

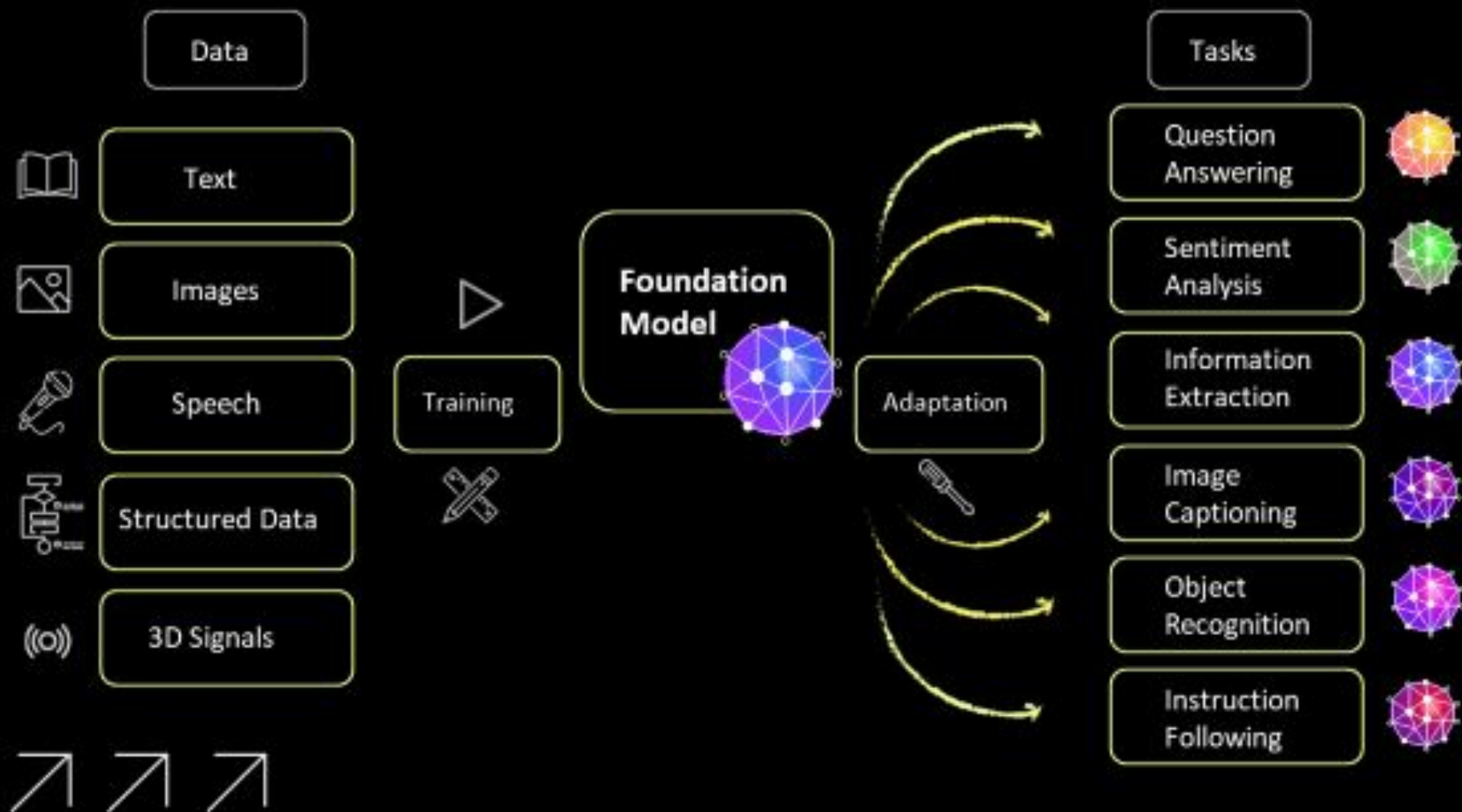
Приложения генеративного ИИ и CV

Текст

Выход: текст

Задачи:

- summarization
- writing assistance
- code generation
- language translation
- sentiment analysis



ИИ-болтун

основа - Large Language Models

GPT-3, PaLM, Galactica, and LLaMA

генерируют текст

здорово рассуждают

обладают пространственным мышлением

зачатки математического мышления

стилилизация текста

в бизнесе: прогнозирование, оценка рисков, оптимизация запасов

Текст-в-науку (Text-to-Science)

Galactica - LLM что хранит, комбинирует и рассуждает о (может на?) научном языке

Minerva - LLM для задач с вычислениями, в области математики, научных и инженерных проблем на уровне колледжа

И чем это грозит?

Симуляция авторского стиля (Text-to-Author Simulation)

GrammarlyGO - умеет писать черновики, наброски, ответы и исправления

PEER - еще и пояснит за свои действия и затюнена под академические статьи

Текст-в-Диагноз (Text-to-Medical Advice)

дисclaimer: модели слушать не безопасно, консультируйся с докторами!

Chatdoctor, GlassAI, Med-PaLM 2, YourDoctor AI

крупнейший стартап: Hippocratic AI

Документы-в-текст (Doc-to-Text)

ChatDOC, MapDeduce

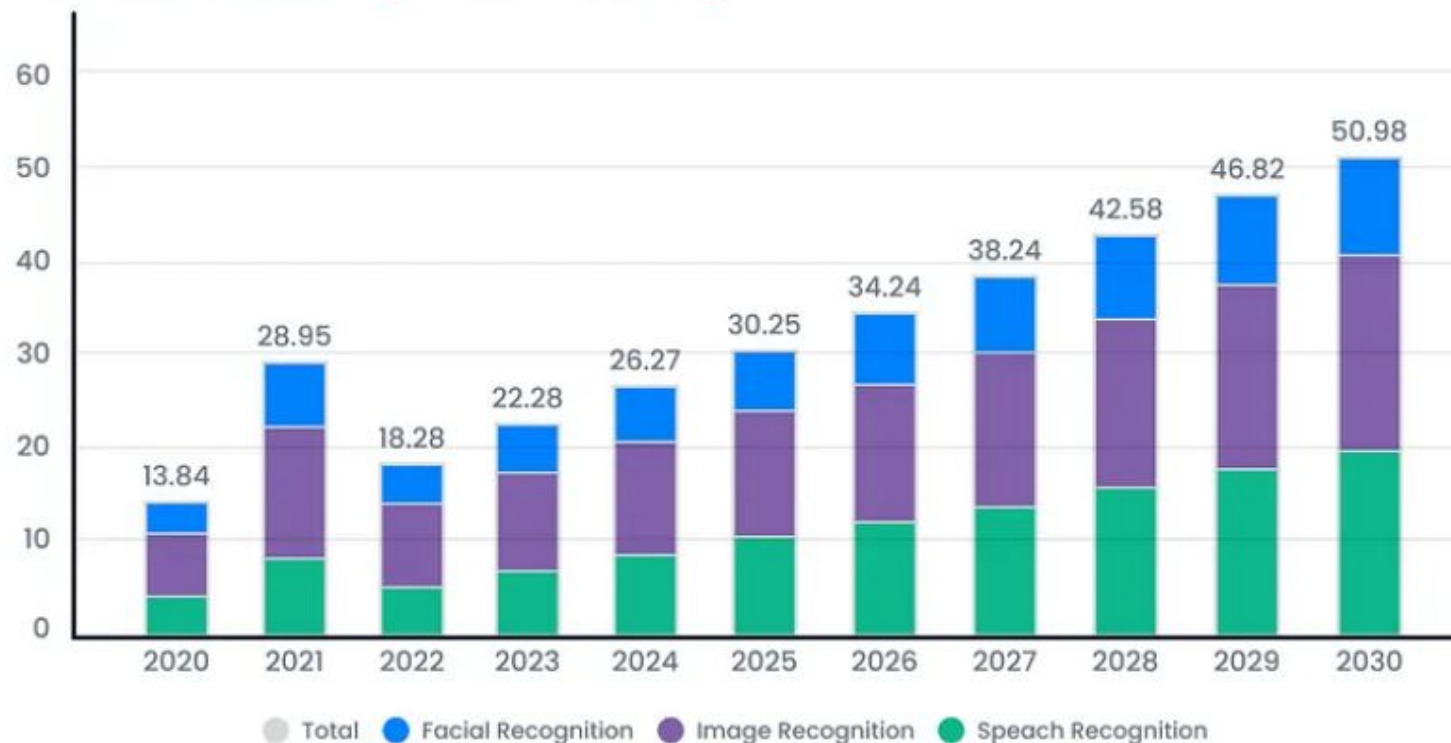
извлечение

локализация

суммаризация

Computer vision market

MARKET SIZE (in billion USD)



Research Areas in **Computer** Vision

Deepfakes

3D
Computer Vision

Real-Time
Computer Vision

Augmented Reality

Synthetic Data &
Generative AI

Vision Models

Edge Computing

Advanced
Satellite Vision

Ethics in
Computer Vision

Healthcare

Изображения

редактирование

генерация

генерация высокореалистичных изображений



Картины, созданные
искусственным интеллектом
Botto, были проданы
на аукционе за \$1 млн



Редактирование (Image Editing)

Alpaca AI , I2SB, Facet AI, Potoroom AI, Photoshop

возможности: inpainting, outpainting, upscaling, superresolution, deblurring and depth map generation

Художественные изображения (Artistic Images)

OpenART, Midjourney, Stable Diffusion, Mage.Space, NightCafe, CLIP-Guided Diffusion, VQGAN+CLIP, Neural Style Transfer



Художник Джейсон Аллен прошел 900 итераций и потратил более 80 часов, чтобы создать и доработать свои картины в Midjourney. Он видит себя творцом и художником и не откажется от заслуженного первого места на конкурсе штата Колорадо.



Реалистичные изображения (Realistic Images)

Creator, Craiyon , DALL-E 2 , GLIGEN, Imagen, Midjourney ,Muse, Parti Runway ML, Text-to-Image, Stable Diffusion ML

промт-инженерия

а есть ли минусы?

Видео

активно развивается

несмотря на сложности, какие?

появляется все больше интересных результатов

<https://www.youtube.com/watch?v=qSewd6laj6I>

<https://www.youtube.com/watch?v=cCic6tsnSUc>

<https://github.com/shubham-goel/4D-Humans>

<https://www.technologyreview.com/2023/06/01/1073858/surreal-ai-generative-video-changing-film/>

Текст-в-видео (Text-to-Video)

Imagen Video - каскад диффузионок

Meta Make a Video - text-to-video,image-tovideo and video-editing

Phenaki - видео в несколько минут с помощью промтов

Runway Gen-2 - генерит видео из текста, картинки, другого видео

Генерация видео

Говорящие синтетики - Colossyan AI , Elai AI , Heygen AI Hour One AI Rephrase AI, Synthesia

SuperCreator - мобильное приложение для создания коротких видео из описания

Персонализация видео

Tavus AI - персонализируй видео под различные аудитории

Генерация 3D

Из текста:

Adobe Firefly, Dreamfusion, GET3D, Magic3D, Synthesis AI, Text2Room

Из изображения:

GeNVS, Kaedim, Make-It-3D, RealFusion

Из нескольких изображений:

NVIDIA Lion, EVA3D, Neural-Lift-360, Scenedreamer

Из видео:

Deepmotion, Plask AI,

Изображение-в-текст (Image-to-text)

Flamingo, Segment Anything, VisualGPT

GriT - трансформер, изучающий регионы изображения и текстовые пары

Приложения

Копирайтинг: Anyword, Copy AI, Google Workspace- Gmail and Docs, Hyperwrite, Jasper, Letterdrop, Regie AI, Simplified AI, Type AI, Writesonic

пишут электронные письма, контент для веб-сайтов, черновики, создают маркетинговый контент и описания продуктов

контент для соц медиа: Clips AI, Pictory AI, Predis AI, Tweethunter, Tweetmonk
персонализированные боты: One Reach AI, OpenSight AI, Brainfish, Yuma AI

финансы: BloombergGPT, Quilt Labs AI

для научного ресерча: Agolo AI, ArxivGPT, ConsensusNLP, Elicit AI, Koala

юриспруденция: Casetext CoCounsel, Darrow AI, Harvey AI, Spellbook Legal

Приложения

игровая индустрия:

CSM, Illiad AI, Latitude

MarioGPT - генерация уровней с помощью LLM

Character AI, ConvAI , InWorld AI, RCT AI Chaos Box - генерация NPC

нейрофизиология:

Speech From Brain, Non-Invasive Brain Recordings

Генеративные AI продукты

BASE10 TREND MAP: GENERATIVE AI

Companies are grouped based on medium produced and segmented by use case within each medium. Companies that offer products across segments are grouped in the segment of the core product offering.

Base10



Первоисточник

<https://arxiv.org/abs/2306.02781>