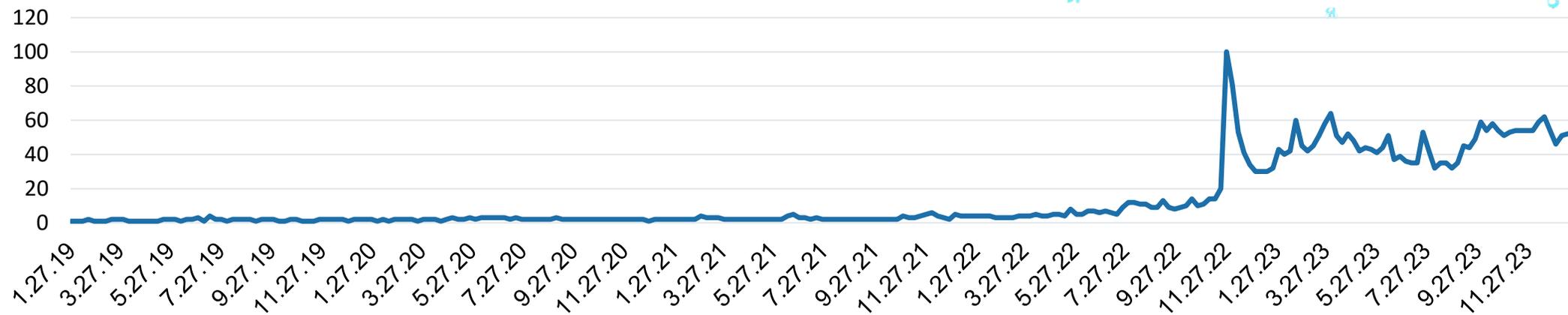




# Слово года – 2023? нейросеть\*

\* Институт русского языка имени А. С. Пушкина

## Динамика популярности (запрос: нейросеть, регион: Россия)



около 4 млн запросов в месяц



# Что произошло в ноябре 2022 года?

# Что произошло в ноябре 2022 года?

состоялся запуск ChatGPT

# Что произошло в ноябре 2022 года?

состоялся запуск ChatGPT

Сбер представил Kandinsky 2.0 — первую российскую нейросеть для генерации изображений по тексту



# ALPHAGO

- AlphaGo — программа для игры в го, разработанная в 2015 году. AlphaGo стала первой в мире программой, которая выиграла матч без гандикапа у профессионального игрока в го на стандартной доске  $19 \times 19$ , и эта победа ознаменовала собой важный прорыв в области искусственного интеллекта, так как большинство специалистов по искусственному интеллекту считало, что подобная программа не будет создана ранее 2020—2025 годов



# ALPHAGO

- В марте 2016 года программа выиграла со счётом 4:1 у Ли Седоля, профессионала 9-го дана (высшего ранга), во время исторического матча, широко освещавшегося в прессе. После победы в матче Коре́йская ассоциация паду́к присвоила AlphaGo «почётный 9-й дан» за «искренние усилия» программы в овладении мастерством игры.



# ALPHAGO

## ДОКУМЕНТАЛЬНЫЙ ФИЛЬМ ALPHAGO

- Фильм, получивший награды на многих фестивалях документального кино. Здесь рассказывается о том, как машина играла матч с чемпионом мира по игре Го корейцем Ли Седодем.

# ТРИ ЗАКОНА РОБОТЕХНИКИ

- Три закона роботехники в научной фантастике — обязательные правила поведения для роботов, впервые сформулированные Айзеком Азимовым в рассказе «Хоровод» (1942).
  - Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинён вред.
  - Робот должен повиноваться всем приказам, которые даёт человек, кроме тех случаев, когда эти приказы противоречат Первому Закону.
  - Робот должен заботиться о своей безопасности в той мере, в которой это не противоречит Первому или Второму Законам.

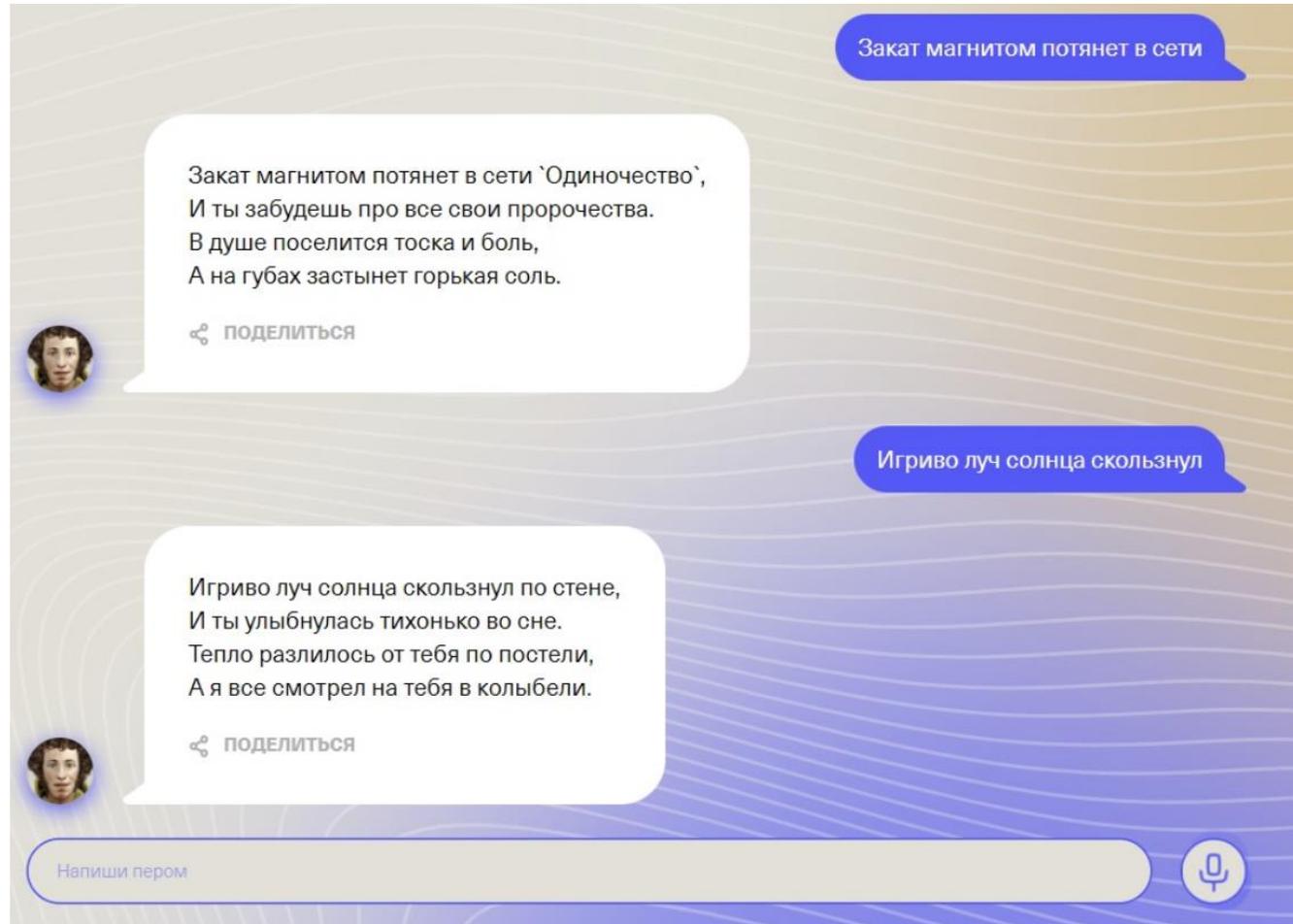


# «ЦИФРОВАЯ РЕИНКАРНАЦИЯ»: ЧЕЛОВЕКА ПОСЛЕ СМЕРТИ МОЖНО ПРЕВРАТИТЬ В ЧАТ-БОТ

- **Что такое цифровая реинкарнация в чат-бота?**
- Это цифровые копии живых или умерших личностей в виде интеллектуального цифрового собеседника. Они будут выполнены в форме программы-собеседника, которая выясняет потребности пользователей, а затем помогает удовлетворить их. Автоматическое общение с ведется с помощью текста или голоса



# «АІ ДА ПУШКИН»: НЕЙРОСЕТЬ ПИШЕТ СТИХИ В СТИЛЕ ПОЭТА



# УГРОЗА ЧЕЛОВЕЧЕСКОМУ ДОСТОИНСТВУ

- Джозеф Вейзенбаум утверждал в 1976 году, что технология искусственного интеллекта не должна использоваться для замены людей на должностях, требующих уважения и заботы, таких как:
  - Сотрудник службы поддержки клиентов (технологии ИИ уже используется сегодня в телефонных системах интерактивных голосовых ответов)
  - Терапевт (как было предложено Кеннетом Колби в 1970-х годах)
  - Сиделка для пожилых людей (как рассказала Памела МакКордак в своей книге «Пятое поколение»)
  - Солдат
  - Судья
  - Полицейский

# СУЩЕСТВУЮЩИЕ НОРМЫ

В России в 2021 году был принят первый в своём роде «Кодекс этики искусственного интеллекта», описывающий правила поведения бизнеса в этой сфере. Он был разработан отраслевым «Альянсом ИИ» и Аналитическим центром при Правительстве РФ. Под ним подписались такие компании и организации как «Сбербанк», «Яндекс», «Росатом», НИУ ВШЭ, МФТИ, ИТМО, «Ростелеком», ЦИАН и другие



# ЧТО ТАКОЕ КОДЕКС ЭТИКИ В СФЕРЕ ИИ?

- Это единая система рекомендательных принципов и правил, предназначенных для создания среды доверенного развития технологий искусственного интеллекта в России
  - Носит рекомендательный характер
  - Присоединение осуществляется на добровольной основе
  - Распространяется только на гражданские разработки



Яндекс



ИСП РАН



ITMO

INNOVATION UNIVERSITY

Skoltech  
Skolkovo Institute of Science and Technology

## ОГЛАВЛЕНИЕ

Кодекс этики в сфере искусственного интеллекта	4
<b>РАЗДЕЛ 1. ПРИНЦИПЫ ЭТИКИ И ПРАВИЛА ПОВЕДЕНИЯ</b>	
<b>1. ГЛАВНЫЙ ПРИОРИТЕТ РАЗВИТИЯ ТЕХНОЛОГИЙ ИИ В ЗАЩИТЕ ИНТЕРЕСОВ И ПРАВ ЛЮДЕЙ И ОТДЕЛЬНОГО ЧЕЛОВЕКА</b>	<b>6</b>
1.1. Человеко-ориентированный и гуманистический подход.	6
1.2. Уважение автономии и свободы воли человека.	6
1.3. Соответствие закону.	6
1.4. Недискриминация.	7
1.5. Оценка рисков и гуманитарного воздействия.	7
<b>2. НЕОБХОДИМО ОСОЗНАВАТЬ ОТВЕТСТВЕННОСТЬ ПРИ СОЗДАНИИ И ИСПОЛЬЗОВАНИИ ИИ</b>	<b>8</b>
2.1. Риск-ориентированный подход.	8
2.2. Ответственное отношение.	8
2.3. Предосторожность.	9
2.4. Непричинение вреда.	9
2.5. Идентификация ИИ в общении с человеком.	9
2.6. Безопасность работы с данными.	9
2.7. Информационная безопасность.	10
2.8. Добровольная сертификация и соответствие положениям Кодекса.	10
2.9. Контроль рекурсивного самосовершенствования СИИ.	10
<b>3. ОТВЕТСТВЕННОСТЬ ЗА ПОСЛЕДСТВИЯ ПРИМЕНЕНИЯ СИИ ВСЕГДА НЕСЕТ ЧЕЛОВЕК</b>	<b>10</b>
3.1. Поднадзорность.	10
3.2. Ответственность.	11

<b>4. ТЕХНОЛОГИИ ИИ НУЖНО ПРИМЕНЯТЬ ПО НАЗНАЧЕНИЮ И ВНЕДРЯТЬ ТАМ, ГДЕ ЭТО ПРИНЕСЁТ ПОЛЬЗУ ЛЮДЯМ</b>	<b>11</b>
4.1. Применение СИИ в соответствии с предназначением.	11
4.2. Стимулирование развития ИИ.	11
<b>5. ИНТЕРЕСЫ РАЗВИТИЯ ТЕХНОЛОГИЙ ИИ ВЫШЕ ИНТЕРЕСОВ КОНКУРЕНЦИИ</b>	<b>12</b>
5.1. Корректность сравнений СИИ.	12
5.2. Развитие компетенций.	12
5.3. Сотрудничество разработчиков.	12
<b>6. ВАЖНА МАКСИМАЛЬНАЯ ПРОЗРАЧНОСТЬ И ПРАВДИВОСТЬ В ИНФОРМИРОВАНИИ ОБ УРОВНЕ РАЗВИТИЯ ТЕХНОЛОГИЙ ИИ, ИХ ВОЗМОЖНОСТЯХ И РИСКАХ</b>	<b>13</b>
6.1. Достоверность информации о СИИ.	13
6.2. Повышение осведомлённости об этике применения ИИ.	13

## РАЗДЕЛ 2. ПРИМЕНЕНИЕ КОДЕКСА

<b>1. ОСНОВЫ ДЕЙСТВИЯ КОДЕКСА</b>	<b>15</b>
1.1. Правовая основа Кодекса.	15
1.2. Термины.	15
1.3. Акторы ИИ.	15
<b>2. МЕХАНИЗМ ПРИСОЕДИНЕНИЯ И РЕАЛИЗАЦИИ КОДЕКСА</b>	<b>16</b>
2.1. Добровольность присоединения.	16
2.2. Уполномоченные по этике и/или комиссии по этике.	16
2.3. Комиссия по реализации Национального кодекса в сфере этики ИИ.	17
2.4. Реестр участников Кодекса.	17
2.5. Разработка методик и руководств.	17
2.6. Свод практик.	18

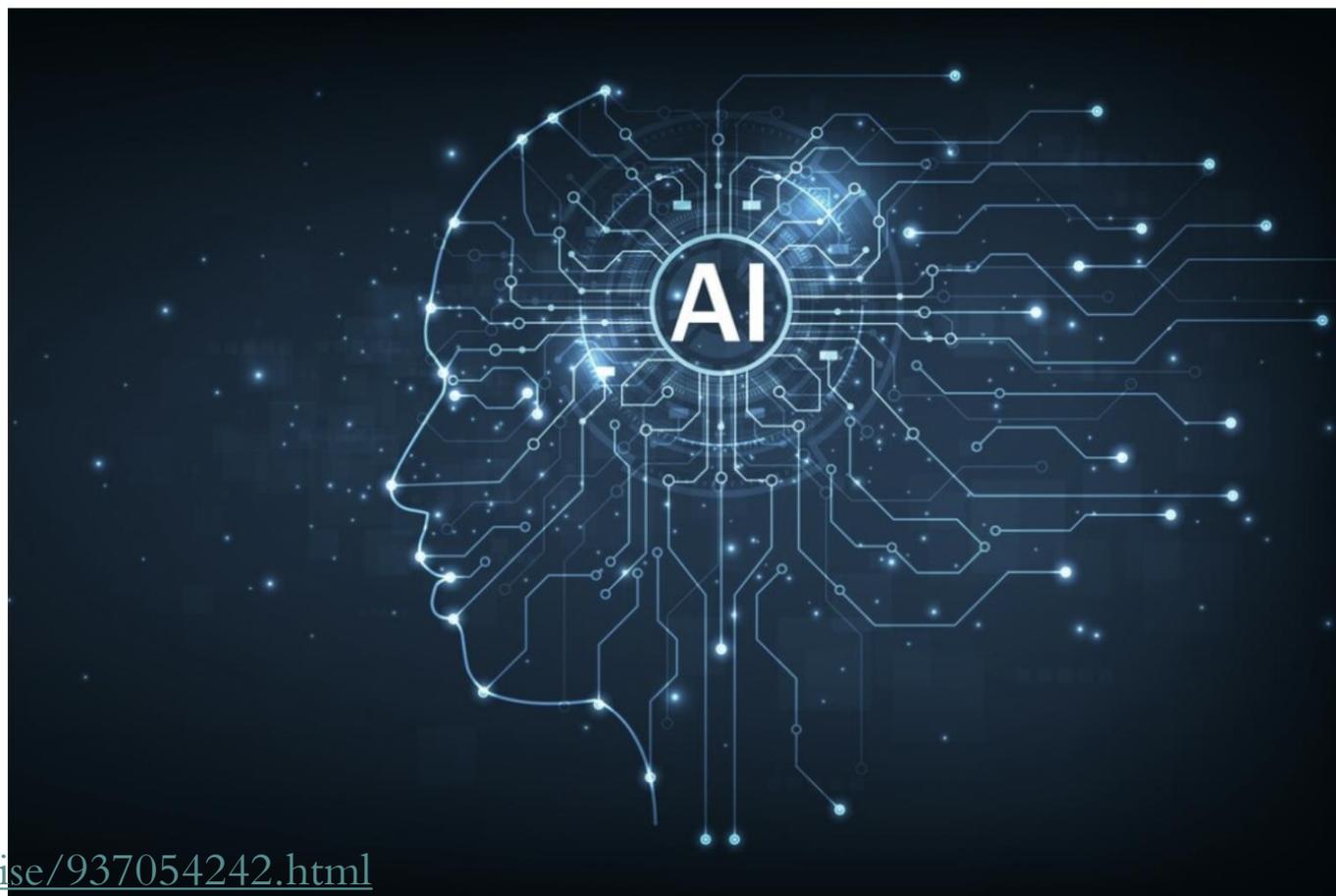


27

ИЮН  
2024

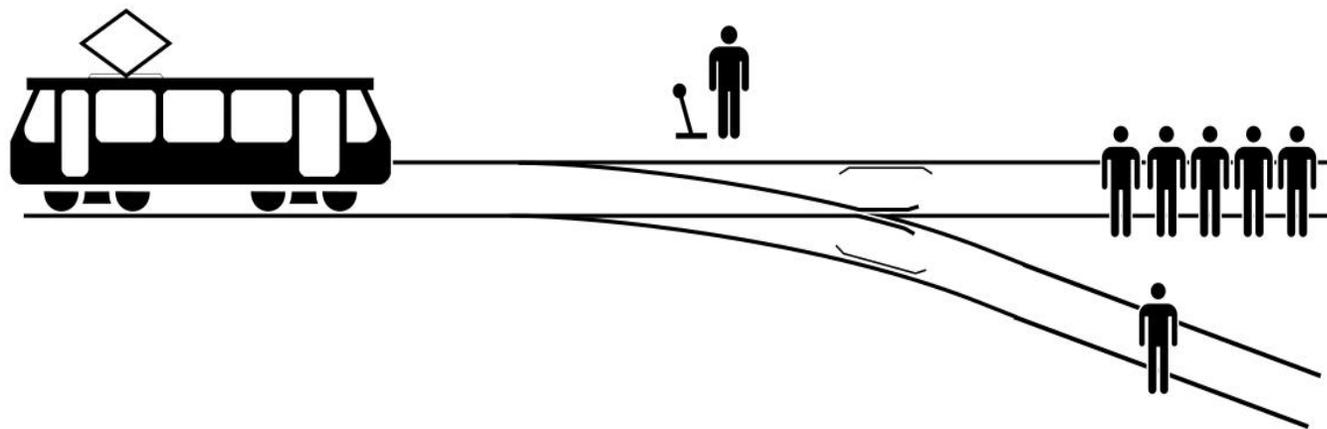
👁 2233

# НИУ ВШЭ первым из российских университетов разработал и утвердил Декларацию этических принципов использования ИИ



# ПРАКТИЧЕСКИЕ ПРИМЕРЫ И КЕЙСЫ

- Проблема вагонетки и ИИ



# Возможности сервисов ИИ

## Что умеет:

- генерация идей и тезисов (по ключевым словам или темам)
- перевод, стилизация и «перефразирование» текста
- объяснение сложных вопросов простым языком
- генерация текста по некоторым вводным
- анализ речи и текста по различным критериям
- написание и анализ компьютерного кода на разных языках программирования

## Что не умеет:

- отсутствует верификация с внешними источниками и какой-либо подтвержденной базой знаний
- может выдать неправильные ответы, сослаться на несуществующих людей и несуществующие научные работы для подтверждения своих тезисов
- в модели отсутствует система логического вывода, поэтому она может ошибаться в рассуждениях, и делать неверные выводы, хотя текст выглядит согласованным и убедительным

## Возможности сервисов ИИ

### Основная опасность

# — отключение критического мышления пользователя в процессе работы с ИИ

Что умеет:

- генерация идей и тезисов (по ключевым словам)
- перевод, стилизация и «перефразирование» текста
- объяснение сложных тем простым языком
- генерация текста по заданным тезисам
- анализ речи и текста по различным критериям
- написание и анализ компьютерного кода на разных языках программирования

Что не умеет:

- отсутствует верификация с внешними источниками информации и подтвержденной базой знаний
- может выдать неправильные ответы, которые могут повлиять на мнение людей и несуществующие научные работы для цитирования
- в модели отсутствует система логического вывода, поэтому она может ошибаться в рассуждениях, и делать неверные выводы, хотя текст выглядит согласованным и убедительным

# Новые профессии – мы можем научить этому студентов?

## Вырос спрос на профессии:

- Специалист по машинному обучению
- Инженер данных
- Data Scientist
- Продюсер в образовании
- Слесарь
- Электрик
- Фрезеровщик
- Инженер-электронщик и т.д.

# Каждому по EdGPT?

- С обучением моделей на еще больших данных мы сталкиваемся с ленью, они начинают отвечать хуже
- К 2030–2050 мы будем испытывать недостаток качественных данных, на которых можно было бы обучать общие модели. Сейчас сложности со сбором размеченных данных
- Необходимо развитие фундаментальных технологий (квантовые, нейроморфные вычисления) – нам не хватает вычислительных мощностей, особенно в условиях санкций. Проблема отсутствия технологических прорывов и фокуса на software
- Это ОЧЕНЬ дорого. Каждый день у OpenAI, например, уходит порядка 700 тысяч долларов на поддержание ChatGPT.

# Фотографии с лишними объектами





\* Скринкаст работы с сервисом jrgtm

# Черно-белые фотографии



Colorizerbot

bot



**What can this bot do?**

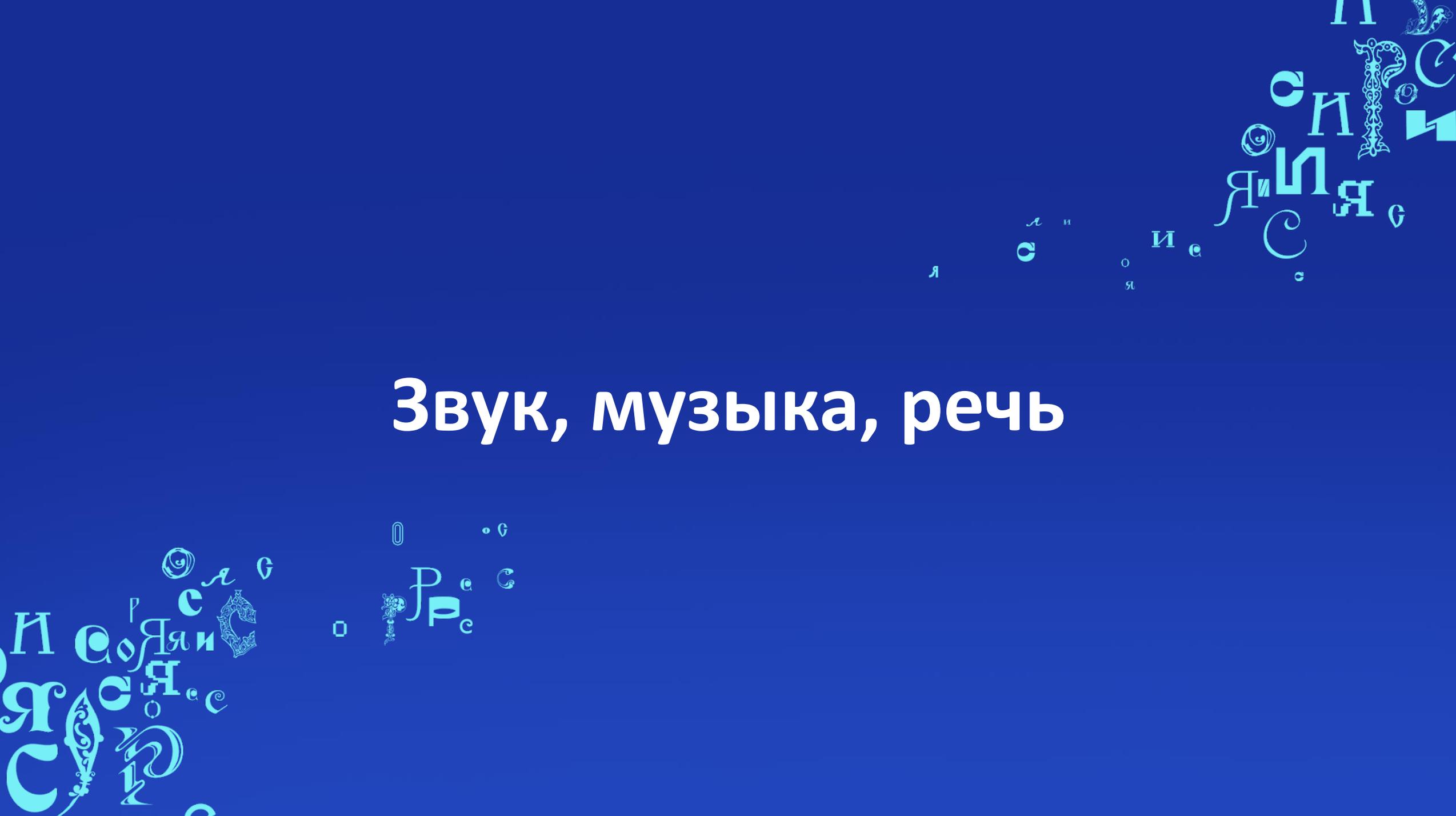
Я бот который может сделать из черно-белой фотографии цветную.

Канал с новостями [@photocolorizer](#)

START



# Звук, музыка, речь



**Образца голоса в 3 секунды  
достаточно, чтобы его симитировать**

# Видео, фильмы и даже прогноз погоды

Модель  
Анимация

Размер  
16:9 (854 x 480)

00:00 | 00:00

Приблизительно 1,5 минуты Отменить

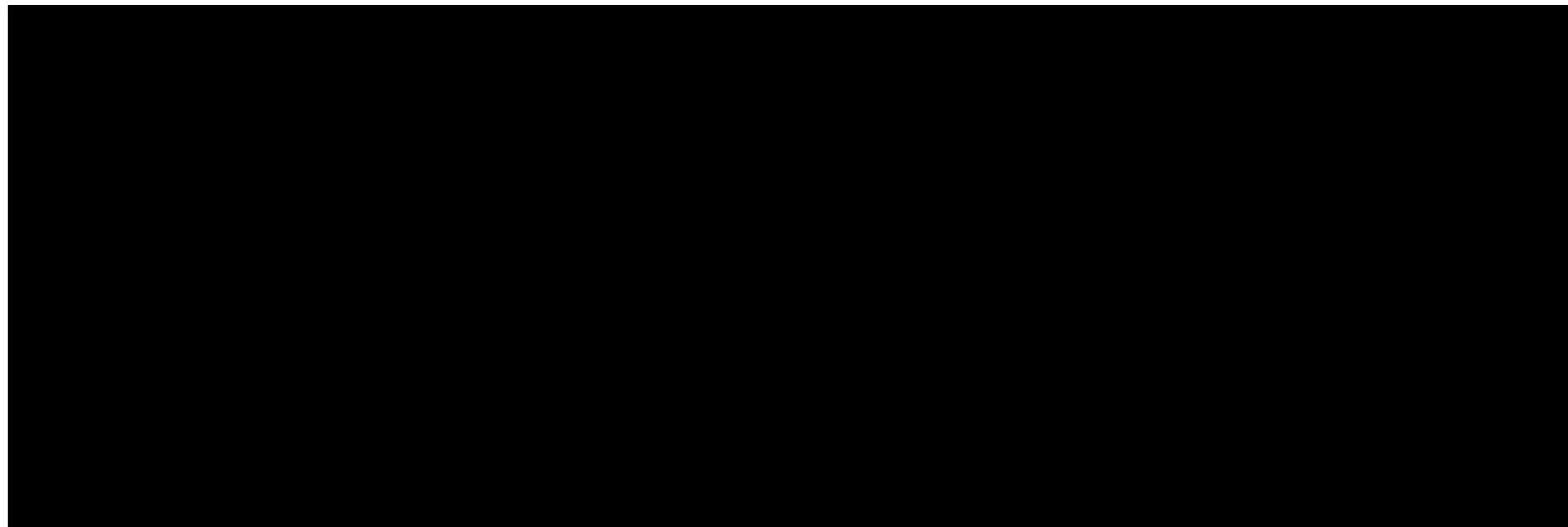
= Сцена 1 · 4 секунды

Девушка на сцене в черном костюме читает лекцию, на слайде у много фотографий. Видео с разных ракурсов. Зал сделан очень современно в

Направление камеры ← →

Image prompts: tiger, smiley face, moon, moon

Девушка на сцене в черном костюме читает лекцию, на слайде много фотографий. Видео с разных ракурсов. Зал сделан очень современно в черно-серых цветах.

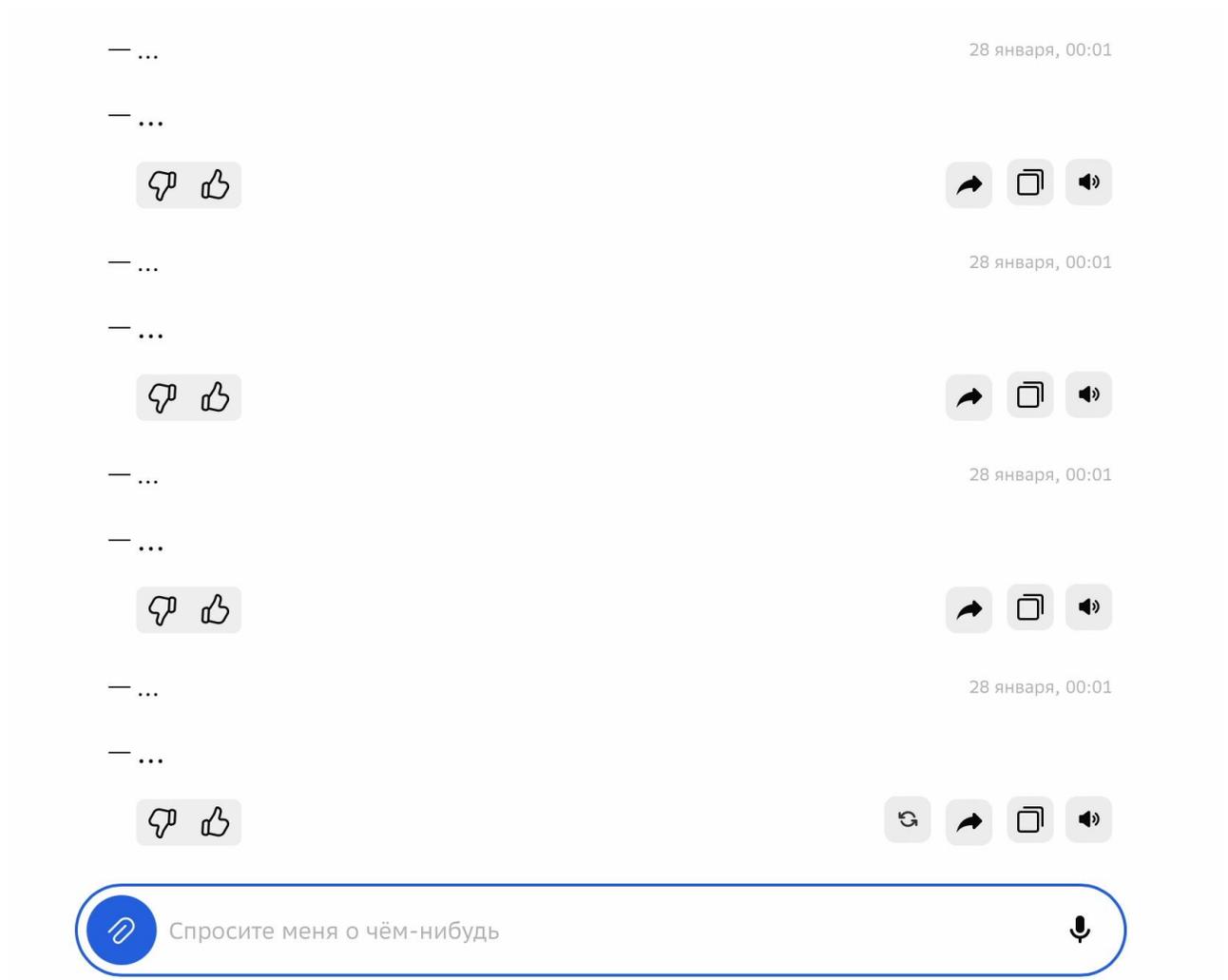


Пример дипфейка



# Тексты



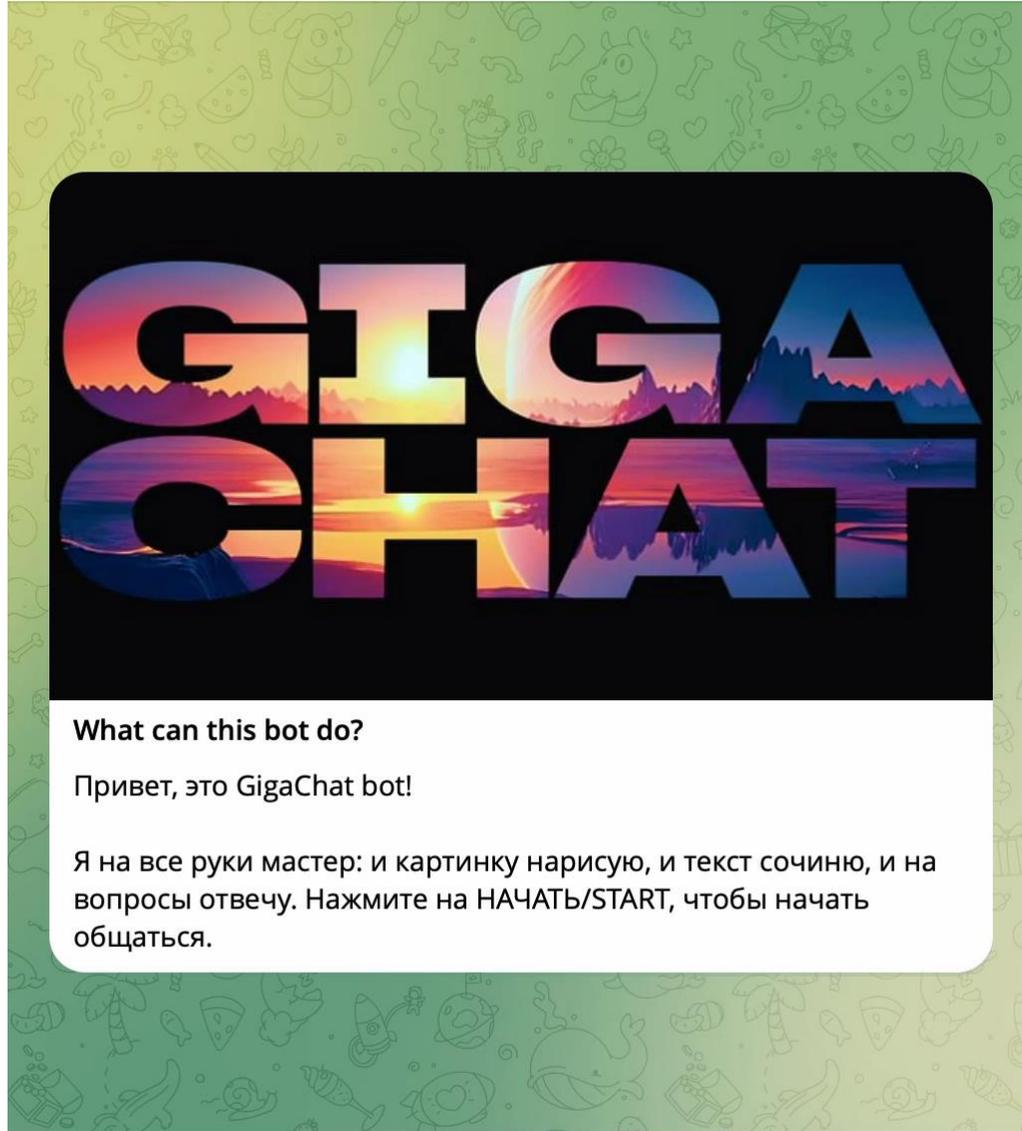


\* Скринкаст работы с диалоговой нейросетевой моделью GigaChat

63



GigaChat  
bot



**What can this bot do?**

Привет, это GigaChat bot!

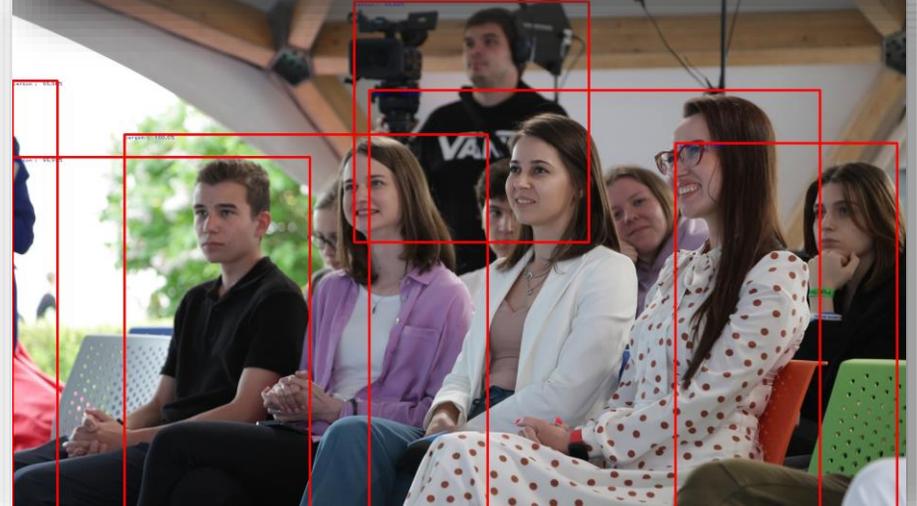
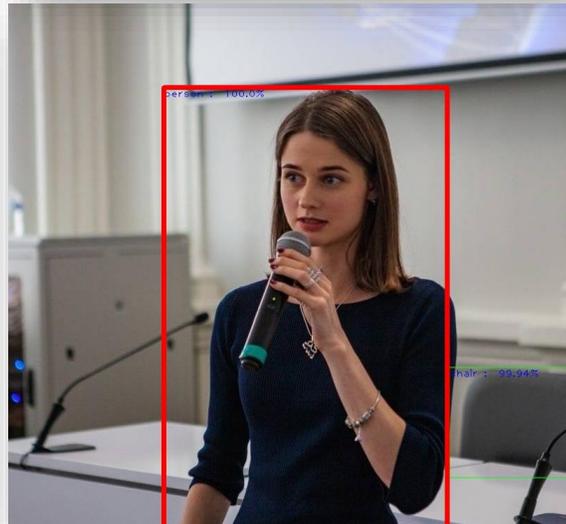
Я на все руки мастер: и картинку нарисую, и текст сочиню, и на вопросы отвечу. Нажмите на НАЧАТЬ/START, чтобы начать общаться.

[START](#)



\* Скринкаст работы с диалоговой нейросетевой моделью GigaChat

# Распознавание фото, видео, речи





## Опасения

- ошибки;
- предвзятость;
- конфиденциальность;
- замена людей/потеря работы;
- этика;
- ограничения.



## Какие профессии может заменить нейросеть?\*

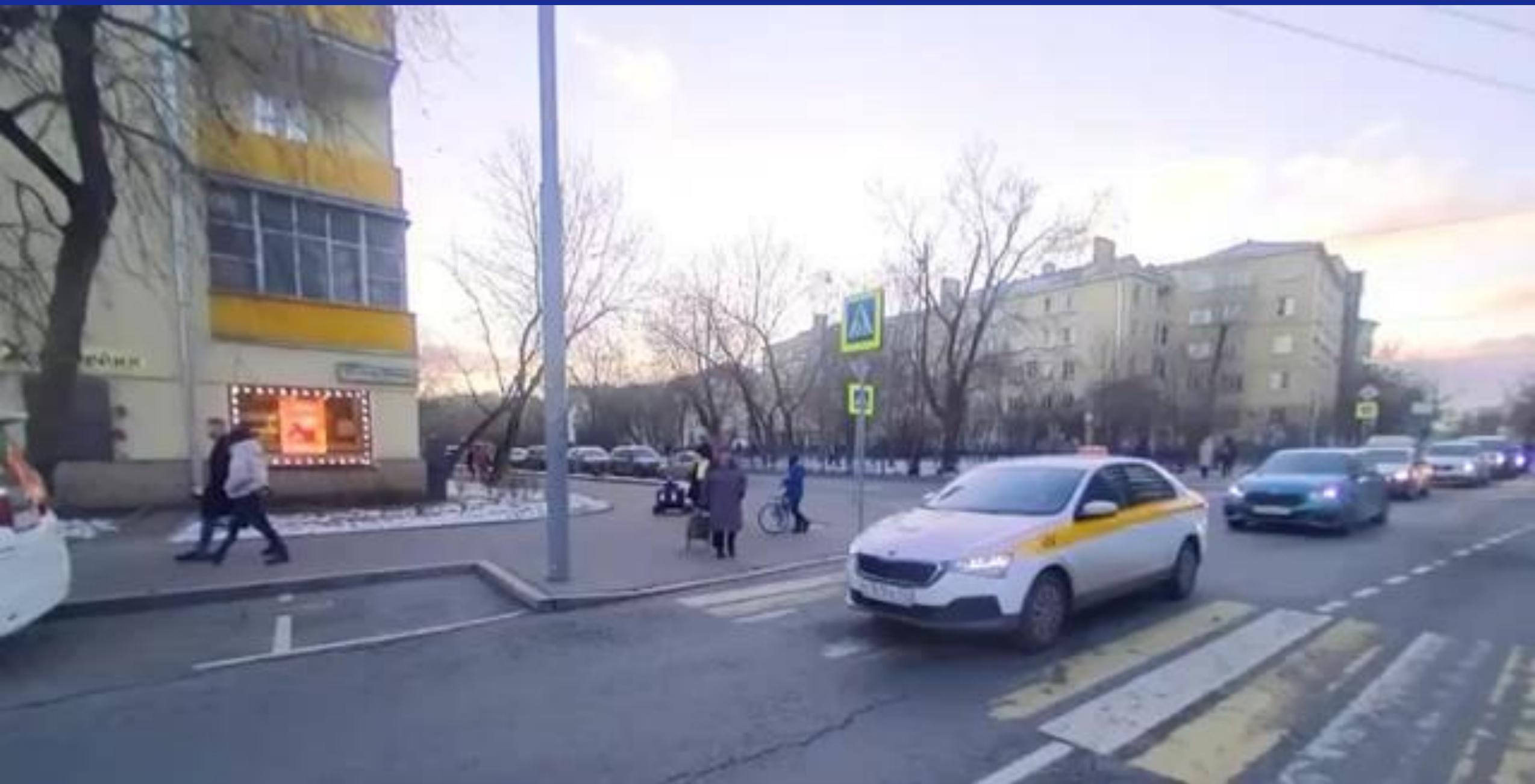
- операторы;
- водители;
- репетиторы;
- журналисты;
- медики;
- дизайнеры;
- юристы.

\* По данным популярной нейросети



Одни профессии исчезнут – другие  
ПОЯВЯТСЯ

Чему учиться сегодня?



Чему учиться сегодня?  
Век живи – век учись.

Робототехника

Образование

Медицина

Транспорт

Развлечения

Финансы

Прогноз погоды

Маркетинг

Энергетика

Искусство

# Спасибо за внимание

## Горденко Мария

- Заместитель генерального директора по информационным технологиям Завода Промышленного Машиностроения
- Академический руководитель магистерской программы «Анализ данных в девелопменте»
- Старший преподаватель НИУ ВШЭ, НИТУ МИСИС
- Лектор Российского общества «Знание»

