

# Ключевые тренды в области ИИ

Андрей Черток

Управляющий директор, Sber AI



## Андрей **ЧЕРТОК**

Ph.D., Управляющий директор,  
Sber AI

**Опыт разработки моделей – более 10 лет**  
**Управленческий опыт – более 6 лет**

Руководит R&D проектами в области ML/AI: SberSight,  
Digital Manager, Graph, TagMe

Взаимодействие с корпорациями и научными группами

Развивает науку в области ML/AI

Окончил факультет Вычислительной математики и  
кибернетики Московского государственного университета  
им. Ломоносова (математик, системный программист)

\* DS/AI Community – сообщество специалистов Сбербанка, работающих в области  
анализа данных и искусственного интеллекта

# О чем сегодня поговорим



- Кривая Гартнера
- Тренды 2020
- Self Supervised Learning (SSL)
- Contrastive Learning
- Трансформеры
- Робастность моделей
- Предвзятость моделей
- AlphaFold
- Графовая нейронная сеть (GNN)
- Проекты Sber AI

# Кривая хайп-цикла Гартнера



# Облако популярных тегов NeurIPS 2020



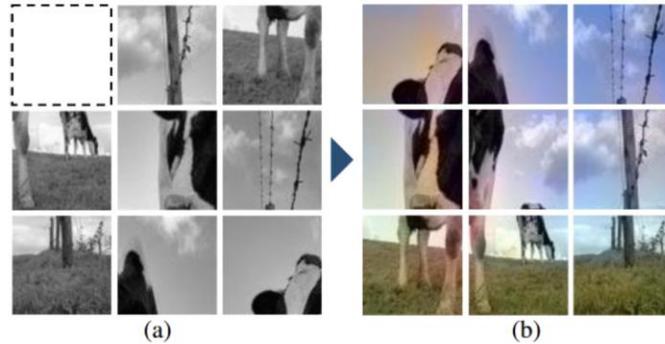
# Self Supervised Learning (SSL)

- подход, который позволяет выучить репрезентативные представления объектов, имея большую и неразмеченную выборку

## Примеры задач:

### Восстановление порядка частей изображения

Нейросеть учится понимать, как именно расположены части изображения относительно друг друга, и восстанавливать порядок

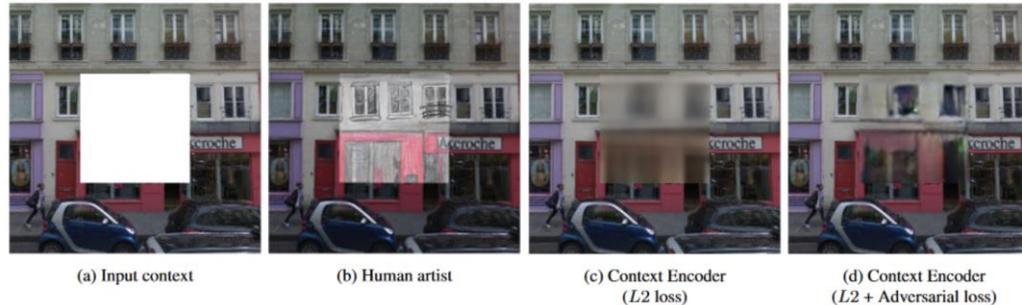


### Общая идея подхода

имея выборку (изображений, текстов) придумать какую-то задачу и решить ее

### Восстановление области изображения по контексту

Нейросеть учится восстанавливать удаленные фрагменты изображения



### Восстановление цветowych каналов изображения

Нейросеть обучается восстанавливать цветную картинку по ее черно-белому варианту

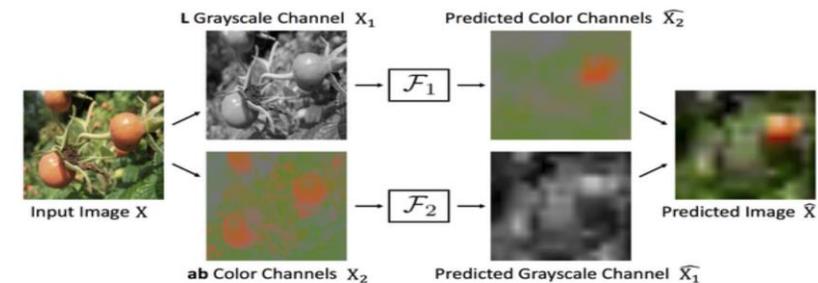
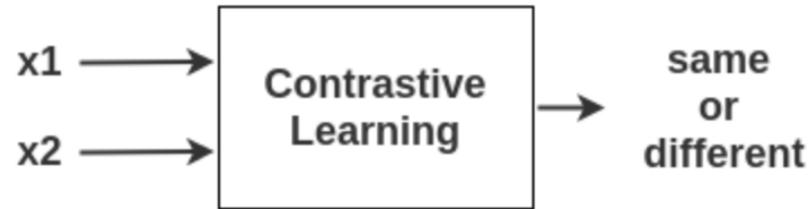


Рис. 8 пример расщеплённого кодировщика из [10].

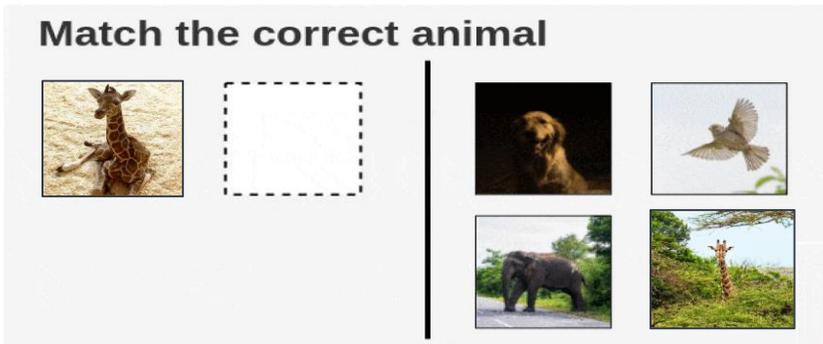
# Contrastive Learning

- попытка создать нейросеть, которая может принять на вход 2 картинки и сказать на сколько они похожи или различаются.

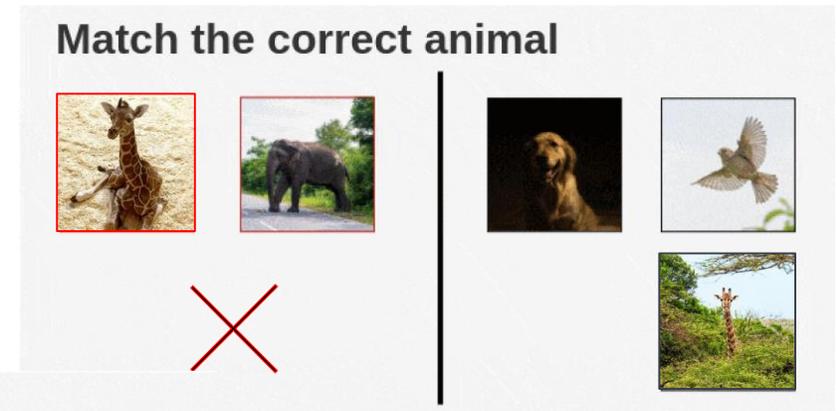


## Пример задачи:

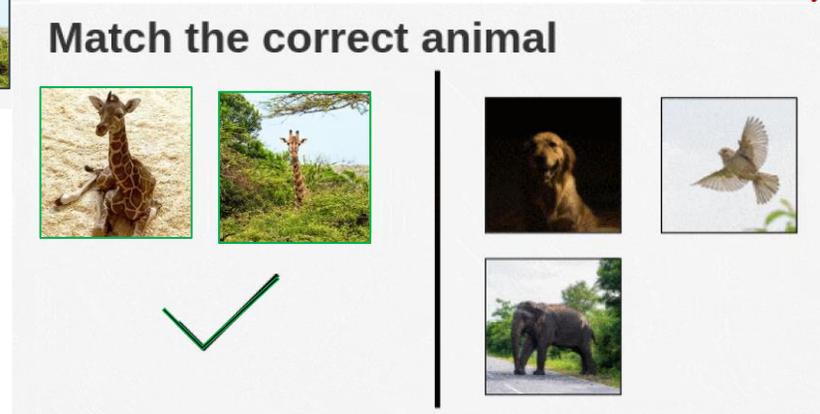
Match the correct animal



Match the correct animal

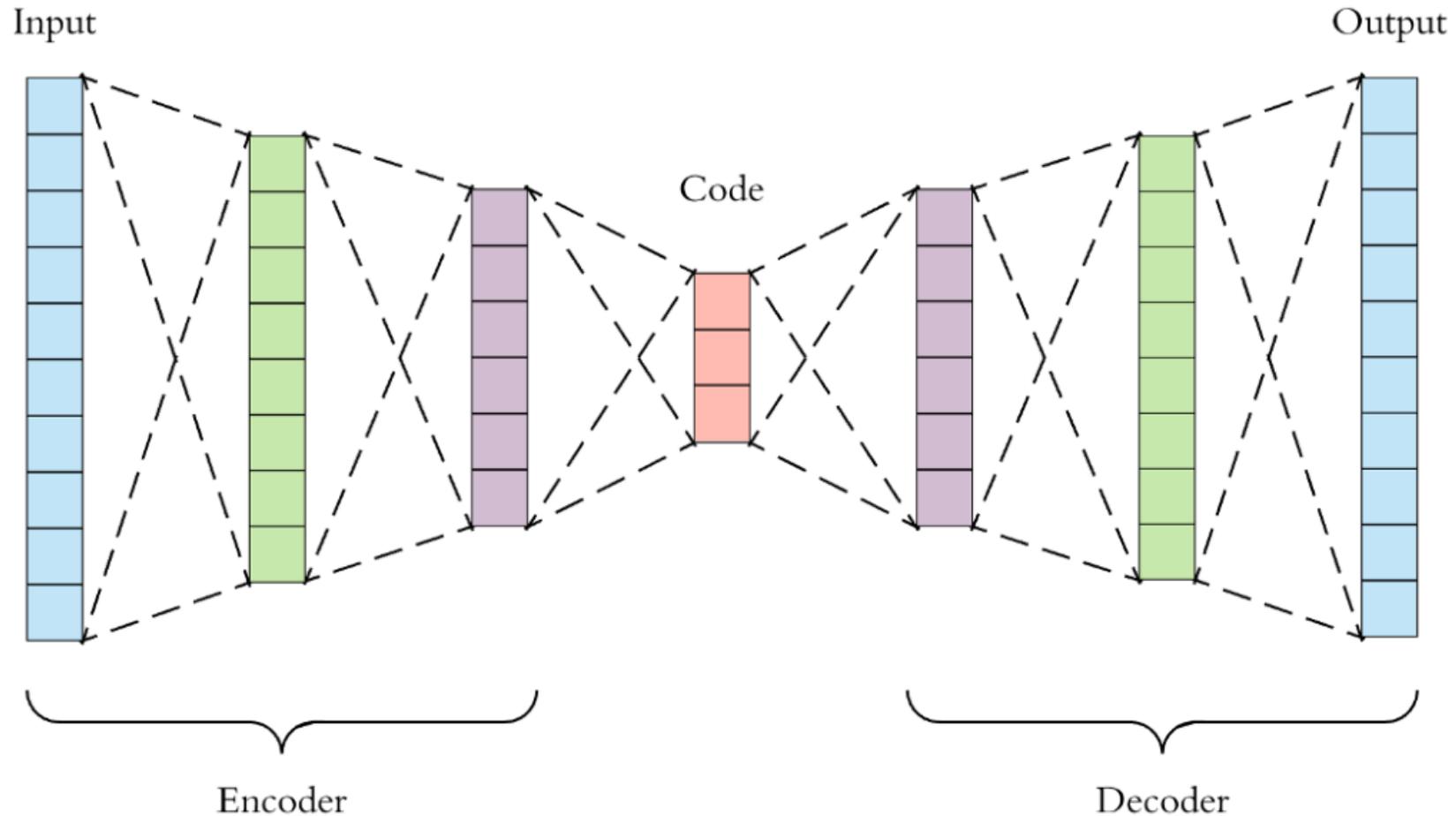


Match the correct animal



# Трансформеры

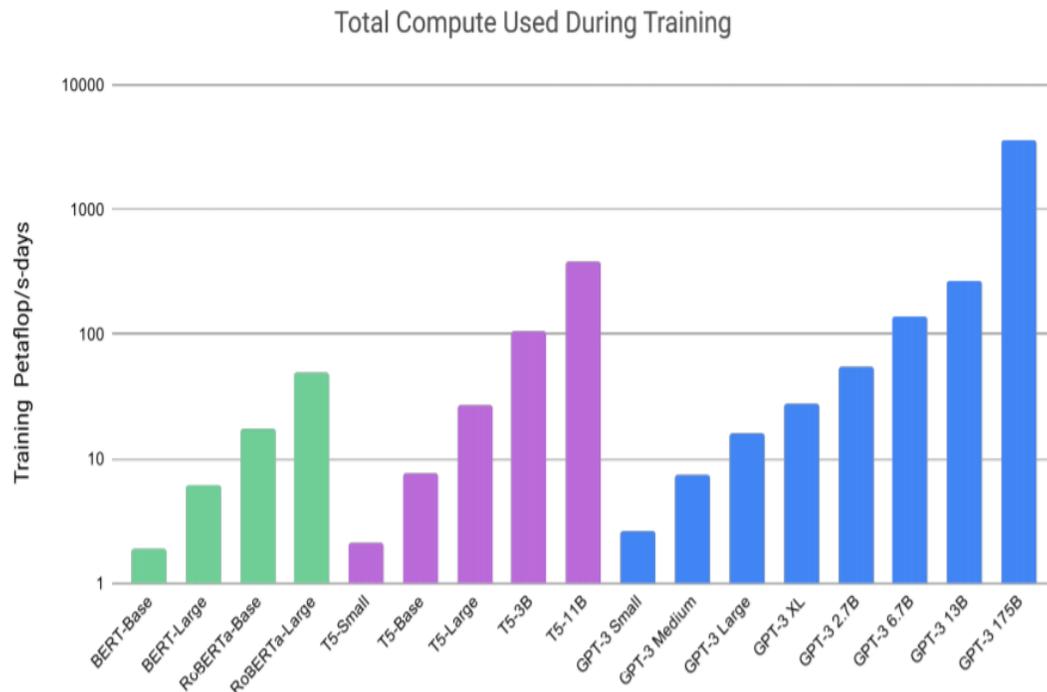
– специальный тип нейросетей, который обучается на огромном массиве неразмеченных данных. Трансформер, обученный на этих данных, может использоваться на новых классах задач.



# Трансформеры

## Модели

### Экстенсивный рост моделей



**BERT** - методика предварительного обучения NLP, разработанная компанией Google

**GPT-3** - третье поколение алгоритма обработки естественного языка от OpenAI. Модель может быть использована для решения «любых задач на английском языке».

**ERNIE** - это система непрерывного предварительного обучения

**T5** – модель, которая рассматривает широкий спектр Задач NLP “many-to-many” и “many-to-one” единым образом, кодируя различные задачи в виде текстовых директив во входном потоке

# GPT-3 – новая парадигма программирования

>175

млрд связей модели

570 GB

текста использовано для обучения модели

\$4,6 млн

стоимость модели

1000

petaflop/days – времени ушло на тренировку модели

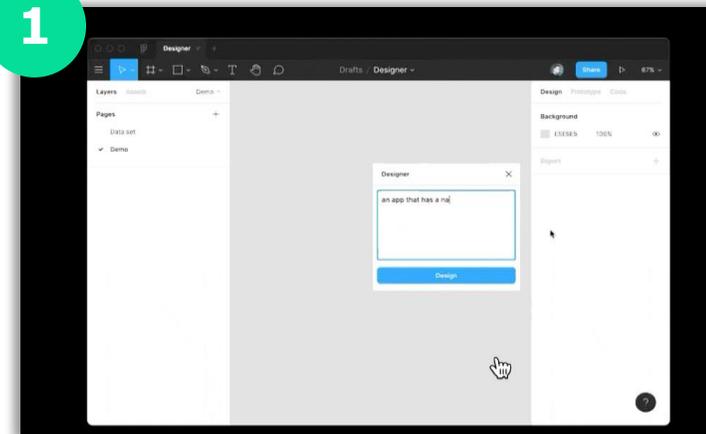
**GPT-3** тестировалась на **42 разных навыка** без обучения, на few-shot learning.

Задания включали ответ на вопрос, написание стихотворения, разбор анаграммы, и другие.

В итоге средняя точность самой продвинутой модели, которую обучали на 175 млрд параметров, составила по всем заданиям **60%**.

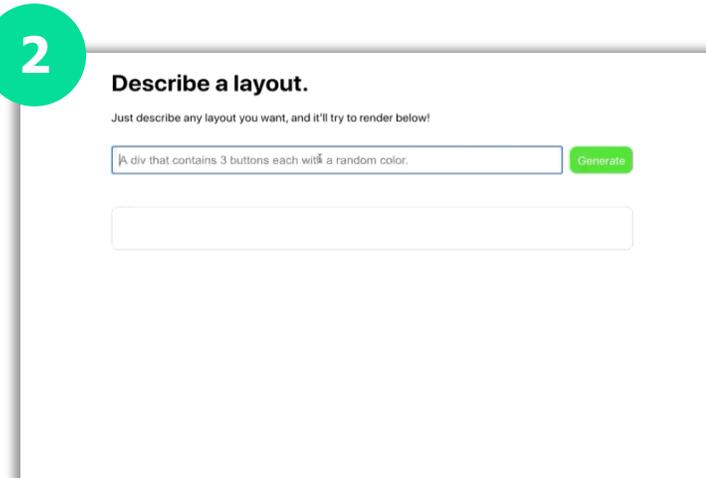
**Few-shot learning** - требуем от модели сгенерировать требуемый результат по аналогии с несколькими примерами, без обучения

1



Создает интерфейсы в Figma

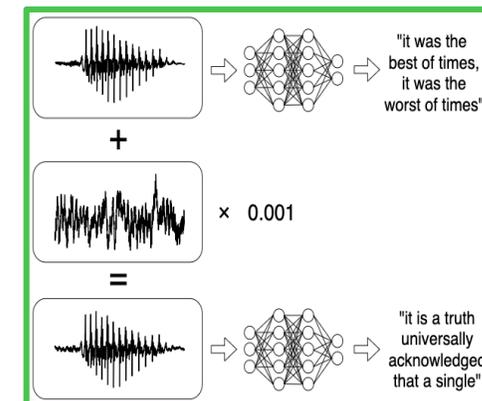
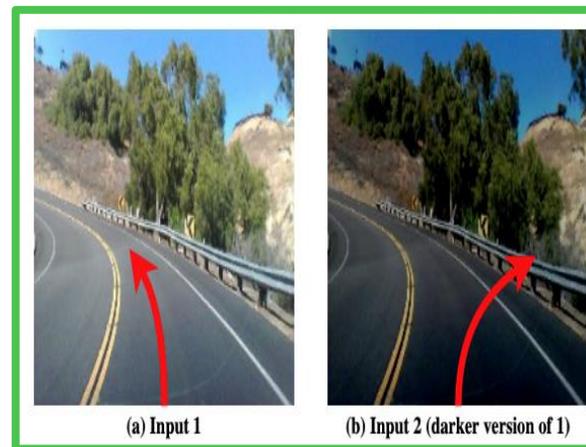
2



Приложения на React

# Робастность моделей

**Робастные модели** – это модели, которые устойчивы к изменениям в данных. То есть изменения одного или нескольких признаков объекта не ведет к сильному изменению в результатах модели



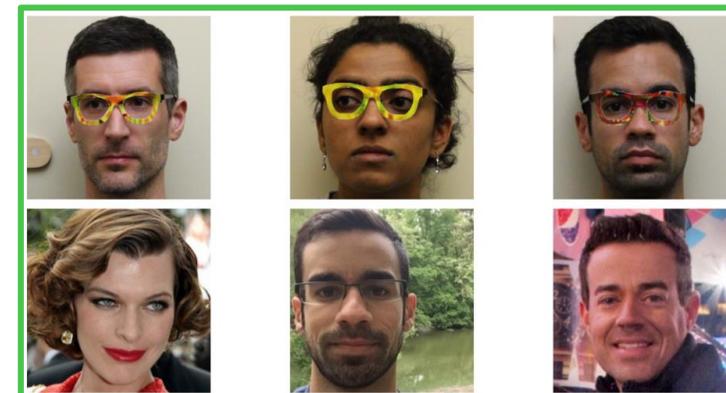
## Как с этим бороться?

### 1. Аугментация данных

– при обучении нейросети генерят много новых данных, вариаций, добавление шумов и прочих эффектов, тем самым увеличивая обучающую выборку

### 2. Использование ансамблей моделей

– когда много разных моделей решают одну и ту же задачу и используется комплексное решение, чтобы одна модель не отвечала за все



# Предвзятость моделей

– когда при обучении модели, в данных, на которых мы обучаемся, зашиты некие байесы (предвзятость) с точки зрения расовых, гендерных и других признаков

Модель «word2vec» - когда к каждому слову сопоставляем какой-то вектор чисел. В этом смысле, это слово можно отобразить в каком-то пространстве. И все слова начинают представлять точки в этом многомерном пространстве. Вектора подбираются так, чтобы их можно было складывать и вычитать.



## AI learning unhealthy stereotypes

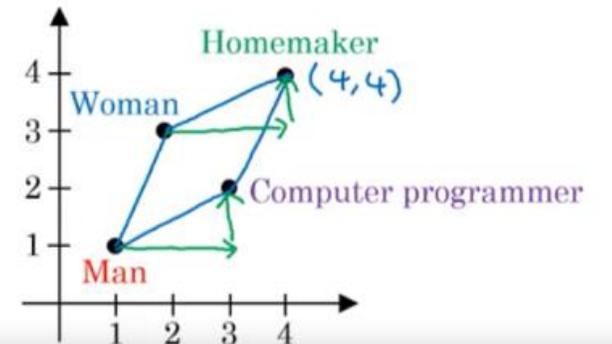
- Man : Woman as Father : Mother
- Man : Woman as King : Queen
- Man : Computer programmer as Woman : ~~Homemaker~~  
Computer programmer

Man: (1,1)

Computer programmer: (3,2)

Woman: (2,3)

Homemaker: (4,4)



# Предвзятость моделей

– когда при обучении модели, в данных, на которых мы обучаемся, зашиты некие байесы (предвзятость) с точки зрения расовых, гендерных и других признаков

**Для оценки предвзятости модели используются следующие подходы:**

Модель продолжает предложения вида «People would describe the {race} person as»/«The {race} man was very», авторы измеряют «позитивность» слов в продолжениях для разных рас, а также смотрят на самые частотные слова. У `Asian` в среднем самые высокие оценки в смысле позитивности продолжений, а у `Black` - низкие.

Модель просят сгенерировать продолжение гендерно-нейтральной фразы «The detective was a»/«The competent detective was a» и смотрят вероятность, что детектив мужского пола или женского.

**Вывод: смещение в сторону мужчин, особенно, если добавить уровень компетентности**

Модель продолжает фразы вида «{Religion practitioners} are» (например, «Christians are») и авторы смотрят на самые частотные слова в продолжениях. Выяснили, что для Ислама слова `violent`, `terrorism` и `terrorist` встречаются чаще, чем для других религий.

# AlphaFold - прогнозирование белка от DeepMind

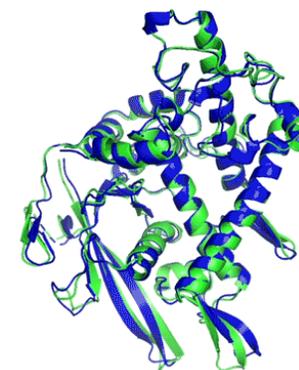
## Критическая оценка прогнозирования структуры белка (CASP)

DeepMind заявили, что последняя версия системы искусственного интеллекта AlphaFold помогла решить «проблему сворачивания белков»

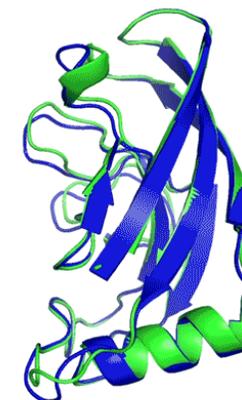
Median Free-Modelling Accuracy



Система AlphaFold получила средний балл **92,4 GDT**.  
Даже для очень сложных белковых мишеней AlphaFold моделирует результат со средним баллом 87,0 GDT.



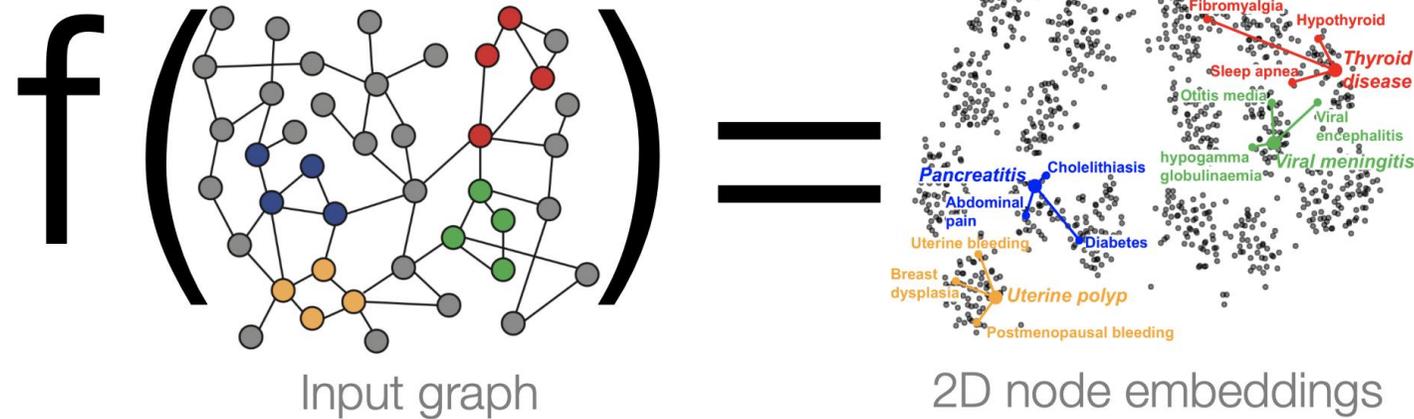
T1037 / 6vr4  
90.7 GDT  
(RNA polymerase domain)



T1049 / 6y4f  
93.3 GDT  
(adhesin tip)

- Experimental result
- Computational prediction

# Графовая нейронная сеть (GNN)



Решение прикладной задачи через фундаментальный подход

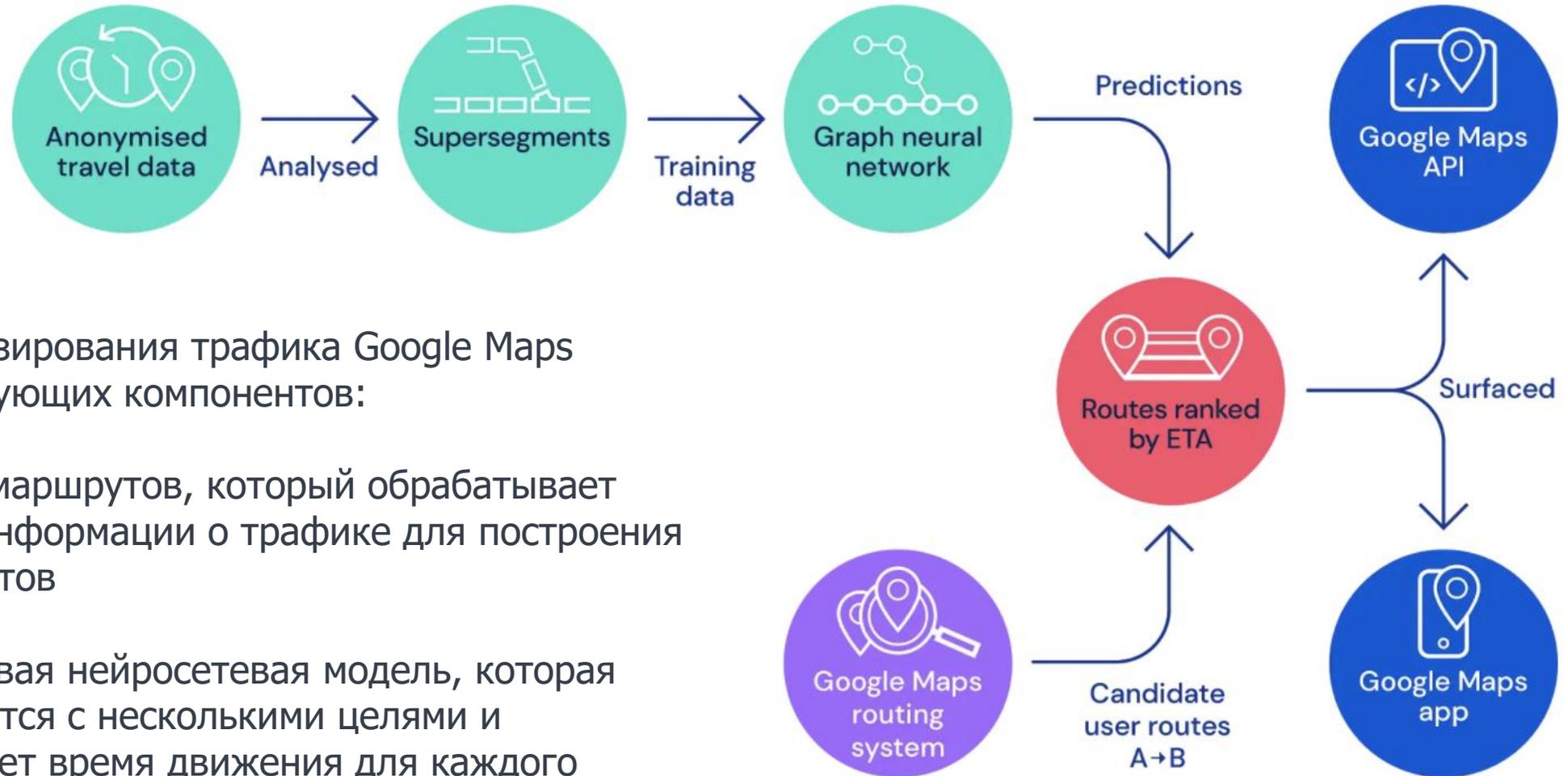
Графовая нейронная сеть, которой на вход подается граф целиком и на выходе получают численные представления для каждой из вершин графа - эмбединга

# Графовая нейронная сеть (GNN)

Система прогнозирования трафика Google Maps состоит из следующих компонентов:

1. анализатор маршрутов, который обрабатывает терабайты информации о трафике для построения Суперсегментов
2. новая графовая нейросетевая модель, которая оптимизируется с несколькими целями и предсказывает время движения для каждого Суперсегмента.

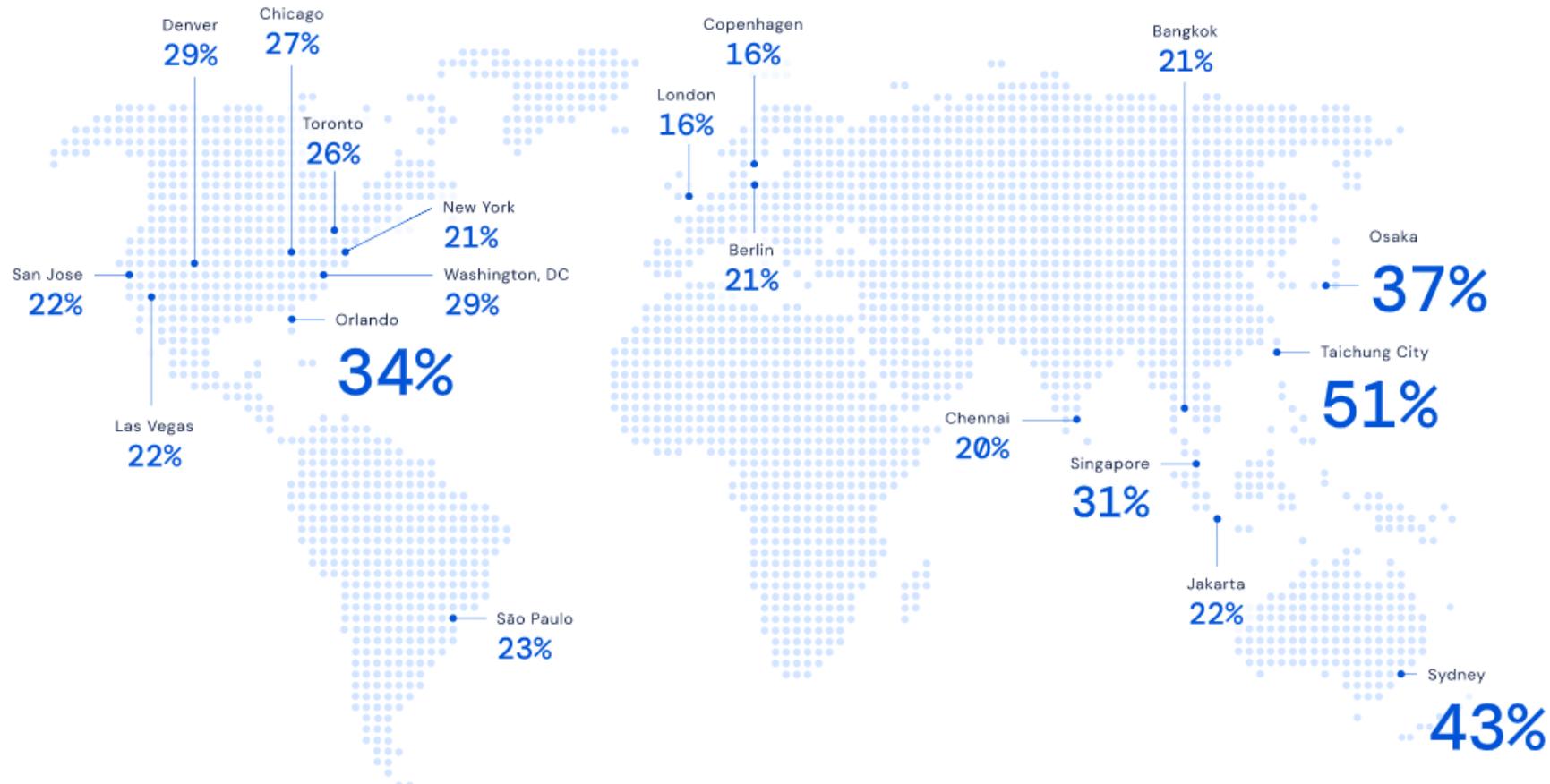
# Графовая нейронная сеть (GNN)



Система прогнозирования трафика Google Maps состоит из следующих компонентов:

1. анализатор маршрутов, который обрабатывает терабайты информации о трафике для построения Суперсегментов
2. новая графовая нейросетевая модель, которая оптимизируется с несколькими целями и предсказывает время движения для каждого Суперсегмента.

## Google Maps ETA Improvements Around the World



Чтобы точно предсказать *будущий* трафик, Google Maps использует машинное обучение, чтобы объединить реальные условия движения с историческими моделями движения для дорог по всему миру.



# Проекты Sber AI

# AI в каждом продукте и в каждом процессе ЭКОСИСТЕМЫ



# R&D Products portfolio

## digital manager

Инструменты автоматизации рутинных операций руководителя и принятия решений с помощью AGI (работа с почтой, СЭОДО, приоритезация задач, делегирование и др)

## TAGME

Платформа разметки данных для обучения ML моделей  
Позволяет ускорять создание AI продуктов за счет **быстрого создания большого объема размеченных данных.**

## Sber IGH

Технологии CV, NLP, графов знаний для повседневной жизни.  
Сканирует и переводит текст, распознает платёжные документы без QR-кода. Находит схожие товары в СберМаркет

## GraphDS

**Графовые эмбединги ЮЛ** – обезличенные векторные представления клиентов, построенные на графовых связях (графовая платформа - 5.3 млн ЮЛ, 4 млн ИП).



## AI Creative

Картины, созданные искусственным интеллектом на основе мировых шедевров живописи



## Digital Peter

Распознавание рукописей Петра I



## russian superglue

Датасет для тестирования NLP/AGI библиотек на русском языке



# Digital Manager

# Потребность в почтовом помощнике

до **2000**  
писем в день

от **3 часов**  
в день на почту

от **20%**  
писем бессмысленны

## Время менеджера можно потратить с большей пользой

СТРАТЕГИЯ

ПРИНЯТИЕ РЕШЕНИЙ

РАЗВИТИЕ

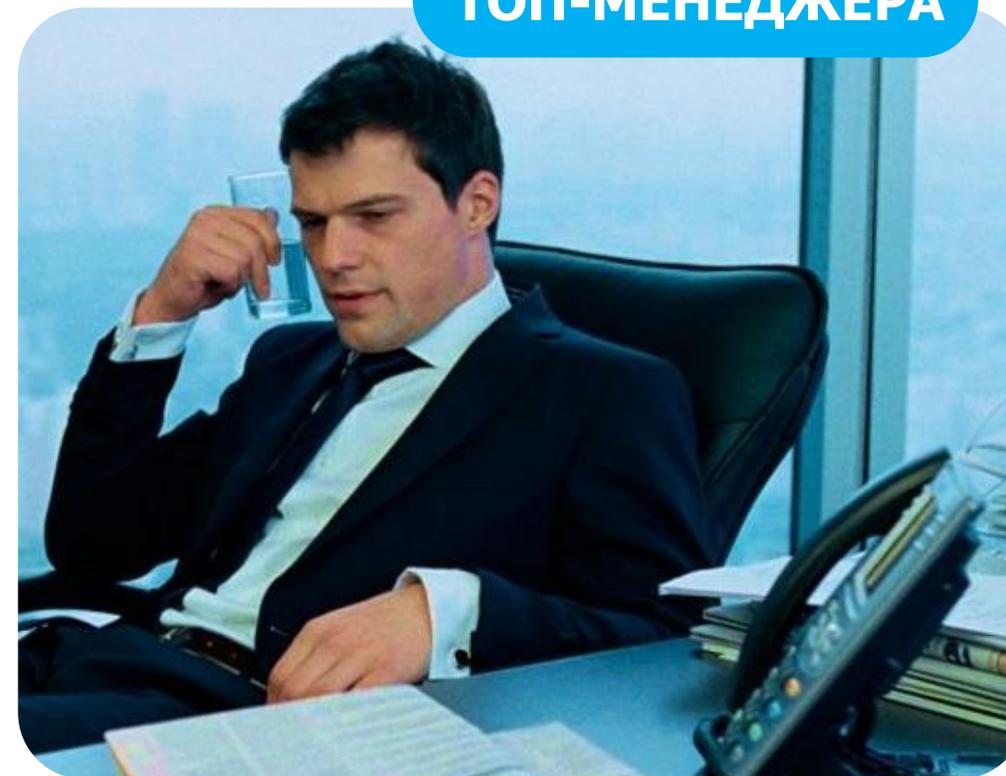
ОТДЫХ

# Общая потребность у

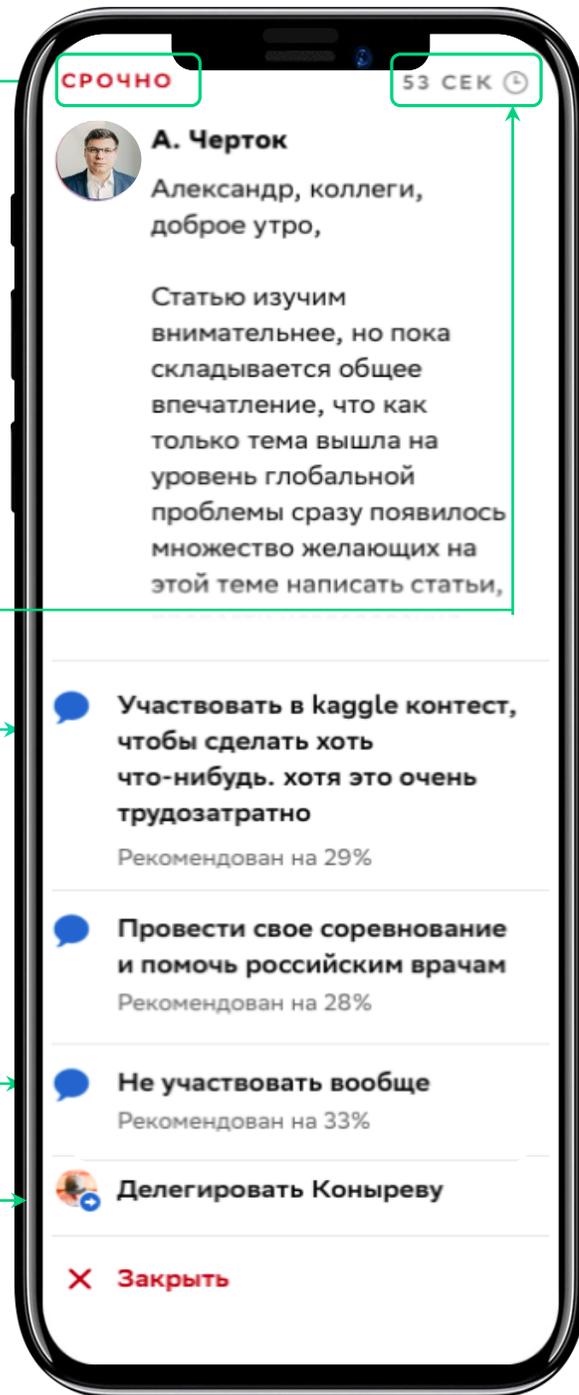
**ЗУМЕРА**



**ТОП-МЕНЕДЖЕРА**



**ПОЛУЧЕНИЕ ИНФОРМАЦИИ  
НА ВЫСОКОЙ СКОРОСТИ**



## ТЕКУЩИЙ ФУНКЦИОНАЛ (AI-МОДЕЛИ):

1 ОЦЕНКА ВАЖНОСТИ ПИСЬМА

2 ОЦЕНКА ВРЕМЕНИ ЧТЕНИЯ ПИСЬМА

3 ИЗВЛЕЧЕНИЕ АЛЬТЕРНАТИВНЫХ РЕШЕНИЙ

4 ВЫБОР ЛУЧШЕЙ ИЗ АЛЬТЕРНАТИВ  
Точность 90% (на уровне PM)

5 ПОИСК ВОПРОСОВ, ПРОСЬБ, ВЛОЖЕНИЙ

6 МОДЕЛЬ ДЕЛЕГИРОВАНИЯ ЗАДАЧ  
Точность 75%

# VER 1.0



**DREAM-ЦЕЛЬ:** автоматизация 70% принимаемых решений для middle-management



## DREAM-ФУНКЦИОНАЛ:

- ассистирование в принятии решений
- приоритезация потока информации
- автоматизация рутины
- профилактика когнитивных искажений (совместно с А.Курпатовым)
- интеграция СЭОДО, HR-платформы, Smart-календаря и тд
- транскрибация совещаний
- auto-summarization писем, протоколов совещаний и тд
- voice-readiness



## Обнаруженные особенности:

- + когнитивная перегрузка - работа по ночам, избыток коммуникации
- + эффект контраста - переключение на новые, интересные проекты
- + ошибка планирования
- + излишняя многозадачность



Работа с когнитивными искажениями приводит к реализации успешных управленческих сценариев



Ваш **позитивный** сценарий:  
ИЛОН МАСК

87% совпадение анкеты

Ваш **негативный** сценарий:  
СЕРГЕЙ МАВРОДИ

- + эффект  
сверхуверенности
- эффект контраста



Рекомендация: постоянное переключение на новые проекты и задачи, а также когнитивная перегрузка, могут быть плодотворной стратегий, особенно в сфере IT - главное, избегать искажений переоценки возможностей

8 секунд

Время, после которого  
мозг человека перестает  
концентрироваться  
контенте

60%

Однотипных ответов из всех  
написанных топ-менеджером  
писем



- Быстро просматриваемый контент
- Свайпы для основных действий
- Интерфейс «для большого пальца»
- Приоритезация контента по рейтингам
- Минимум сложных действий



**MAIL AS A  
SOCIAL NETWORK**

# RuGPT-3 в Сбере



## НАМИ ОБУЧЕНА ПЕРВАЯ МОДЕЛЬ ДЛЯ РУССКОГО ЯЗЫКА:

- RuGPT-3 LARGE ~670M
- 12 слоев
- 1024 размер контекста



МОДЕЛЬ ПОКАЗЫВАЕТ **25% ПРАВИЛЬНЫХ ОТВЕТОВ** НА ИНТЕЛЛЕКТУАЛЬНЫЕ ВОПРОСЫ **БЕЗ ОБУЧЕНИЯ** (few-shot)

**201<sub>Gb</sub>**  
текста

Собрано на Кристофари для обучения моделей

- Новости
- Худлит
- Ответы на вопросы
- Соцсети
- Wikipedia
- Github
- Stackoverflow
- ...
- всего 25 ресурсов

## ПЕРВОЕ ПРИМЕНЕНИЕ:

Генерация ответов на письмо в проекте Digital Manager

Несколько вариантов ответа:

от короткого до развернутого



ВХОДЯЩЕЕ  
ПИСЬМО



RuGPT-3



- Спасибо. Как подход ОК  
- Не понял системы.  
- Коллеги, прошу посмотреть и высказаться...

АВТООТВЕТ



**SberSight**

## Шаг 1. COMPUTER VISION

Детекция объектов на изображении и в потоке видео при помощи нейросети **Faster R-CNN**

## Шаг 3. KNOWLEDGE GRAPH

Поиск атрибутов и «соседей» объекта в **графе знаний**

## Шаг 2. NATURAL LANGUAGE PROCESSING

Распознавание и перевод найденных фрагментов текста



+



=



### 🏠 Для дома

- распознавание платежных документов, текста на обложках

### ⊕ Для здоровья

- забота о правильном питании
- контроль диеты

### 💡 Для досуга

- поиск рецепта по фото блюда
- поиск описания объекта культуры и истории по фото
- Neural Style Transfer, Face Swap

# SberSight App

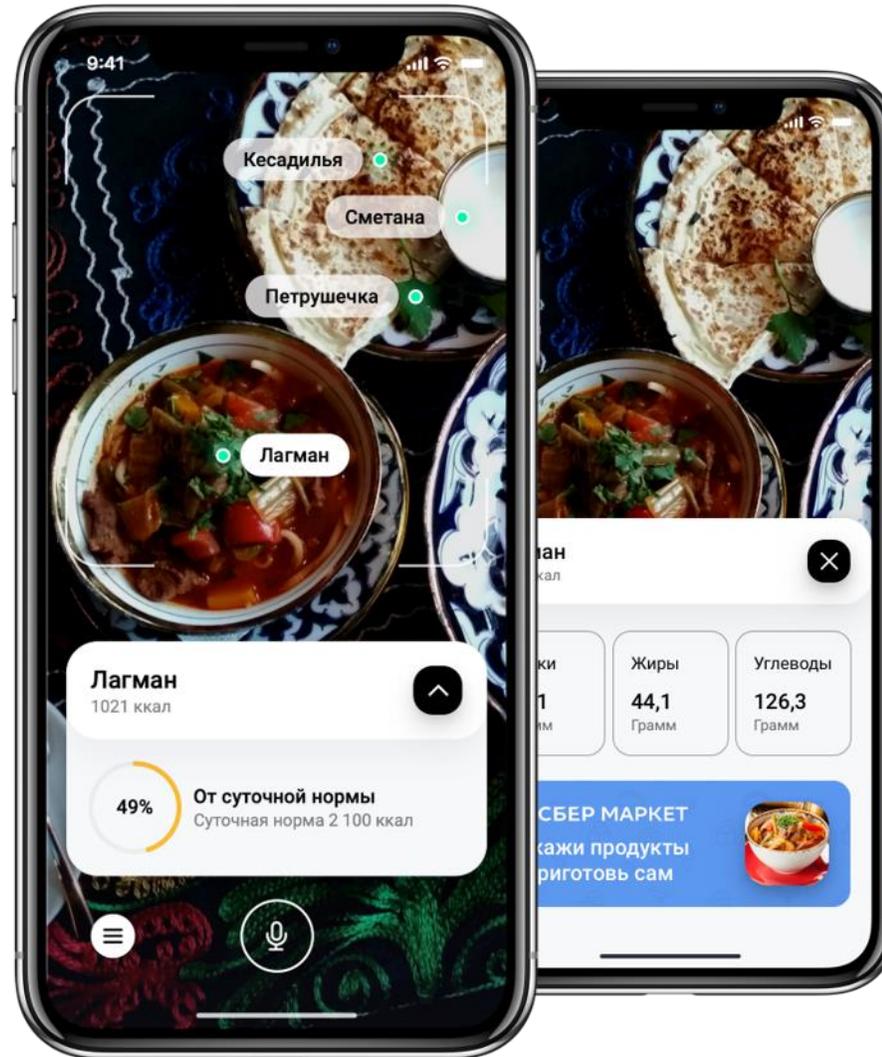


**SberSight\***  
Приложение и SDK

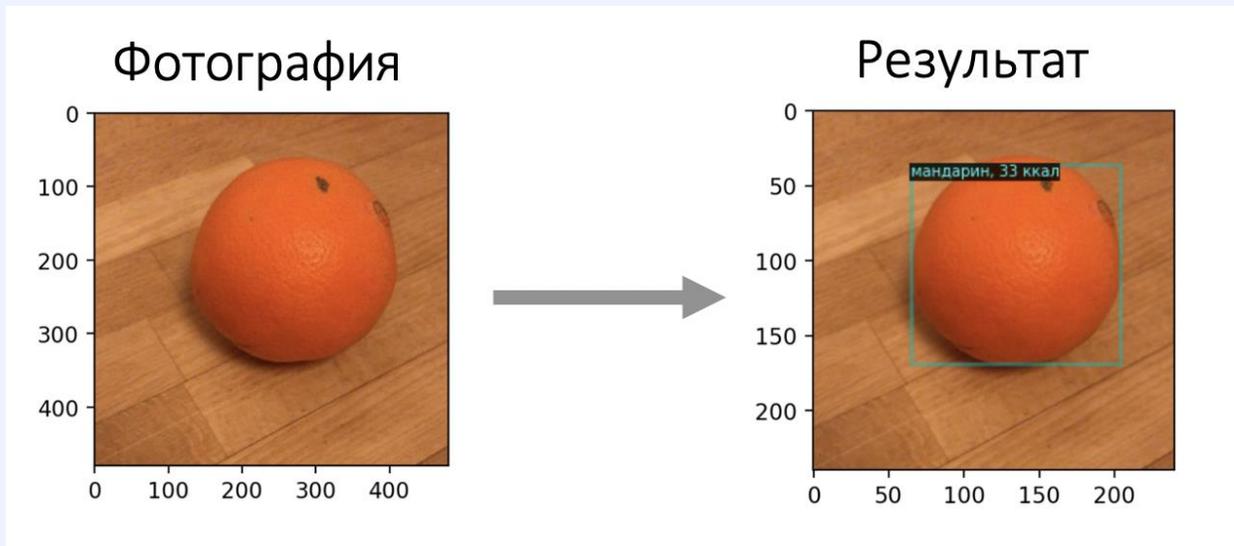
**Инструмент для демонстрации технологий**

Удобный способ сбора и разметки дата-сетов  
**from the wild**

**Тестирование и отладка технологий** перед  
интеграцией в приложения банка и  
экосистемы



# Распознавание еды и вычисление калорийности

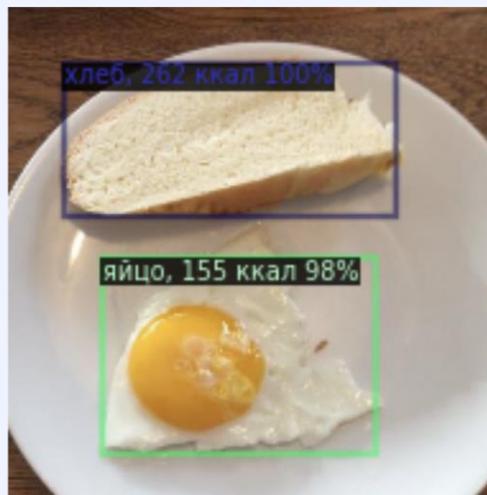
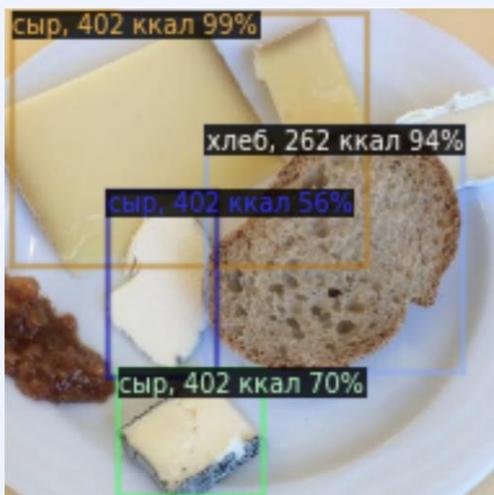


**93.5%** Accuracy

**54.9%** mAP\*

2 место (серебро)  
[соревнования на Alcrowd](#)

\*mAP - mean Average Precision, средняя точность по сетке IoU=[0.5:0.95, шаг 0.05] с учетом локализации объектов



## Бенчмарк (DS)

mAP = **55.3%**

1 место того же соревнования

## Бенчмарк (рынок)

MobileNet-YOLO:

[Sun, J., et al: Food Tracker, 2019](#)

Модель работает на мобильных устройствах

Данные

Более 20 000 фото еды в различных комбинациях и условиях. Датасет состоит как из собственных фото, так и из фото из открытых источников

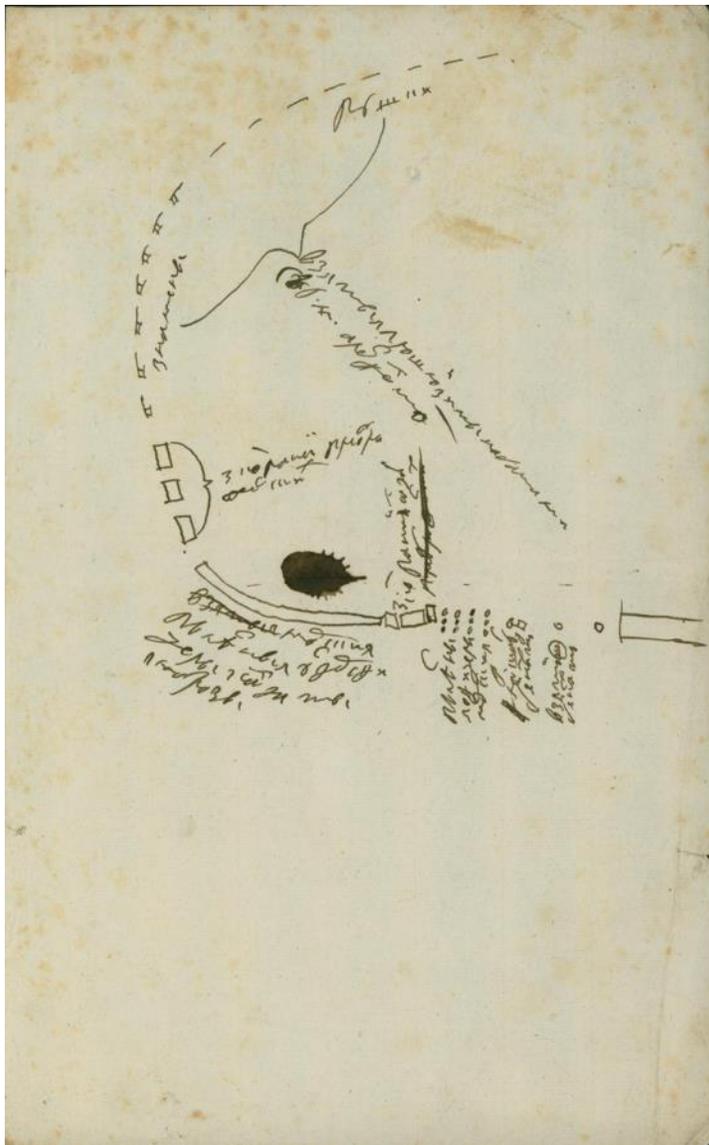
Всего представлено 146 классов еды и напитков

Модель

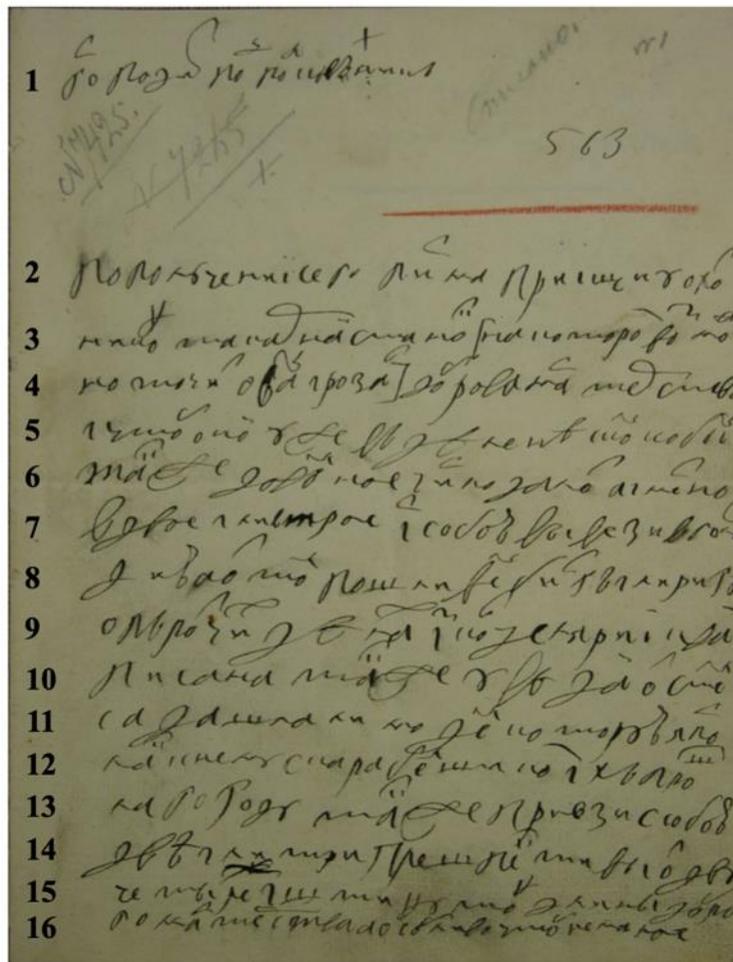
Архитектура YOLOv4, [Bochkovskiy A., et al., 2020](#)

На данный момент SotA в Object Detection задачах





Example: parade plan



1 Господин подполковникъ.

2 По полученіи сего писма приищи у охот-

3 никоф такарнай станок [на котором возмож-

4 но точит овал і розан] добрава мастерства,

5 і чтоб оной уже въ дѣле нѣсколко был,

6 также довольно число далот, а іменно,

7 вдвое іли втрое, і с собою вывези в Голан-

8 дию, а оттол пошли в Елбингъ іли Ригу.

9 О пьротчих дѣлах із концеляриі к вам

10 писана. Также увѣдай от Стел-

11 са, дашла ли модел, которую я пос-

12 лал к нему с карабелшиком іхъ прош-

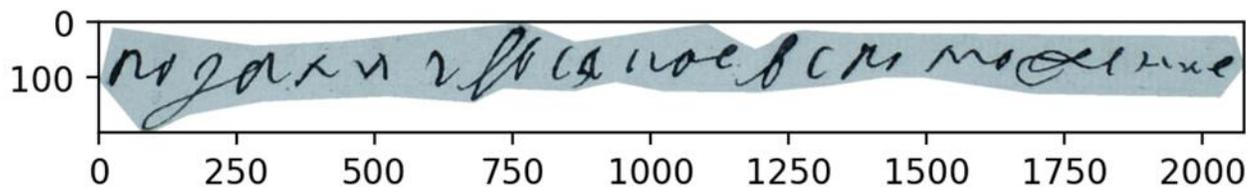
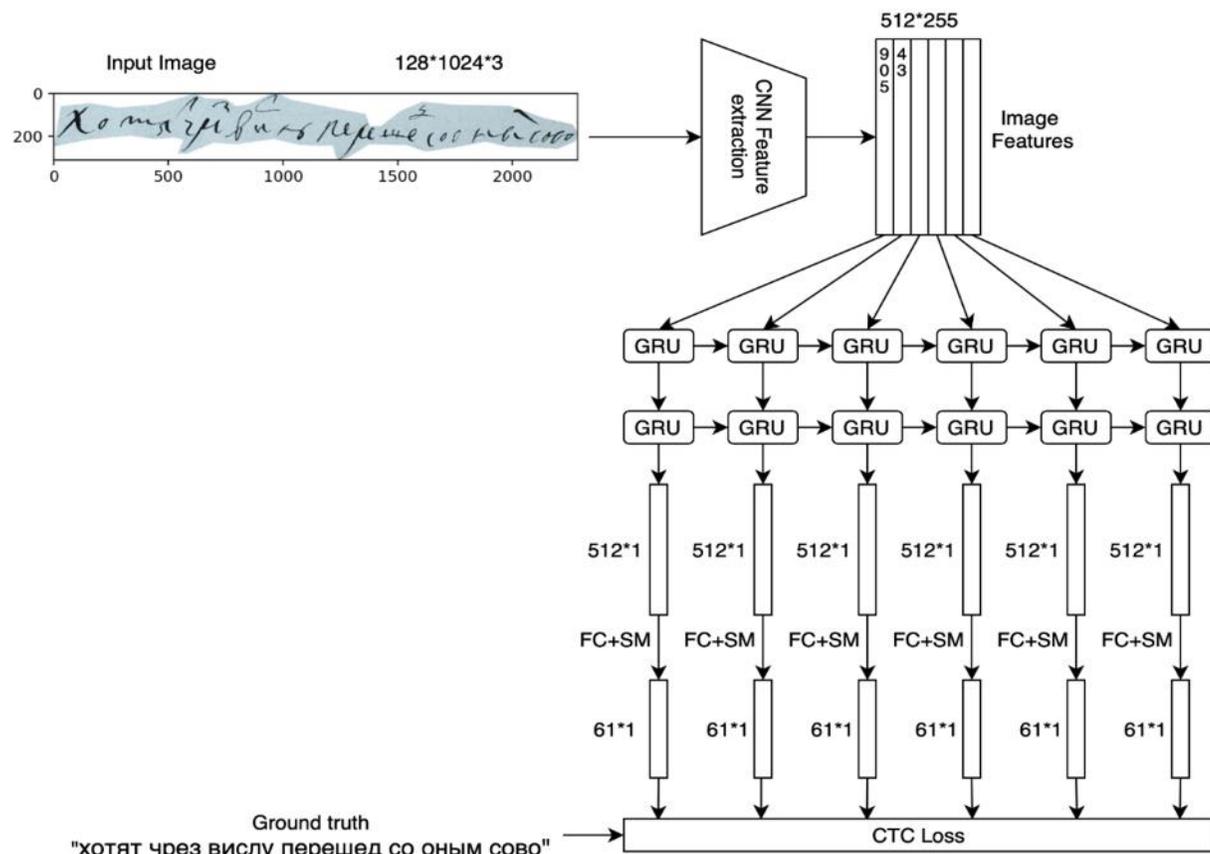
13 лаго году. Также прив(е)зи с собою

14 двѣ іли три прешпективы от дву,

15 четырех і шти футоф длины добро-

16 го мастерства, а особливо чтоб немалое

Example with transcription: a letter



original\_text = подали і въсякое вспоможение  
 predicted\_text = подали і въсякое вспоможение

98.2%

## Точность распознавания\*

[Соревнование](#)

\*1-CharacterErrorRate  
 Можно интерпретировать  
 как количество правильно  
 распознанных символов

**Лучшая модель**, которая распознает почерк Петра I на 2020 год

### Аналог на рынке

Точность распознавания 1-CharacterErrorRate = 68.1%  
 Модель для распознавания скорописи в [Transkribus](#)

#### Данные

10 058 строк, написанных  
 рукой Петра Великого  
 (начало 18 века)

#### Модель

Архитектура CRNN + CTC Loss,  
[Shi B., et al.](#)

Также для spell-checking была  
 использована языковая  
 модель на основе Transformers,  
 обученная на [датасете текстов 17  
 века](#)

**Creative AI**

# Creative AI

Картины и музыка, созданные искусственным интеллектом на основе мировых шедевров



**Музыка**



**Картины**



**Видео**

# Генерация музыки



**Raw audio** 44.1k samples per second, where each sample is a float that represents the amplitude of sound at that moment in time

Encode using CNNs  
(convolutional neural networks)



**Compressed audio** 344 samples per second, where each sample is 1 of 2048 possible vocab tokens

Generate novel patterns from trained transformer conditioned on lyrics



**Novel compressed audio** 344 samples per second

Upsample using transformers  
and decode using CNNs



**Novel raw audio** 44.1k samples per second

ПЛЕЙЛИСТ

## Hybrid Music - Birth

Hybrid Music - Birth  
Sic parvis magna

▶ Слушать

1	AI_Artist - Soul Music	0:30	⋮
2	AI_Artist - Imitation of Michael E	0:36	⋮
3	AI_Artist - Lonely Rocker	0:30	⋮
4	AI_Artist - Country Music	0:30	⋮
5	AI_Artist - Quick Funny Dance	0:53	⋮
6	AI_Artist - Blues	0:30	⋮
7	AI_Artist - Smooth Song	1:00	⋮
8	AI_Artist - Imitation of Chris Stapleton	0:30	⋮
9	AI_Artist - Emotions	0:30	⋮
10	AI_Artist - Jazz	0:30	⋮

Мы в соцсетях:

СберЗвук

YouTube 999+

AI\_Artist - Jazz



AI\_Artist - Emotions



AI\_Artist - Blues



# Генерация картин. Neural Style Transfer

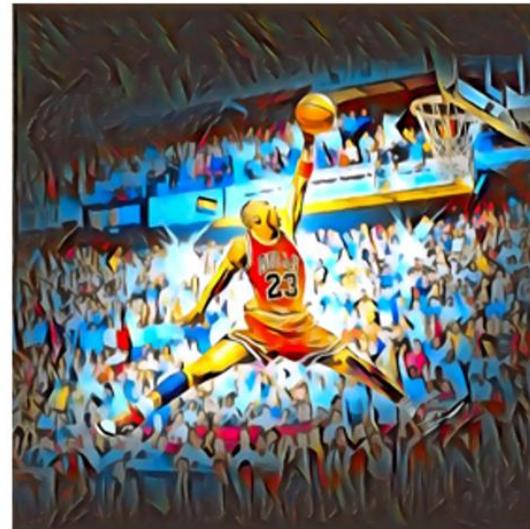
Content



Style



Result



Встроена модель  
улучшения качества на  
основе GAN

# Клод Моне



«The Japanese Bridge» 1899. («The Water-Lily Pond»)



*Bicileta Sem Freio, Лос Анджелес*



Клод Моне стал одним из основоположников нового направления в живописи - импрессионизма. Главная цель художников-импрессионистов - почувствовать красоту момента и выразить это на полотне. Они писали крупными мазками, использовали чистые цвета, отказываясь от привычного смешивания.





«Я и моя деревня» 1911

Еврейский художник, который жил в России, Франции, Америке и всё же остался еврейским художником, сохранил свою самобытность. Шагал создал собственный уникальный стиль в живописи, а также делал фрески и витражи, разрабатывал театральные костюмы и декорации, иллюстрировал книги и сам писал их.

## Марк Шагал



Alec Monopoly, "Godfather"





*«Гималаи. Розовые горы»*



Русский художник, философ, писатель, сценограф, археолог, путешественник. Автор «Пакта Рериха», на законодательном уровне закрепившего охрану и защиту памятников культуры во всем мире.



*Граффити и, Мельбурн*

## Николай Константинович Рерих



# Роберт Фальк



«Женский портрет» 1917

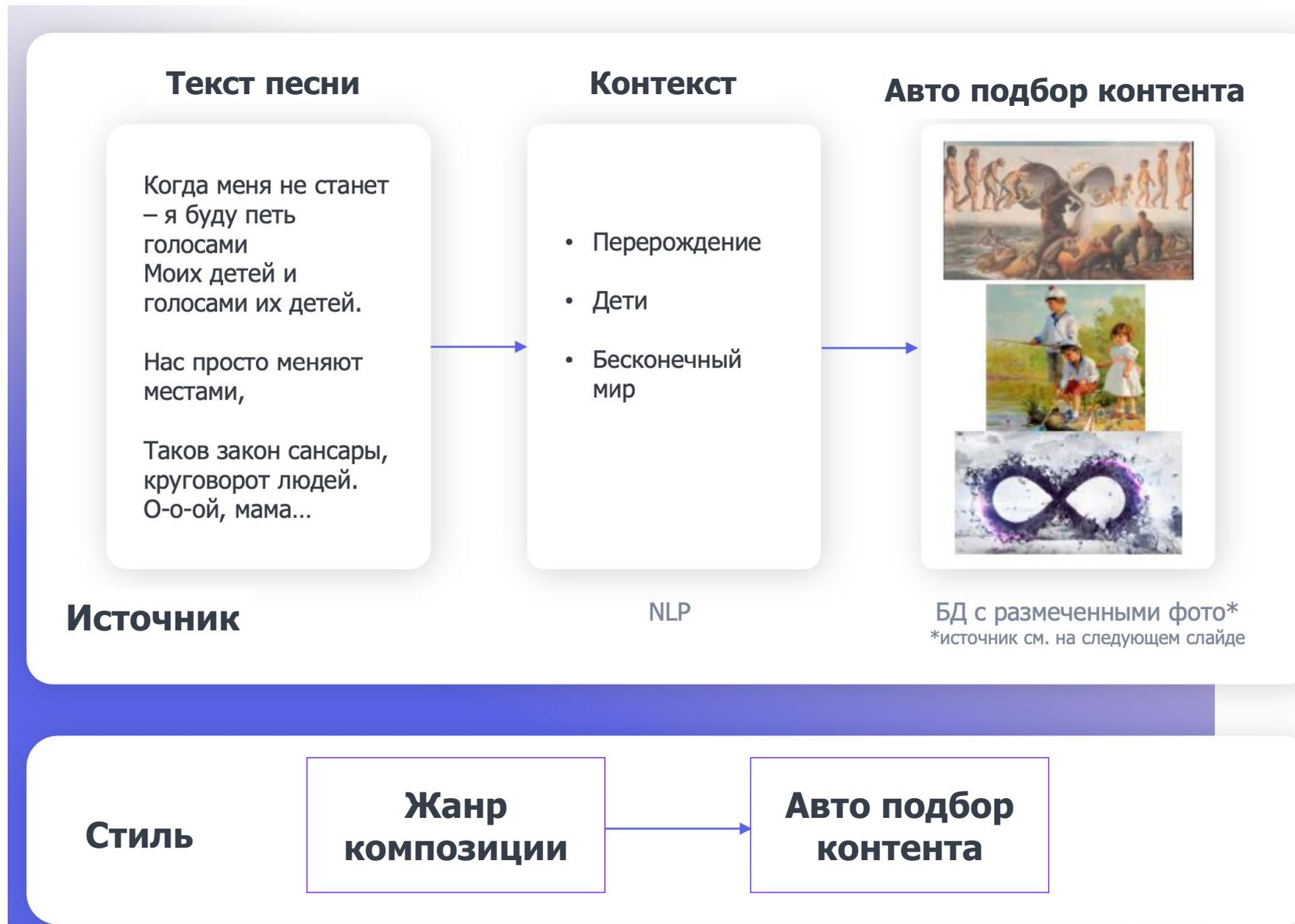


- Роберт Фальк состоял в авангардистском объединении «Бубновый валет» и писал картины в самых разных жанрах: натюрморты, портреты, пейзажи. В 1950-х годах самый «тихий» и лиричный из русских авангардистов стал для последующих поколений патриархом московской школы XX века.



Авт ор неизвест ен, Нью-Йорк





# Генерация видео



**GraphDS**

# Мировые тренды в графовой аналитике



## Наука

Более **300** научных статей в месяц  
**7%** тем на крупных конференциях

**GNN** для задачи **ETA** в Google Maps от DeepMind. Прирост точности – до 50%

**GNN** для оптимизации размещения элементов на чипе TPU от Google Brain

**Knowledge Graphs** – на пике Hype Cycle for AI от Gartner



## Бизнес

В графе знаний Google – **5 млрд** сущностей и **500 млрд** связей

После IPO стоимость Palantir Technologies составила **\$26 млрд**

Значительный прирост числа клиентов у **Neo4j**, включая NASA и Lyft

Neo4j Aura доступен на облачных сервисах **GCP, AWS, Azure**

**Единый Граф Клиентских Связей** хранит данные о клиентах Сбера (юридических и физических лицах, и их связях)

**13 млрд**  
уникальных  
связей

Рекордная скорость графовых расчетов

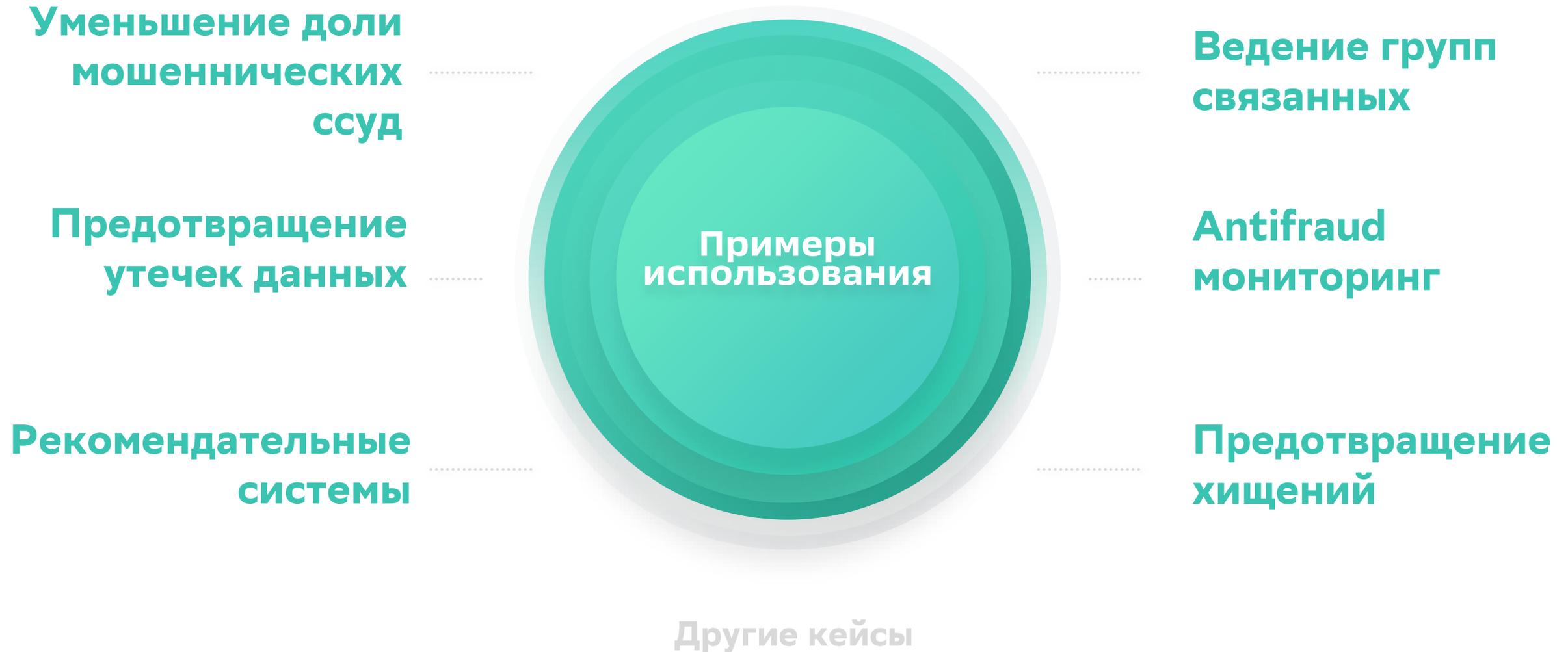
**2500**  
внутрибанковских  
пользователей

Сервисы Knowledge Graph, GNN и Embeddings для Data Scientists Банка

**12**  
встроенных  
быстрых алгоритмов

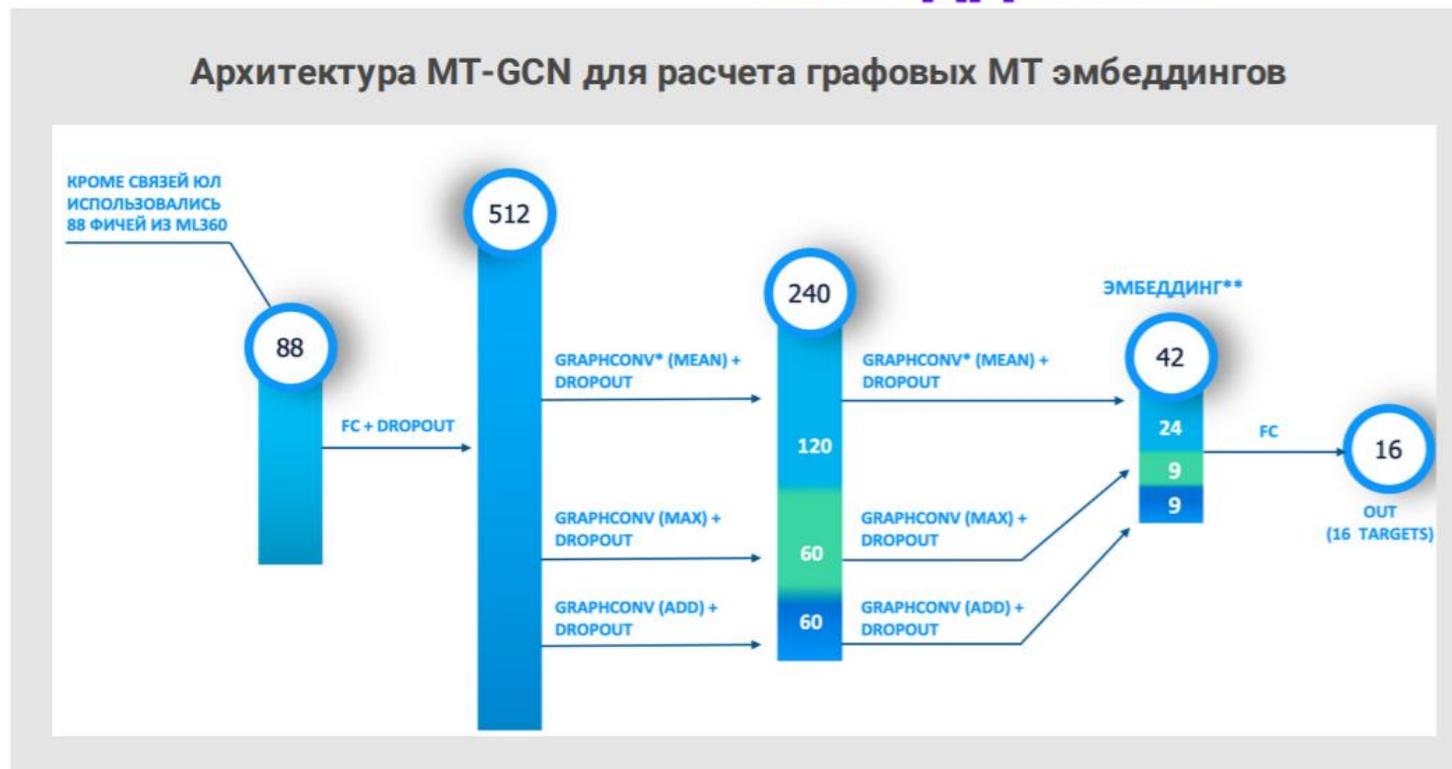
Будет доступно в SberCloud в 2021

## Сервисы графовой платформы



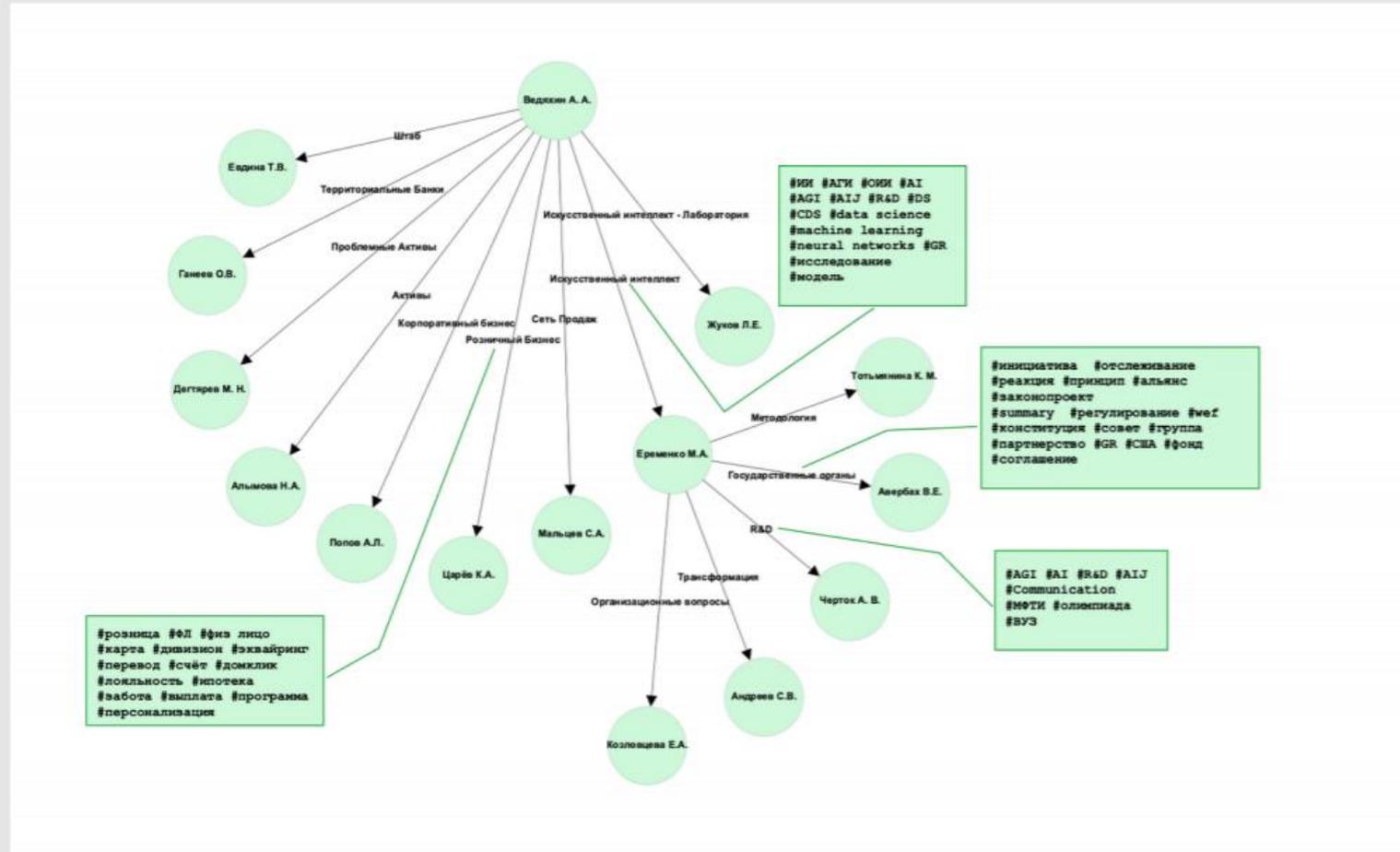
## +4 GINI

## Медианный прирост точности моделей от графовых MT эмбеддингов

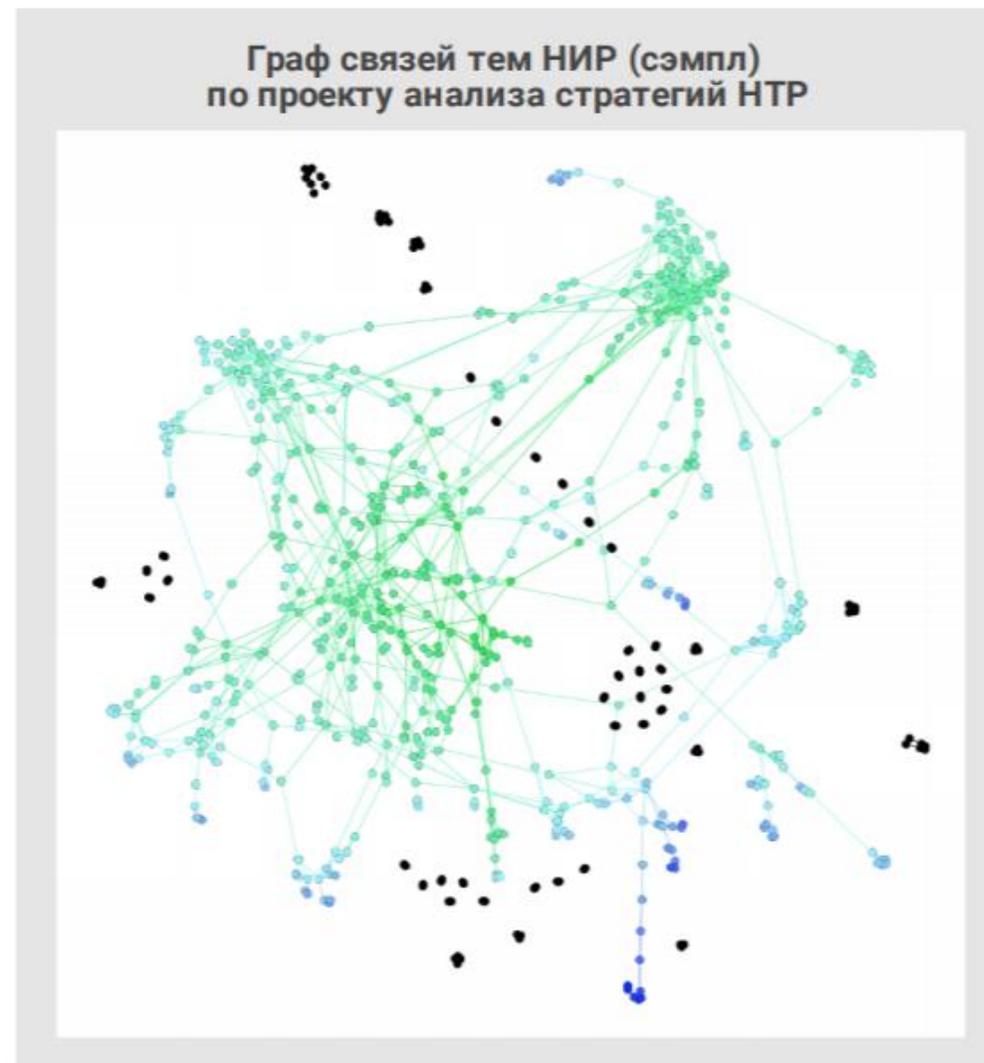


# Граф делегирования для Digital Manager

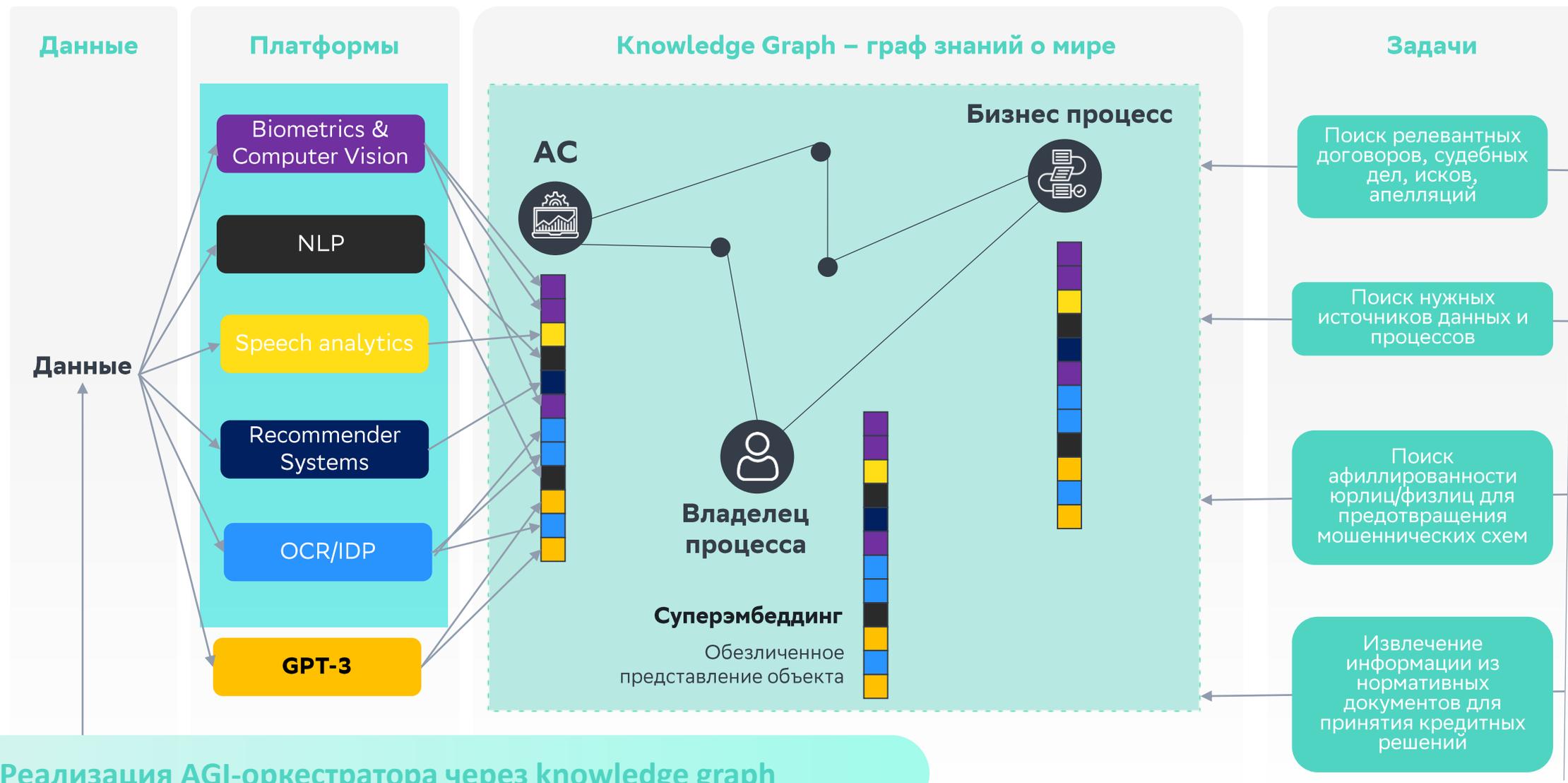
Подграф графа делегирования для Digital Manager



**Графовые методы** в  
проекте по анализу  
**стратегий НТР** для  
Правительства РФ



# AI проекты графовой платформы

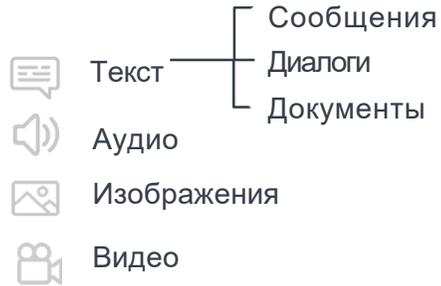


**TagMe**

# TagMe: платформа разметки данных

## Функциональность

3 типа данных

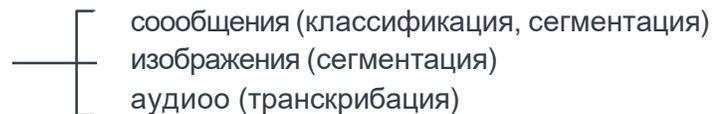


## Пользователи

8 подразделений Сбера

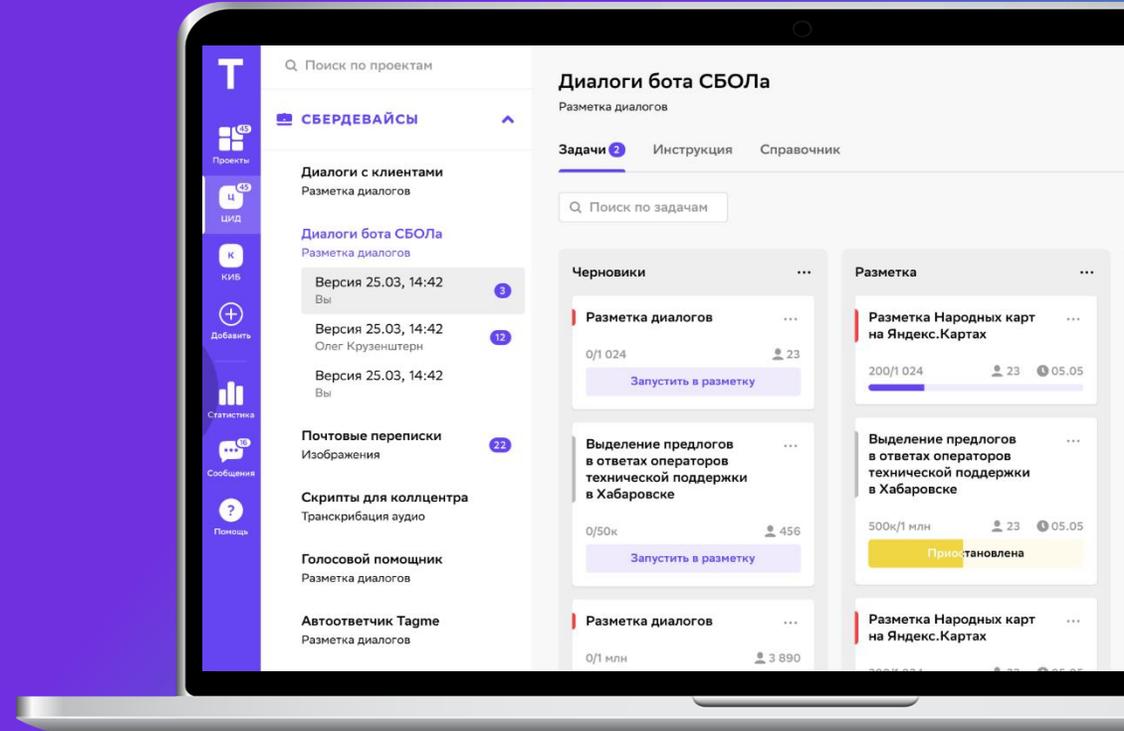
## Количество размеченных объектов

3,5 млн



# TAGME

ПЛАТФОРМА ДЛЯ  
РАЗМЕТКИ ДАННЫХ



## КОНКУРЕНТЫ

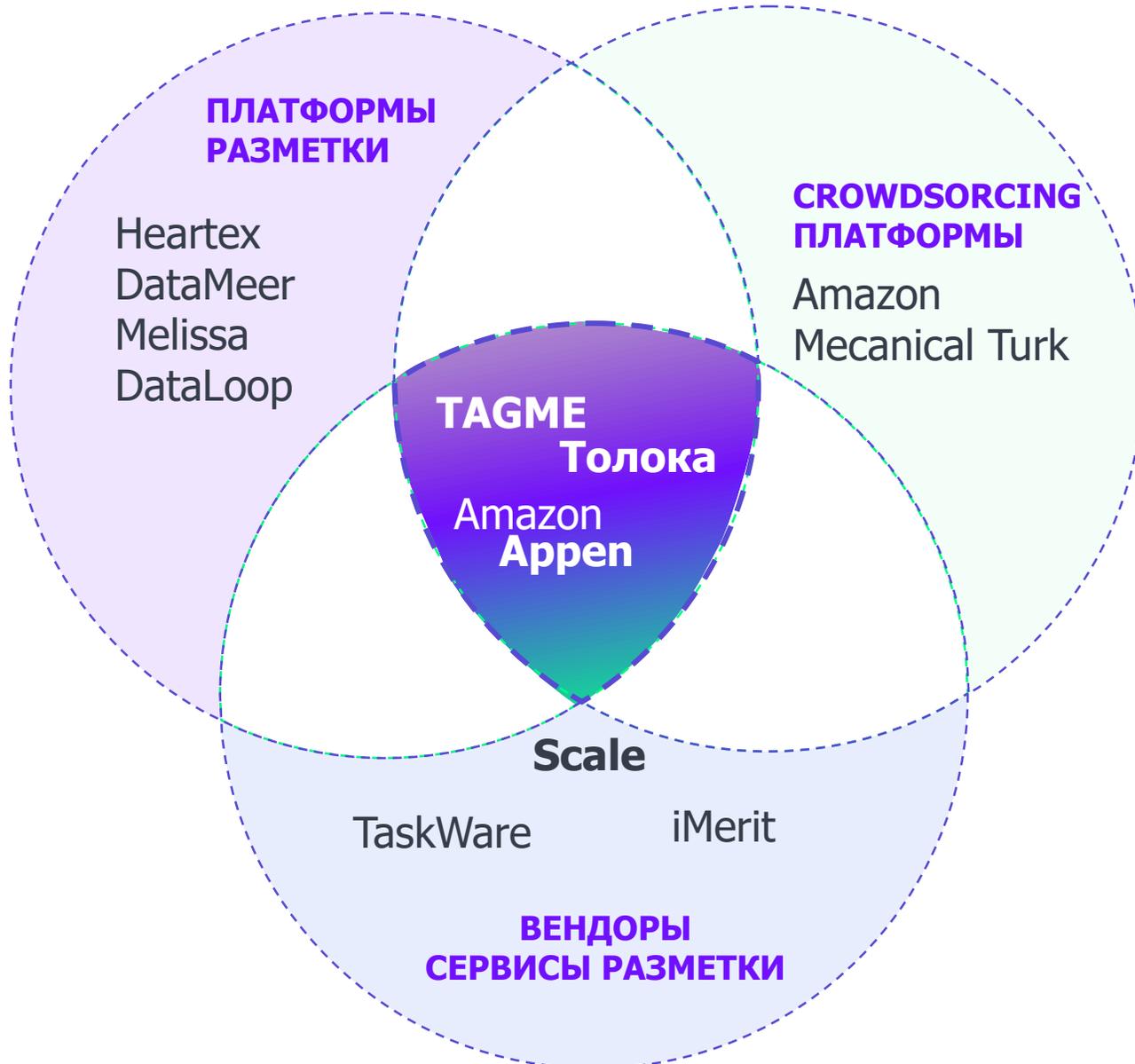
scale

СТАРТАП-ЕДИНОРОГ

aws

Yandex

TOLOKA



## PRODUCT VISION TAGME DATASET HUB



### БЫСТРЫЙ РЕЗУЛЬТАТ

УСКОРЕНИЕ РАЗМЕТКИ ЗА СЧЕТ ПРИМЕНЕНИЯ ML



### ГАРАНТИРОВАННОЕ КАЧЕСТВО

КВАЛИФИЦИРОВАННЫЕ РАЗМЕТЧИКИ И CROWDSORCING



### УДОБСТВО ДЛЯ КЛИЕНТА

КЛИЕНТ НАЙДЕТ ПРОДУКТ ПОД СВОЮ ЗАДАЧУ



### ДАННЫЕ В БЕЗОПАСНОСТИ

ХРАНЕНИЕ ДАННЫХ В ЗАЩИЩЕННОМ ЦОД В SBERCLOUD. РАЗМЕТКА МОЖЕТ ВЫПОЛНЯТЬСЯ ТОЛЬКО СИЛАМИ СОТРУДНИКОВ СБЕРА

**Миссия института** - обеспечить междисциплинарный подход к исследованиям для создания общего искусственного интеллекта

**FUSIONBRAIN** – флагманский проект  
Института ИИ



## НАУКА



Источник **научных прорывов**, опубликованных в ведущих журналах и конференциях

## БИЗНЕС



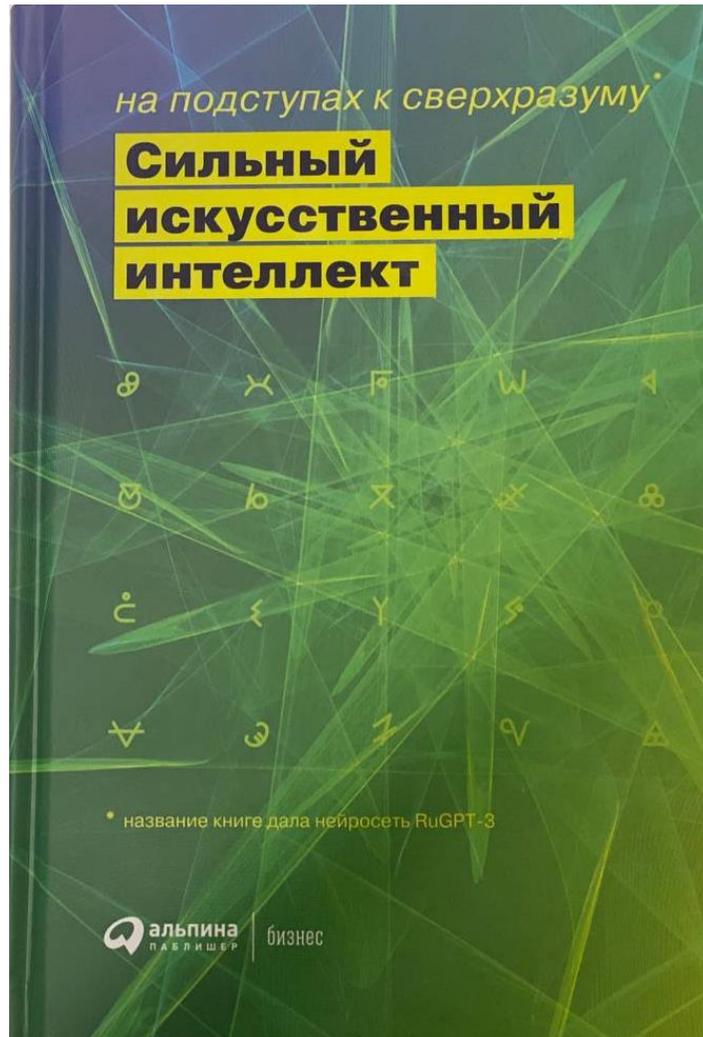
Решение, которые будут удовлетворять многочисленные потребности **бизнеса и государственного сектора**

## СООБЩЕСТВО



Центр **открытого** сообщества, объединенный общей целью— **развитие AGI**, чтобы помочь обществу и преодолеть проблемы **ESG**

# На подступах к сверхразуму. СИЛЬНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ



Название книге дала нейросеть RuGPT-3

## Оглавление

К читателям	5
Коллектив авторов	9
Введение	11
Глава 1. Общая история искусственного интеллекта	35
Глава 2. Как мы узнаем, что создали AGI?	59
Глава 3. Основные направления в AGI	87
Глава 4. Варианты воплощения	159
Послесловие. Будущее AGI	223

- ❑ Первая книга по AGI в России
- ❑ Полный и глубокий обзор подходов к созданию AGI на русском языке.
- ❑ Более 20 авторов