



Как большие мультимодальные модели (LMM) поменяют видеоаналитику в 2024-2025 году ?



Комплексный поставщик средств безопасности И автоматизации бизнес-процессов.

www.dssl.ru



Разработчик системы видеонаблюдения, основанной на передовом ПО, для качественного решения задач автоматизации бизнес-процессов и безопасности.

www.trassir.ru

ИВАН ТАРАСКИН

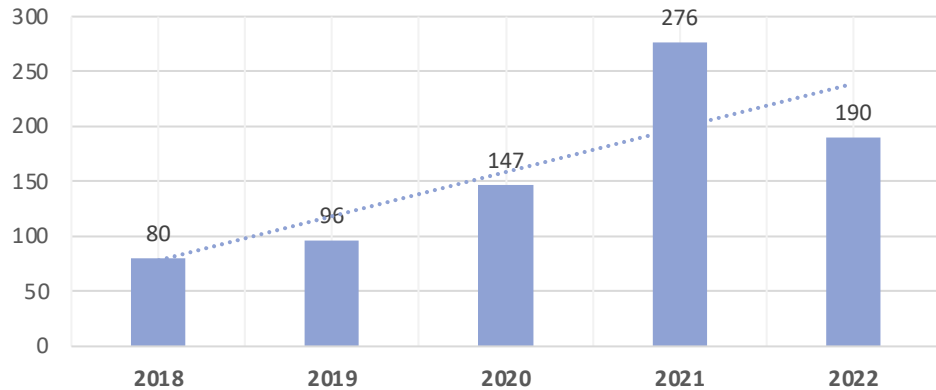
Руководитель направления
развития продуктов на базе ИИ



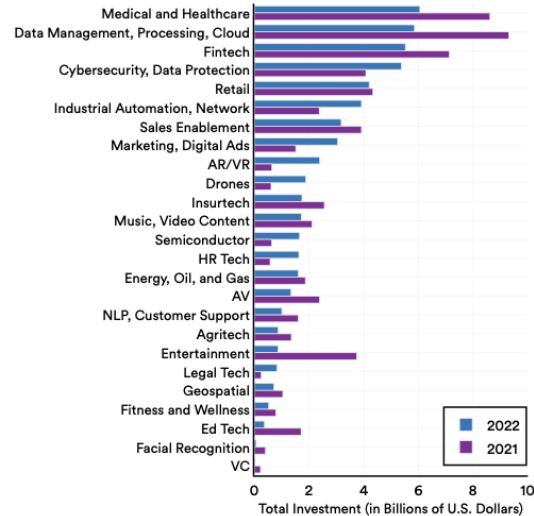
Обзор рынка ИИ

Динамика, игроки, ключевые продукты, тренды, факторы роста, место медиаконтента и видеоаналитики на рынке ИИ
Драйвером роста рынка AI на ближайшие 10 лет выступит генеративный AI и мультимодальные модели
50% компаний в 2022 году внедрили AI хотя бы в рамках одной функции или бизнес юнита**

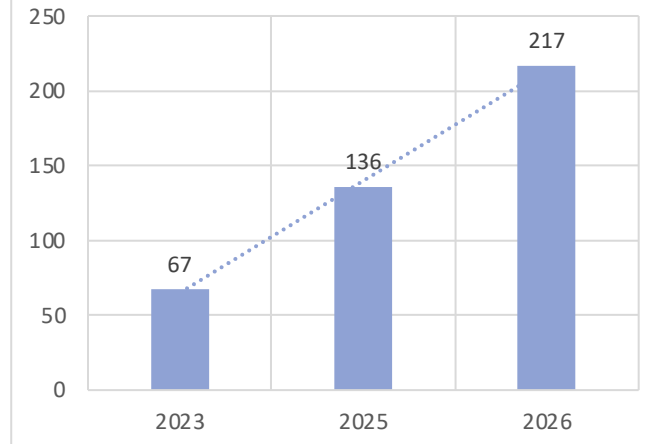
1. Глобальные корпоративные инвестиции в AI (вкл. M&A, IPO и Private Equity), в млрд долл. США*



Private Investment in AI by Focus Area, 2021 Vs. 2022
 Source: NetBase Quid, 2022 | Chart: 2023 AI Index Report



3. Выручка от генеративного AI, в млрд долл. США****



- Глобальные инвестиции в AI впервые сократились за 5 лет, что связано со зрелостью технологий и снижением кол-ва сделок;
- В ноябре 2022 года Open AI выпустил ChatGPT 3.5, что резко подогрело интерес к рынку AI решений и должно привести к дальнейшему росту инвестиций в 2023 и далее (CAGR до 38%)
- 76% компаний планируют увеличить свои инвестиции в AI проекты в 2023 году***

- Частные инвестиции в медиаконтент и технологии распознавания лиц составили около 2 млрд \$ в 2022;
- Значимые сделки по facial recognition отсутствовали в 2022.

- Темп роста CAGR составит 42% на 10 лет;
- Основными бенефициарами роста будут являться облачные провайдеры благодаря обучению и инференсу больших языковых моделей (до 55%)

* https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

** Source: McKinsey & Company Survey, 2022

**** Source: Bloomberg Intelligence, 2023

Гипотезы продуктов для видеоаналитики и видеонаблюдения (1/2)

Драйверами улучшений в продуктах видеоаналитики могут стать технологии visual reasoning ...

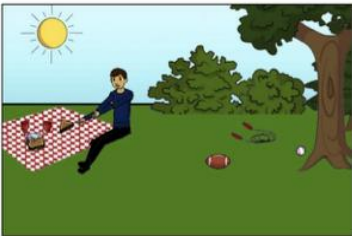
1. Visual QA reasoning – Вопросы по видео на естественном языке



How many slices of pizza are there?
Is this a vegetarian pizza?



Does it appear to be rainy?

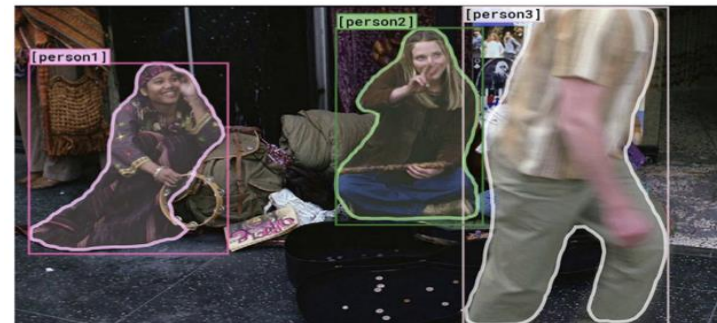


Is this person expecting company?
What is just under the tree?

В 2022 году на датасете VQA V2 наилучшие результаты показала мультимодальная модель PaLI от Google. Accurasy – 84.3%, что в среднем лучше чем у человека – 80.78%. Сопоставимые результаты у BEiT-3 от Microsoft.



2. Visual commonsense reasoning – Вопросы на понимание смысла



How did [person2] get the money that's in front of her?

- a) [person2] is selling things on the street.
 - b) [person2] earned this money playing music.**
 - c) She may work jobs for the mafia.
 - d) She won money playing poker.
- I chose b) because...
- a) She is playing guitar for money.
 - b) [person2] is a professional musician in an orchestra.
 - c) [person2] and [person1] are both holding instruments, and were probably busking for that money.**
 - d) [person1] is putting money in [person2]'s tip jar, while she plays music.

В 2022 году лучшие модели демонстрировали результат по точности распознаваний до 75.6% , что пока ниже человеческого порога – 85%.



➔ Гипотезы продуктов для видеоаналитики и видеонаблюдения (2/2)

Пользовательские сценарии на базе технологий visual reasoning и activity recognition

1. **Visual QA reasoning** – Вопросы по видео на естественном языке

- ➔ Улучшение пользовательского опыта за счет удобного поиска на естественном языке без необходимости пересмотра видеоархива => «Покажи мне кто приходил на склад ночью?»
- ➔ Выполнение простых аналитических расчетов по видео => «Посчитай сколько труб у меня осталось на улице» или «Приходил ли к нам сегодня в офис курьер?»

2. **Visual commonsense reasoning** – Вопросы на понимание смысла

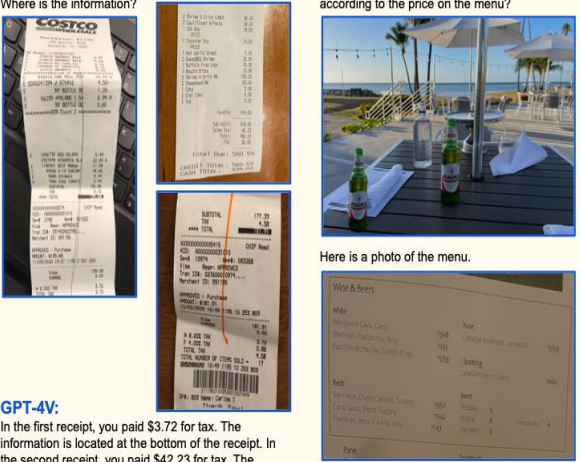
- ➔ Выполнение сложных аналитических запросов в том числе на понимание контекста происходящего => «Чем занимается человек на видео?»

Мультимодальные модели

ChatGPT (V)

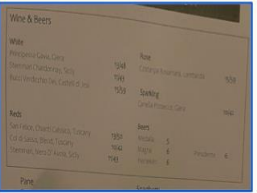
Interleaved Image-text Pairs

Prompt:
How much did I pay for tax?
Where is the information?



GPT-4V:
In the first receipt, you paid \$3.72 for tax. The information is located at the bottom of the receipt. In the second receipt, you paid \$42.23 for tax. The information is located near the bottom of the receipt, under the "Sales Tax" line. In the third receipt, you paid \$4.58 for tax. The information is located at the bottom of the receipt, under the "TAX" line. In total, you paid \$50.53 in tax across all three receipts.

Prompt:
How much should I pay for the beer on the table according to the price on the menu?



GPT-4V:
According to the menu, the price for a Magna beer is \$6. Since there are two Magna beers on the table, the total cost for the beers would be \$12.

Ноябрь 2023

<https://export.arxiv.org/pdf/2309.17421>

OMNIFUSION



Is it suitable to eat? 19:11 ✓

No, it is not suitable to eat. 19:11

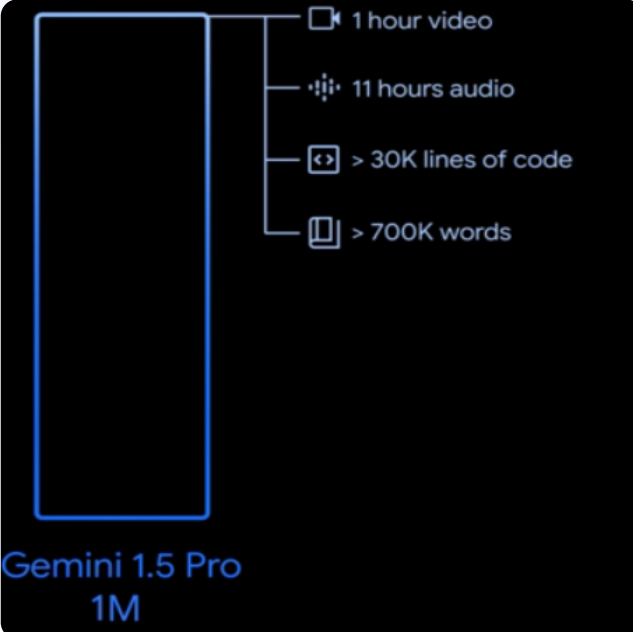
Why not? Explain please 19:11 ✓

The piece of toast on the plate is covered in mold, which is a sign of food spoilage or contamination. Eating moldy food can lead to various health issues, such as allergic reactions, digestive problems, and infections. In this case, it is best to discard the moldy toast and avoid consuming it. 19:11

Ноябрь 2023

<https://habr.com/ru/companies/airi/articles/775108/>

GEMINI



- 1 hour video
- 11 hours audio
- > 30K lines of code
- > 700K words

Gemini 1.5 Pro 1M

Февраль 2024

TRASSIR AI Incident Expert

AI Анализ видеопотока

- поиск и разбор любых инцидентов
- детальные отчеты о происшествиях их причинах и вовлеченных лицах



ОТЧЕТ ОБ ИНЦИДЕНТЕ

- **Название инцидента:** Авария на алкогольном складе
- **Описание обстоятельств инцидента:** На фотографии видно, что произошло нарушение целостности полок с алкогольной продукцией, в результате чего бутылки разных видов алкоголя оказались разбросаны по полу. Один из сотрудников пытается убрать последствия, скорее всего, собирая неразбитые бутылки и осколки стекла.
- **Место и время инцидента:** Конкретные место и время инцидента по фотографии определить невозможно, однако судя по освещению и внутреннему убранству, это склад или магазин с алкогольной продукцией, событие произошло внутри помещения.
- **Количество участников в инциденте и их описание:** На фотографии виден только один человек, предположительно работник или владелец, который занимается уборкой. Одежда у него неформальная, что может говорить о непредвиденном характере событий.
- **Потенциальный размер ущерба от инцидента (в штуках товаров и деньгах):** Точное количество разбитых товаров установить невозможно без подробного учета, но исходя из видимой площади разлива и количества бутылок на полу, можно предположить, что разбито более десятка бутылок. Ущерб в деньгах будет зависеть от стоимости каждой разбитой бутылки; если предположить среднюю стоимость бутылки в 15-20 долларов, ущерб может составить от 150 до 200 долларов и выше.



РАССЛЕДОВАНИЕ ОБСТОЯТЕЛЬСТВ ИНЦИДЕНТА И ЕГО ПРИЧИН

Оценка вероятности причин инцидента на основе одного фото не может быть точной и останется чисто гипотетической, так как необходима дополнительная информация и контекст. Но исходя из того, что мы видим:

- **Физический дефект полок:** Если полки выглядят старыми или поврежденными, вероятность может быть высокой. Однако, если они кажутся в хорошем состоянии, вероятность снижается. Вероятность: 30%.
- **Ошибки при укладке товара:** Если видно, что бутылки уложены неравномерно, возможно, что это причина. Вероятность: 20%.
- **Столкновение с полками:** Если в магазине есть следы, указывающие на возможное столкновение (например, следы на ковре), это увеличивает вероятность. Вероятность: 10%.
- **Сейсмическая активность:** В случае если магазин находится в сейсмически активном регионе, вероятность может быть выше. Вероятность: 5%.
- **Нарушение техники безопасности:** Без дополнительных деталей сложно оценить. Вероятность: 15%.
- **Производственный дефект полок или их креплений:** Если полки относительно новые и обрушились, это может указывать на производственный дефект. Вероятность: 10%.
- **Вандализм или преднамеренное повреждение:** Требуется дополнительная информация для оценки этой вероятности. Вероятность: 5%.
- **Вандализм или преднамеренное повреждение:** Требуется дополнительная информация для оценки этой вероятности. Вероятность: 5%.

Эти проценты представляют собой грубые оценки, основанные на общих причинах подобных происшествий и могут не отражать действительную вероятность в конкретной ситуации, изображенной на фото.



ОТЧЕТ ОБ ИНЦИДЕНТАХ ЗА 20 НОЯБРЯ 2023 ГОДА

➤ **Магазин:** [Название магазина]

➤ **Дата:** [ДД.ММ.ГГГГ]

№ п/п	Время	Место происшествия	Описание инцидента	Участники	Предпринятые действия	Оценка ущерба
1	10:05	Продовольственный отдел	Падение банок с полки	Сотрудник отдела	Уборка места, проверка полок на надежность	Незначительный
2	12:45	Зона касс	Несанкционированный доступ в служебную зону	Посетитель	Обращение к охране, извлечение посетителя из зоны	Без ущерба
3	14:30	Отдел электроники	Сломанный дисплей смартфона	Покупатель	Запись инцидента, информирование менеджмента, оформление документов для страховки	Средний
4	16:20	Парковка	Столкновение тележек с автомобилем	Покупатель, сотрудник парковки	Оформление протокола, вызов полиции, контакты страховой компании	Значительный
5	17:55	Отдел одежды	Разлив кофе на товар	Покупатель	Уборка, отправка товара на химчистку, информирование клиента о компенсации	Незначительный
6	19:40	Склад	Неправильная погрузка товара	Сотрудник склада	Инструктаж сотрудника, переучет товара, проверка оборудования	Минимальный

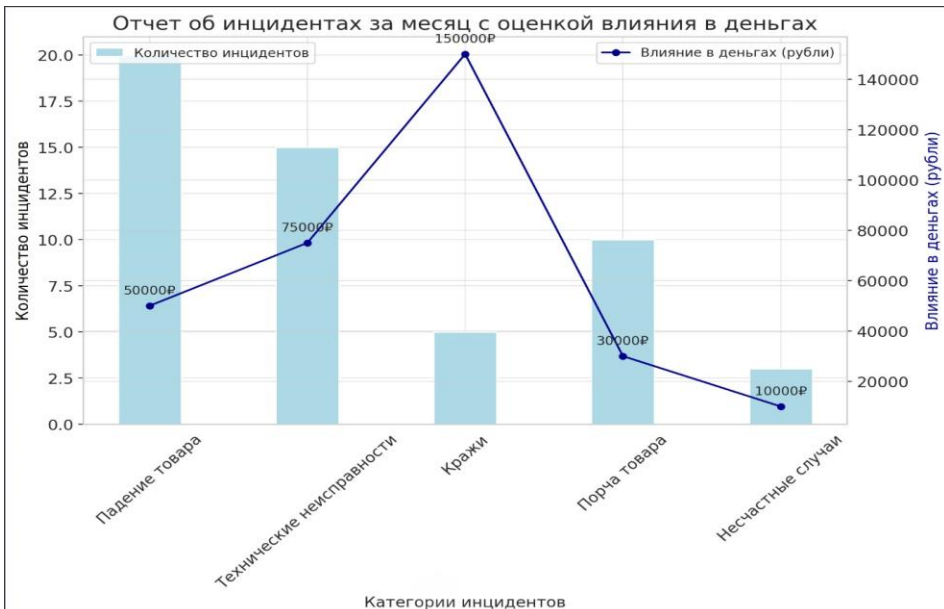
➤ **Заключение:** В течение дня было зарегистрировано 6 инцидентов. Большинство из них не повлекли за собой значительного ущерба и были оперативно урегулированы сотрудниками. Один инцидент на парковке потребовал вмешательства полиции и будет расследован дополнительно. Все инциденты были задокументированы согласно внутренним процедурам безопасности.

➤ **Подпись ответственного:** _____



РЕКОМЕНДАЦИИ ПО СНИЖЕНИЮ РИСКОВ В БУДУЩЕМ

Для предотвращения подобных инцидентов в будущем, как на изображении с разбитыми бутылками в магазине, можно предпринять следующие меры:



- **Укрепление полок и стеллажей:** Убедиться, что все полки и стеллажи соответствуют стандартам безопасности и способны выдерживать вес хранящейся на них продукции.
- **Правильная укладка товаров:** Располагать тяжелые товары на более нижних полках и легкие – на верхних, чтобы снизить риск их падения.
- **Регулярный осмотр и техническое обслуживание:** Вводить регулярные проверки состояния полок и стеллажей, чтобы выявлять и устранять потенциальные проблемы до того, как они приведут к инцидентам.
- **Регулярный осмотр и техническое обслуживание:** Вводить регулярные проверки состояния полок и стеллажей, чтобы выявлять и устранять потенциальные проблемы до того, как они приведут к инцидентам.
- **Использование защитных барьеров:** Установить барьеры или ограничители на полках, чтобы предотвратить скатывание бутылок и других товаров.
- **Введение системы предупреждения:** Использовать сигнальные системы для предупреждения сотрудников о перегрузке полок или о нестабильности оборудования.
- **Улучшение условий освещения:** Обеспечить достаточное освещение в торговом зале, чтобы сотрудники могли лучше видеть и избегать потенциально опасных ситуаций.
- **Ограничение доступа к зонам с высоким риском:** Ограничить доступ клиентов к зонам, где располагаются тяжелые или легко разбиваемые товары.
- **Инвестиции в технологии:** Рассмотреть возможность внедрения автоматизированных систем для хранения и перемещения товаров, чтобы уменьшить риск человеческой ошибки.
- **Страхование:** Поддерживать актуальный полис страхования для покрытия потенциальных убытков от подобных инцидентов.



Внедрение и соблюдение этих мер поможет снизить риск подобных инцидентов и их потенциальное влияние на работу магазина

СПАСИБО ЗА ВНИМАНИЕ

ИВАН ТАРАСКИН

Руководитель направления
развития продуктов на базе ИИ

📧 @ihorrible

☎ +7 (916) 315-62-20

