

ГЕНЕРАТИВНЫЕ МОДЕЛИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И КАК ОНИ МЕНЯЮТ БИЗНЕС

Андрей Самарин

Архитектор клиентских решений
Cloud

Cloud – это



лидирующий провайдер услуг
и сервисов по моделям IaaS и PaaS
для физлиц, стартапов, малого, среднего,
крупного бизнеса, государственных
учреждений и корпораций

Разработчик платформы машинного
обучения полного цикла **ML Space**

Создатель суперкомпьютеров
Christofari и **Christofari Neo**

№1

на рынке функционального PaaS
по выручке (IKS Consulting), 2021

№2

на рынке облачной инфраструктуры
России по выручке (IKS Consulting), 2021

№3

в списке быстрорастущих
ИТ-компаний (TAdviser), 2022

№50

в списке самых производительных
суперкомпьютеров в мире

Прогнозирование спроса и управление запасами

Прогнозирование спроса на продукцию в зависимости от региона и других параметров, например, погодных условий, мероприятий и т.д.

Умные дороги

Повышение безопасности дорожного движения и сокращение задержек в цепочках поставок

Управление маршрутом доставки и оптимизация цепочек поставок

Оптимизация дистанции и траектории путей доставки, а также снижение времени доставки

Автоматизация складов и центров обработки

Повышение эффективности работы складов и уменьшение времени обработки

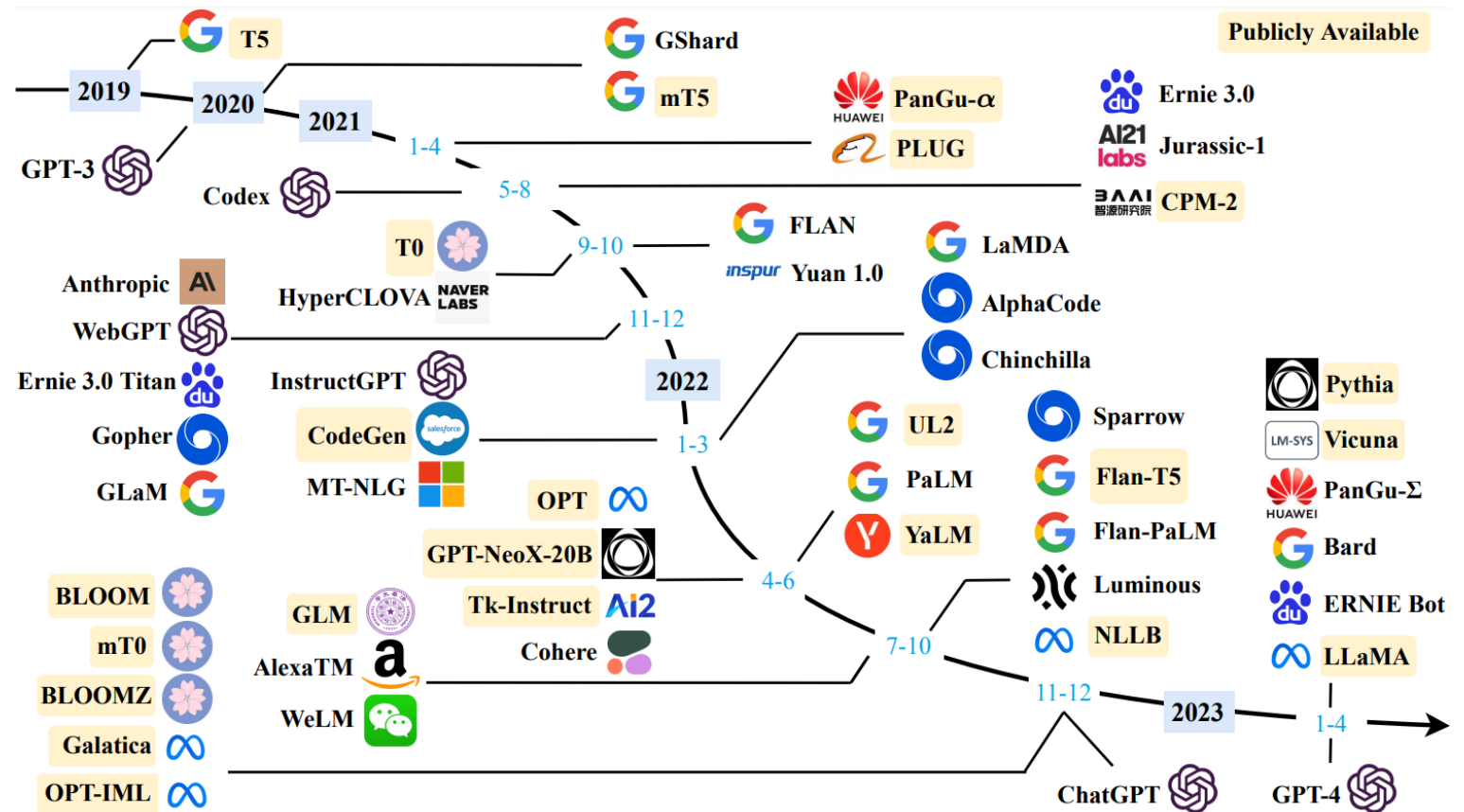
Автономный транспорт

Повышение безопасности, сокращение сроков доставки и повышение эффективности

БОЛЬШИЕ ЯЗЫКОВЫЕ МОДЕЛИ

ChatGPT изменил все,
но технология возникла
не за один день

Благодаря более качественным,
разнообразным и связанным текстам, а
также простоте использования и
доступности, в том числе по API, ChatGPT
получил 1 млн. пользователей уже на 5
день после запуска.



Источник: A Survey of Large Language Models, 2023, arXiv. Хронология крупных языковых моделей, обученных более, чем на 10 млрд параметров. Общедоступные модели отмечены желтым цветом.

Еще более **универсальная** модель, которая:

1. Может распознавать, понимать и анализировать изображения, в том числе текст и код, написанные от руки
2. Генерирует качественный контент на 26 языках, в том числе на русском
3. Запоминает еще больше контекста и более последовательна, чем ChatGPT
4. Модель сложнее сбить с толку или обмануть
5. Умеет писать код на разных языках программирования

User What is unusual about this image?



Source: [Barnorama](#)

GPT-4 The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.

ПОКА СУЩЕСТВУЮЩИЕ НЕДОСТАТКИ CHATGPT И GPT-4



Мы спросили саму модель

1. **Ограниченное понимание.** Обученные на основе статистических закономерностей, они не могут осмыслить предложения, как делают люди.
2. **Зависимость от больших объемов данных.** Требуют огромного количества данных для дообучения.
3. **Непостоянство результатов.** Могут быть непредсказуемыми, порождая качественно разные результаты при запросах даже с небольшими изменениями ввода.
4. **Сложность контроля.** Ограничение и регулирование вывода моделей может быть сложным процессом.
5. **Неполнота знаний.** Обучение заканчивается на определенной дате, и модель не может учесть новые события или информацию.
6. **Ошибки и неточности.** Могут давать неверные или неточные ответы.
7. **Безопасность данных.** Могут представлять риски для безопасности и конфиденциальности данных, особенно при обработке чувствительной информации.

A

Оркестр из 120 музыкантов может сыграть девятую симфонию Бетховена за 40 минут. Сколько времени понадобится 60 музыкантам, чтобы сыграть эту мелодию?



Для решения задачи можно использовать правило обратно пропорциональности. Если количество музыкантов уменьшается вдвое, то время необходимое для исполнения увеличивается вдвое.

Таким образом, время, необходимое 120 музыкантам для исполнения девятой симфонии Бетховена, пропорционально количеству музыкантов:

120 музыкантов - 40 минут

60 музыкантов - x

Подставим значения в формулу:

$$120/60 = 40/x$$

Решим уравнение:

$$120x = 60 * 40$$

$$x = (60 * 40) / 120$$

$$x = 20$$

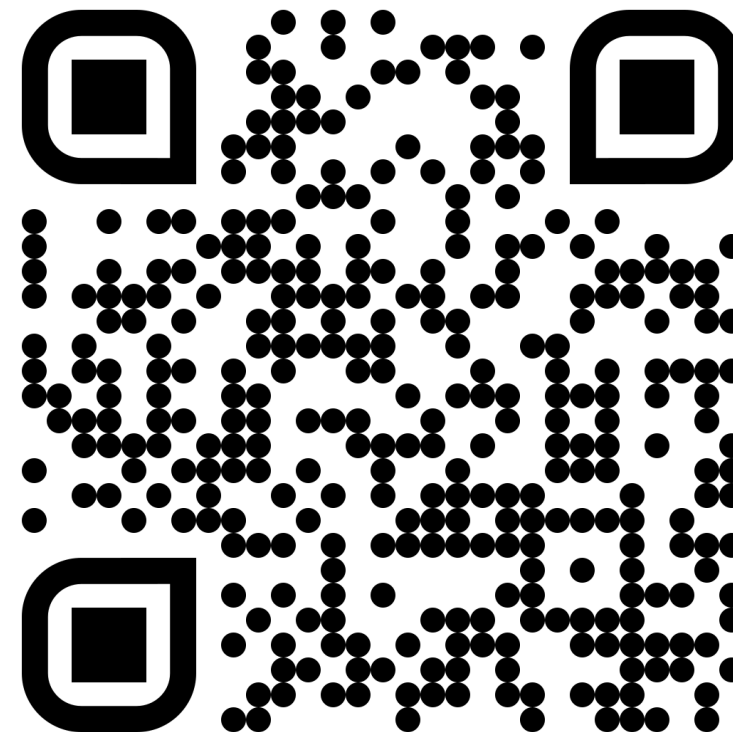
Таким образом, 60 музыкантам потребуется 20 минут для того, чтобы исполнить девятую симфонию Бетховена.

СЦЕНАРИИ ПРИМЕНЕНИЯ CHATGPT



- 01 Поддержка клиентов и обработка запросов
- 02 Мониторинг и отслеживание транспорта
- 03 Системы управления запасами
- 04 Прогнозирование спроса на перевозки
- 05 Оптимизация маршрутов доставки

Опрос от разработчиков языковых моделей



Расскажите, какие задачи вы хотели бы решать при помощи таких моделей

ЧТО ПРЕДЛАГАЕТ CLOUD



Экосистема инструментов
и решений для ML разработчиков
и Data Science специалистов

ML Space

- Платформа полного цикла для разработки моделей машинного обучения
- С возможностью обучения на двух мощных суперкомпьютерах Christofari и Christofari NEO
- Для R&D и больших Data Science команд

AI Marketplace

- Витрина артефактов машинного обучения: моделей, датасетов, окружений
- Предобученные AI-сервисы, готовые для интеграции в приложения
- Сервисы собственной разработки Cloud

On-Premise

- Возможность установки в закрытом контуре заказчика

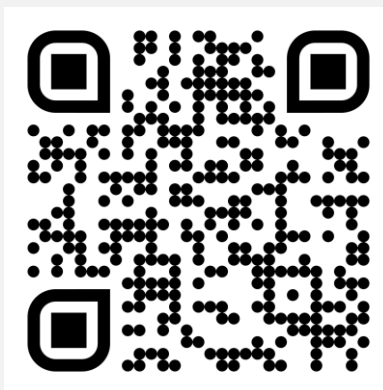
СПАСИБО
ЗА ВНИМАНИЕ!



ML Space

- Платформа для ML-разработки полного цикла
- Совместная работа DS-команд
- На базе двух мощных суперкомпьютеров Christofari и Christofari NEO
- Есть возможность установки on premises

Для юридических лиц и ИП.



**Андрей
Самарин**

Архитектор клиентских решений,
Cloud

Моб. тел.: +7 (985) 990-13-87
E-mail: ansamarin@sbercloud.ru

ООО «Облачные технологии»