На ежегодной конференции для разработчиков Google I/O 2024 компания анонсировала ряд новых продуктов и функций, связанных с искусственным интеллектом. Среди ключевых анонсов — обновленная генеративная модель Imagen 3 для создания изображений, а также нейросеть Veo для генерации видео.

Imagen 3 отличается улучшенным пониманием текстовых запросов, меньшим количеством ошибок и способностью создавать более четкий текст на изображениях. Модель будет доступна в нескольких версиях, оптимизированных под различные задачи — от быстрых эскизов до изображений высокого разрешения.

В свою очередь, Veo позиционируется как ответ Google на разработку OpenAI под названием Sora. Новая генеративная модель позволит создавать видеоролики с разрешением 1080p в различных стилях. На данный момент компания предоставила доступ к Veo некоторым авторам YouTube и представителям киноиндустрии для тестирования.

Кроме того, на конференции был показан прототип мультимодального ИИ-помощника Project Astra. В перспективе он должен стать универсальным виртуальным ассистентом, взаимодействовать с которым можно будет через камеру устройства. Помощник сможет описывать код на экране, помогать в поиске предметов и решении задач в реальном времени.

Google также сообщает о скором расширении возможностей ИИ-ассистента Gemini на устройствах Android. Он сможет отвечать на вопросы пользователей о видео на экране, опираясь на автоматически сгенерированные подписи. Для платных подписчиков Gemini станет доступна функция загрузки и анализа PDF-файлов.

Ещё одним анонсом стал запуск программы Gems для создания пользовательских чат-ботов. По аналогии с GPTS от OpenAI, Gems даст возможность настраивать реакции и специализацию ИИ-помощника Gemini с помощью инструкций.

Также Google объявила о внедрении облегчённой версии языковой модели Gemini Nano в браузер Chrome для десктопов. Встроенный ассистент будет использовать ИИ на устройстве для помощи в генерации текстов постов, обзоров и прочего контента прямо в Chrome.

Помимо этого, последняя версия языковой модели Gemini 1.5 Pro интегрируется в сервисы Google Workspace. ИИ-помощник сможет извлекать информацию из файлов на Google Диске, писать электронные письма и устанавливать напоминания. Функция

Все права защищены

станет доступна платным подписчикам Gemini в следующем месяце.

Наконец, летом Google запустит функцию «Спросите фотографии», которая позволит Gemini анализировать медиатеку Google Фото для ответа на вопросы пользователей. Система сможет находить конкретные детали на снимках, например, номерной знак автомобиля. Кроме того, в Google Lens появится возможность поиска по видео с помощью ИИ.

Все права защищены