



تازه‌ترین اخبار هوش مصنوعی آگوست ۲۰۲۶

اگر احساس می‌کنید هر روز یک مدل جدید، یک ابزار تازه یا یک قابلیت عجیب و غریب در دنیای هوش مصنوعی معرفی می‌شود، کاملاً حق دارید! سرعت پیشرفت هوش مصنوعی به حدی رسیده که دنبال کردن همه خبرها تقریباً به یک کار تمام وقت تبدیل شده است؛ اما در میان این همه سر و صدا، همه خبرها به یک اندازه مهم نیستند.

در این مقاله قرار نیست فقط فهرستی از تازه‌ترین اخبار هوش مصنوعی را مرور کنیم. هدف این است مهم‌ترین اتفاقات و معرفی‌های اخیر را ساده و کاربردی بررسی کنیم و ببینیم هرکدام دقیقاً چه کاری انجام می‌دهند، چه چیزی در آن‌ها بهتر شده و این تغییرات چه معنایی برای کاربران، تولیدکنندگان محتوا و کسب‌وکارها دارند.

از ایجنت‌هایی که می‌توانند به ابزارهای کاری شما متصل شوند و واقعاً بخشی از کار را انجام دهند، تا مدل‌هایی که تولید تصویر و ویدیو را سریع‌تر و واقعی‌تر کرده‌اند؛ از مدل‌های قدرتمند و گران‌قیمت گرفته تا مدل‌های کوچک و ارزان که می‌توانند روی

سیستم شخصی شما اجرا شوند. اگر می‌خواهید تا حدی در جریان مهم‌ترین تحولات هوش مصنوعی باشید، این گزارش تصویری فشرده از اتفاقات اخیر به شما می‌دهد؛ بدون اینکه مجبور باشید هر روز ده‌ها خبر و معرفی جدید را دنبال کنید.

خلق دنیای سه‌بعدی

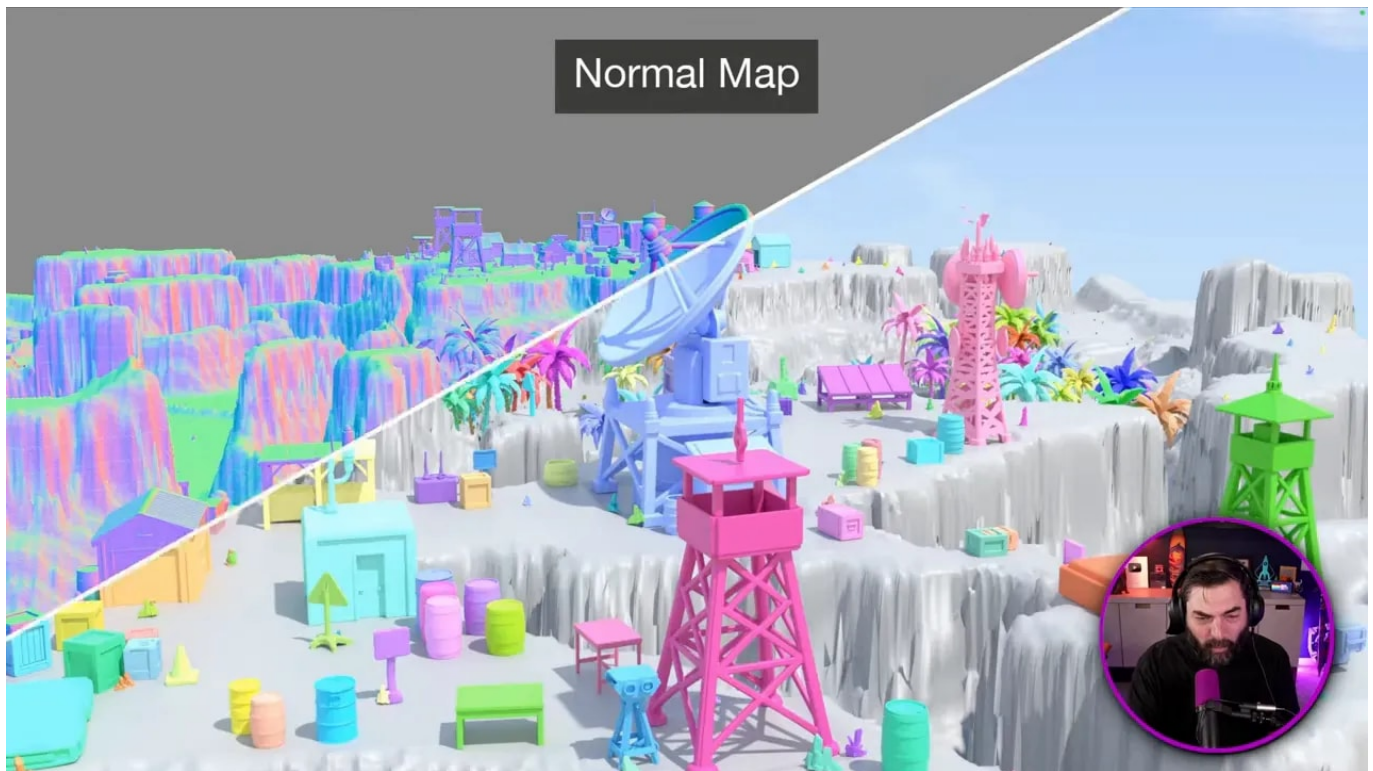
یکی از تازه‌ترین اخبار هوش مصنوعی از شرکت بزرگ Tencent و بخش هوش مصنوعی آن یعنی **Tencent Hunyuan** است که نشان می‌داد آن‌ها با پروژه‌ای جدید به نام WorldClaw توانسته‌اند دنیاهای سه‌بعدی بزرگ و کاملاً قابل ویرایش را به کمک هوش مصنوعی بسازند.

Tencent که یکی از غول‌های تکنولوژی و بزرگ‌ترین شرکت‌های بازی‌سازی دنیا است، فعلاً فقط یک مقاله علمی و چند ویدیو دمو از این دستاورد منتشر کرده است؛ بنابراین اگر الان به صفحه گیت‌هاب (GitHub) آن‌ها سر بزنید، هنوز هیچ کد برنامه‌نویسی یا نرم‌افزاری برای دانلود و استفاده عموم قرار نداده‌اند.

پروژه WorldClaw دقیقاً چیست؟

تا همین چند وقت پیش، نمی‌توانستید با یک خط متن، یک دنیای سه‌بعدی واقعی و قابل ویرایش بسازید. هوش مصنوعی صرفاً یک تصویر تخت (مثل یک نقاشی) تحویل می‌داد و شما نمی‌توانستید درختی را در آن عکس جابه‌جا کنید یا زاویه یک خانه را تغییر دهید؛ چون همه‌چیز ثابت و به هم چسبیده بود.

اما با WorldClaw، همه‌چیز تغییر می‌کند. این ابزار به‌جای یک نقاشی تخت، یک «جعبه لگو» به شما می‌دهد. وقتی می‌نویسید «یک روستای برفی»، یک محیط کامل ساخته می‌شود که در آن تک‌تک اجزا (درخت، خانه، گاری و چراغ‌ها) کاملاً مستقل هستند و می‌توانید هرکدام را به‌دلخواه جابه‌جا یا ویرایش کنید.



این فناوری چگونه کار می‌کند؟

تنسنت در این سیستم از روشی استفاده کرده که به آن «کل به جزء» می‌گویند. یعنی هوش مصنوعی مثل یک تیم شهرسازی واقعی عمل می‌کند، طی چهار مرحله:

۱. تحلیل و برنامه‌ریزی (اتاق فکر)

کاربر فقط یک خط می‌نویسد، مثلاً: «یک دره رودخانه‌ای سرسبز». اما هوش مصنوعی برنامه‌ریز وارد عمل می‌شود و با خودش می‌گوید: «این دره نیاز به کوه، سنگ، درخت، مسیر آب، کلبه و نور آفتاب دارد». او یک نقشه متنی کامل از چیزی که نیاز است می‌نویسد.

۲. ساخت زمین و پستی و بلندی‌ها (زیرسازی)

سیستم در این مرحله هیچ درختی نمی‌کارد! فقط یک نقشه دوبعدی شبیه نقشه جغرافیا می‌سازد که مشخص می‌کند کجاها کوه است، کجاها دشت است و مسیر رودخانه کجاست. سپس به این نقشه مسطح، ارتفاع و حجم می‌دهد تا زمین

سه بعدی ما (تپه‌ها و دره‌ها) شکل بگیرد.

۳. ساخت و چیدن وسایل (مرحله جزئیات)

حالا سراغ مناطقی می‌رود که نیاز به جزئیات دارند. هوش مصنوعی برای هر منطقه یک عکس تولید می‌کند، تک‌تک اشیاء درون عکس (مثل کلبه یا درخت) را از عکس برش می‌زند، آن‌ها را تبدیل به مدل‌های سه بعدی (آبجکت) می‌کند و سر جایشان روی زمین سه بعدی می‌کارد.

۴. بازرسی نهایی (رفع باگ‌ها)

در مرحله آخر، یک هوش مصنوعی ناظر، کل محیط را بررسی می‌کند تا مطمئن شود همه چیز طبیعی است. مثلا چک می‌کند پایه‌های یک گاری روی هوا نمانده باشد یا یک درخت درون یک سنگ فرو نرفته باشد و همه چیز دقیقا روی زمین نشسته باشد.

ماجرا فقط به ساخت بازی ختم نمی‌شود!

قطعا یکی تازه‌ترین اخبار هوش مصنوعی همین امکان ساخت سریع مراحل و محیط‌ها برای بازی بودت؛ اما ماجرا فقط به بازی ختم نمی‌شود.

این ابزارها به افراد متخصص کمک می‌کنند نتایج بهتری بگیرند و سرعت کارشان را در زمینه‌هایی مثل انیمیشن‌های سه بعدی کوتاه و نمونه‌سازی محیط‌های مجازی بالا ببرند. همچنین در آموزش ربات‌ها بسیار کاربردی هستند؛ چراکه ربات‌ها می‌توانند پیش از ورود به دنیای واقعی، با سناریوهای متنوع در این محیط‌های شبیه‌سازی شده و بی‌خطر تمرین کنند.

فعلا WorldClaw برای استفاده عمومی در دسترس نیست و مخزن GitHub آن هم، هنوز کدهای قابل‌توجهی برای اجرا ندارد؛ با این حال، یکی از آن پروژه‌هایی است که کاملا ارزش دارد زیر نظر بماند.

Grok Bot؛ ایجت‌های ابری برای کارهای واقعی

با وجود خرید اخیر Cursor توسط Grok xAI، ابزار Grok Bot اصلا مختص

برنامه‌نویس‌ها طراحی نشده است. این ابزار بیشتر رقیبی برای Claude Cowork محسوب می‌شود؛ محیطی ساده که اجازه می‌دهد بدون نیاز به هیچ‌گونه دانش کدنویسی یا درگیری با مباحث فنی، «ایجنت‌های خودکار» بسازید. در این پلتفرم می‌توانید دستیارهای هوشمندی با وظایف کاملاً مشخص خلق کنید. نقش‌های پیشنهادی خود ابزار شامل موارد زیر است:

- مدیریت و مرتب‌سازی ایمیل‌ها
- دستیار شخصی یا مدیر اجرایی
- مذاکره‌کننده
- پژوهشگر و محقق
- خریدار
- جست‌وجوگر آپارتمان
- نمونه‌ساز محصول

نام ایجنت، نقطه شروع نقش آن است

جالب است بدانید حتی نامی که برای یک ایجنت انتخاب می‌کنید، می‌تواند به شکل‌گیری نقش و نحوه عملکرد آن کمک کند. برای مثال، وقتی نام ایجنت را «Research Bot» می‌گذارید، از همان ابتدا مشخص است قرار است نقش یک دستیار پژوهشی را ایفا کند؛ یعنی منابع را جست‌وجو کند، اطلاعات مرتبط را پیدا کند و نتیجه را به‌شکلی روشن و خلاصه ارائه دهد.

نمونه‌هایی از ایجنت‌های متخصص

مثلاً می‌توانید برای هر کار، یک ایجنت تخصصی داشته باشید که وظیفه مشخصی را بر عهده بگیرد و با اتصال به ابزارهای موردنیاز، آن را به‌صورت خودکار انجام دهد:

۱. ایجنت پژوهشی

هر روز تازه‌ترین اخبار و تحولات هوش مصنوعی را بررسی می‌کند و خلاصه‌ای از

مهم‌ترین اتفاقات، اهمیت آن‌ها و نکات کلیدی ارائه می‌دهد.

۲. ایجنت مدیریت ایمیل

به Gmail متصل می‌شود، ایمیل‌های مهم را شناسایی می‌کند، برای برخی پیام‌ها پیش‌نویس پاسخ می‌سازد، ایمیل‌های تبلیغاتی را سامان می‌دهد و خلاصه‌ای از وضعیت صندوق ورودی ارائه می‌کند.

اتصال ایجنت به ابزارهای واقعی کار

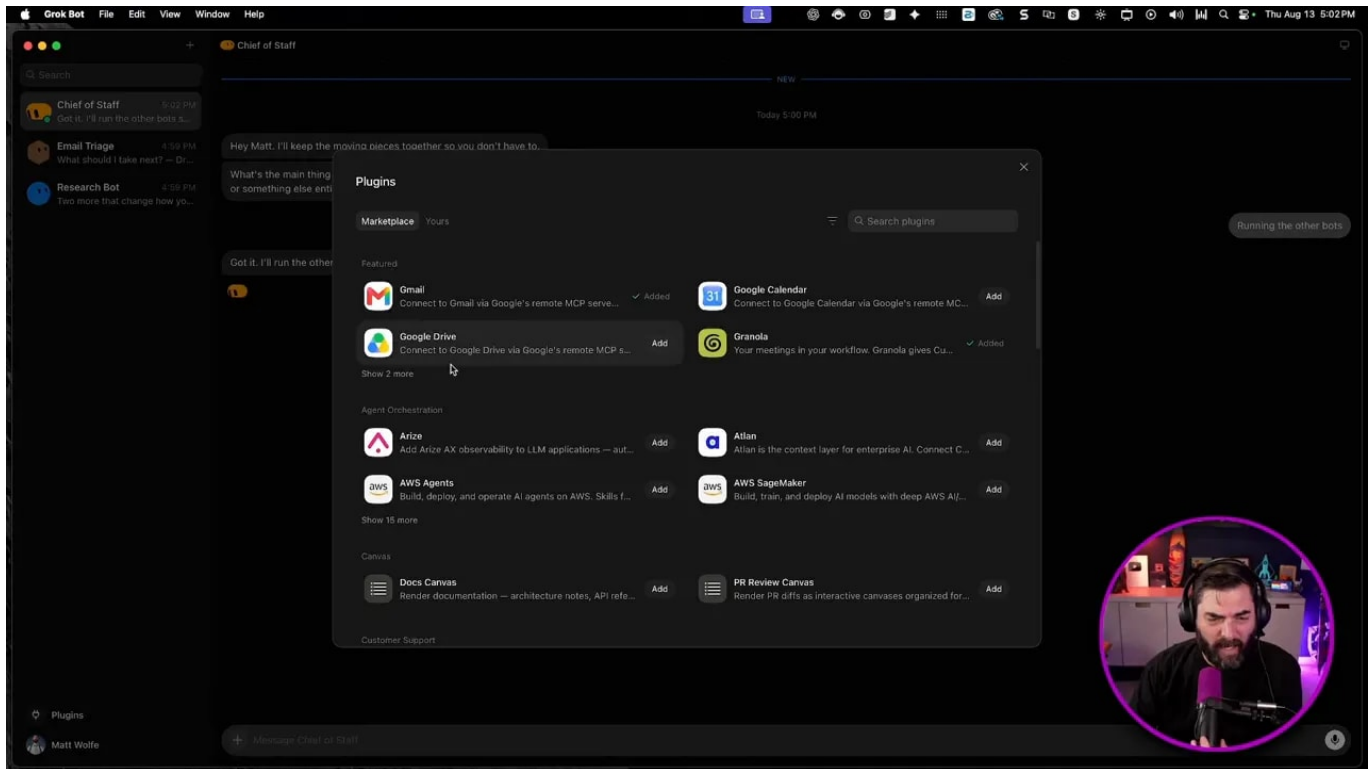
قدرت واقعی ایجنت‌ها زمانی بیشتر مشخص می‌شود که آن‌ها را به ابزارهایی که هر روز با آن‌ها کار می‌کنید متصل کنید. Grok Bot می‌تواند به ابزارها و سرویس‌های مختلفی از جمله موارد زیر متصل شود:

- Gmail
- Google Drive
- Google Calendar
- Notion
- Slack
- Granola
- ابزارهای پشتیبانی مشتری
- ابزارهای تحلیل داده
- ابزارهای هماهنگی میان ایجنت‌ها

بنابراین، قرار نیست ایجنت فقط یک چت‌بات باشد که با آن گفت‌وگو می‌کنید. ایده اصلی این است که ایجنت به محیط واقعی کار دسترسی داشته باشد، اطلاعات را از ابزارهای مختلف دریافت کند، میان آن‌ها ارتباط برقرار کند و بر اساس هدفی که برایش تعیین کرده‌اید، بخشی از کار را به صورت عملی انجام دهد.

به بیان ساده، تفاوت یک چت‌بات معمولی با یک ایجنت در این است که چت‌بات

بیشتر «پاسخ می‌دهد»؛ اما ایجنت می‌تواند داده‌ها را جمع‌آوری کند، ابزارها را به کار بگیرد و اقدام کند.



قابلیت‌های دیگر ایجنت‌های Grok Bot

یکی از جذاب‌ترین ویژگی‌های Grok Bot این است هر ایجنت، انگار یک کامپیوتر شخصی در فضای ابری در اختیار دارد؛ کامپیوتری که لازم نیست خودتان آن را روشن یا مدیریت کنید. ایجنت در همین محیط می‌تواند وارد سایت‌ها شود، فایل‌ها را جابه‌جا کند، با ابزارهای مختلف کار کند و حتی از طریق ترمینال، دستور اجرا کند. بنابراین، اجرای کارها وابسته به روشن بودن لپ‌تاپ شما نیست.

آموزش دادن به ایجنت

برای اینکه این موضوع ملموس‌تر شود، فرض کنید یک کار تکراری دارید که هر هفته انجام می‌دهید. کافی است یک بار آن را به ایجنت نشان دهید:

Teach a Task

ضبط صفحه را روشن می‌کنید و کار را همان‌طور که همیشه انجام می‌دهید، مرحله به مرحله اجرا می‌کنید. ایجنت این فرایند را یاد می‌گیرد و برای انجام دوباره آن، یک مهارت ایجاد می‌کند.

Routines

مشخص می‌کنید این کار چه زمانی یا با چه رویدادی اجرا شود؛ مثلاً هر روز ساعت ۹ صبح، هر هفته یا هر ماه. حتی می‌توانید تعیین کنید با وقوع یک اتفاق، مثل دریافت پیام جدید در Slack، کار شروع شود.

اجرای مستقل

بعد از تنظیم کار، دیگر لازم نیست هر بار خودتان آن را اجرا کنید. ایجنت در رایانه ابری خودش کار را انجام می‌دهد؛ حتی اگر لپ‌تاپ شما خاموش باشد یا اصلاً در دسترس نباشد.

مدیریت از موبایل

با اپلیکیشن موبایل می‌توانید ایجنت‌ها و فعالیت‌هایشان را بررسی و مدیریت کنید؛ بدون اینکه برای انجام کارها دائماً پای کامپیوتر باشید.

در واقع، می‌توانید یک کار را یک‌بار انجام دهید، به ایجنت نشان دهید، زمان یا شرایط اجرای آن را مشخص کنید و بعد اجازه دهید خودش آن را تکرار کند. اینجاست که ایجنت از یک ابزار گفت‌وگویی ساده، به یک نیروی اجرایی همیشه در دسترس تبدیل می‌شود.

Seedance ۲.۵

Seedance ۲.۵ حالا مستقیماً در Artlist در دسترس است و برای تولید ویدیو، مدل بسیار قابل توجهی به نظر می‌رسد. یکی از ویژگی‌های مهم آن این است می‌تواند در یک خروجی، ویدیوهایی تا ۳۰ ثانیه تولید کند. این قابلیت باعث می‌شود برای ساخت یک صحنه، دیگر مجبور نباشید چندین کلیپ چندثانیه‌ای تولید کنید و بعد

آنها را به هم بچسبانید.

به هوش مصنوعی الگو بدهید!

ویژگی جذاب‌تر، امکان استفاده از تا ۵۰ تصویر مرجع است. می‌توانید از این تصاویر برای معرفی شخصیت، محصول، لوکیشن یا حتی تعیین سبک بصری ویدیو استفاده کنید. به این ترتیب، مدل می‌تواند عناصر اصلی را در نماهای مختلف با ثبات بیشتری حفظ کند؛ مثلاً چهره یک شخصیت، ظاهر یک محصول یا فضای یک لوکیشن کمتر دچار تغییر شود.

این موضوع اهمیت زیادی دارد، چون ثبات شخصیت‌ها و عناصر بصری در نماهای مختلف هنوز یکی از چالش‌های اصلی تولید ویدیو با هوش مصنوعی است. ۲.۵ Seedance با استفاده از تصاویر مرجع، تلاش می‌کند این مشکل را تا حد زیادی کاهش دهد و خروجی یکدست‌تری ارائه کند.

واترمارک هوش مصنوعی

هرچه محتوای تولیدشده با هوش مصنوعی واقعی‌تر می‌شود، تشخیص منشأ آن هم دشوارتر خواهد شد. به همین دلیل، شرکت‌های بزرگ دنبال راه‌هایی هستند تا نشانه‌ای از استفاده از AI در محتوا باقی بماند.

Claude و نشانه‌گذاری نامرئی متن

Anthropic اعلام کرده که در خروجی متنی برخی مدل‌های Claude، نشانه‌هایی نامرئی قرار می‌دهد. این نشانه‌ها برای خواننده قابل مشاهده نیستند و نباید روی معنا یا کیفیت متن تاثیری داشته باشند. این نشانه‌ها ممکن است هنگام کپی و پیست یا حتی بعد از برخی ویرایش‌ها همچنان باقی بمانند؛ البته وجود یا نبود آنها به‌تنهایی نمی‌تواند ثابت کند متنی با AI تولید شده یا انسانی است.

استفاده از AI بدون پنهان‌کاری

طبیعی است برخی تولیدکنندگان محتوا نخواهند مخاطب متوجه استفاده از AI شود؛

اما رویکرد منطقی‌تر این است هوش مصنوعی را برای ساخت پیش‌نویس اولیه به کار ببریم و سپس متن را واقعا ویرایش و شخصی‌سازی کنیم تا به لحن و سبک خودمان نزدیک شود.

آیا واترمارک راه‌حل قطعی است؟ احتمالا نه. همان‌طور که روش‌های تشخیص AI پیشرفته‌تر می‌شوند، ابزارهایی برای حذف یا دور زدن واترمارک‌ها نیز شکل خواهند گرفت. با این حال، تلاش برای ایجاد شفافیت همچنان می‌تواند قدم مثبتی باشد.

سیل مدل‌های جدید

معرفی چندین مدل جدید در یک هفته ممکن است نشانه یک جهش بزرگ به نظر برسد؛ اما همه مدل‌ها قرار نیست زندگی کاربران را تغییر دهند. بسیاری از آن‌ها فقط کمی سریع‌تر، ارزان‌تر یا در یک بنچمارک خاص بهتر هستند.

مهم است بین پیشرفت واقعی محصول و بهبود چند درصدی در یک جدول مقایسه تفاوت بگذاریم. یک مدل زمانی واقعا مهم است که این پیشرفت در استفاده روزمره هم خودش را نشان دهد.

Grok ۴.۶؛ عملکرد قوی با قیمت رقابتی

Grok ۴.۶ یکی از مدل‌های قابل‌توجه این موج است که در برخی بنچمارک‌های کدنویسی و کارهای دانشی عملکرد بسیار خوبی نشان داده است. قیمت پایین‌تر آن نیز می‌تواند برای استفاده‌های سنگین جذاب باشد. مهم‌ترین بهبودهایی که در این مدل مشاهده می‌شوند عبارتند از:

- **کدنویسی بهتر:** عملکرد قوی‌تر در حل مسائل واقعی برنامه‌نویسی.
- **مصرف کمتر توکن:** برای رسیدن به نتیجه مشابه، خروجی کوتاه‌تر و کارآمدتری تولید می‌کند.
- **سرعت بالاتر:** پاسخ‌گویی سریع‌تر، به‌خصوص در کارهای طولانی و چندمرحله‌ای.
- **عملکرد بهتر در کارهای دانشی:** از تحلیل و تحقیق گرفته تا کار با فایل‌های

Word، Excel و PowerPoint.

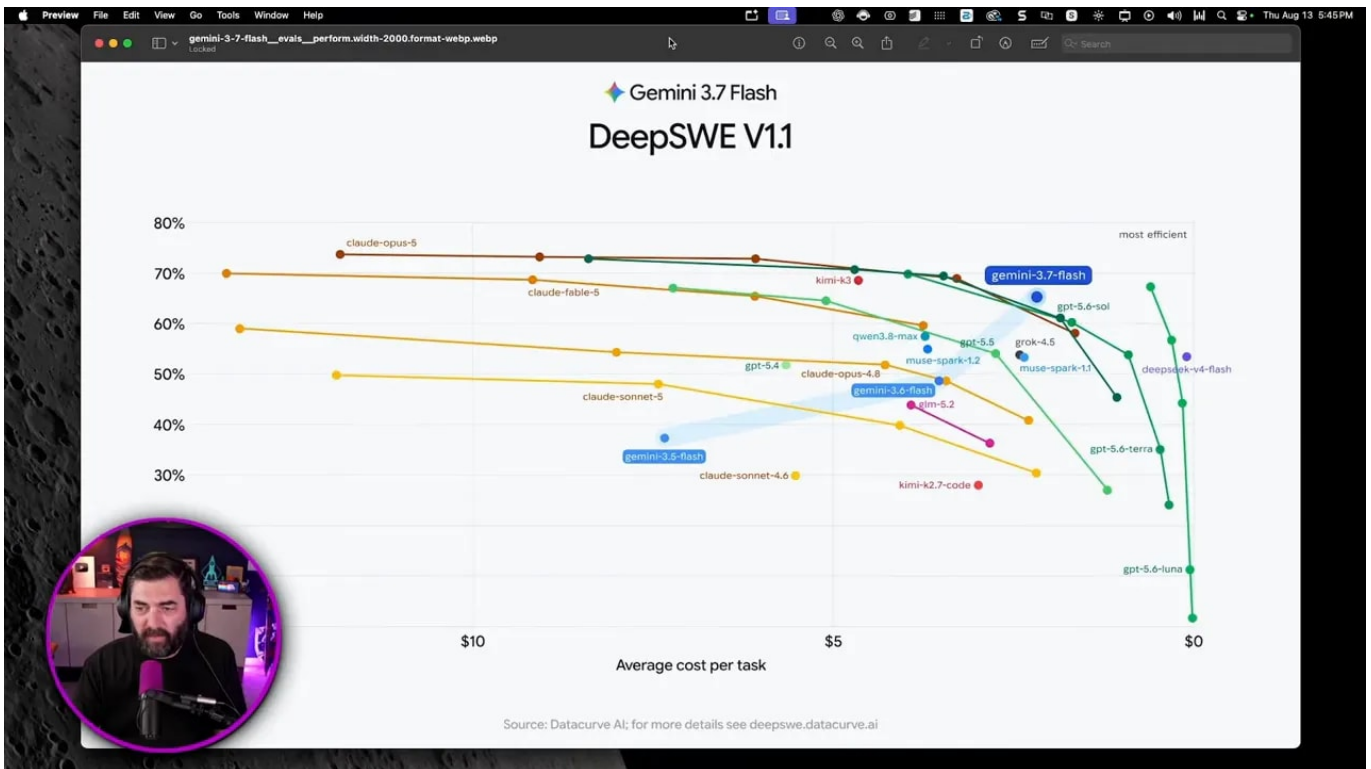
• **قابلیت‌های عامل‌محور:** تمرکز بیشتر روی انجام واقعی کارها؛ نه فقط پاسخ دادن به سوال.

در مجموع، جذابیت Grok در این نسل فقط «هوش بیشتر» نیست؛ ترکیب توانایی، سرعت و هزینه پایین‌تر است که آن را برای استفاده‌های سنگین جذاب می‌کند.

Gemini ۳.۷ Flash؛ ارزش بالا در برابر هزینه

گوگل با Gemini ۳.۷ Flash بیشتر روی تعادل میان توانایی و هزینه تمرکز کرده است. این مدل قرار نیست قدرتمندترین مدل بازار باشد؛ بلکه می‌خواهد با هزینه‌ای پایین، عملکرد قابل‌قبولی ارائه دهد. قیمت معرفی‌شده آن تا پایان سال، ۰/۷۵ دلار برای هر یک میلیون توکن ورودی و ۳/۷۵ دلار برای خروجی است. امتیاز آن در DeepSuite نیز ۶۵/۳ درصد گزارش شده است.

این مدل بیشتر توسط چه کسانی استفاده می‌شود؟ برای بسیاری از کاربران و کسب‌وکارها، همیشه قوی‌ترین مدل بهترین انتخاب نیست. گاهی مدلی که با هزینه بسیار کمتر، کیفیتی نزدیک به مدل‌های گران‌تر ارائه می‌دهد، انتخاب منطقی‌تری است.



مدل‌های Flash گوگل، مسیر واضحی دارند

کاهش مداوم هزینه در کنار پیشرفت محسوس توانایی. Gemini ۳.۷ Flash در آزمایش تولید SVG هم بسیار خوب ظاهر شد و در رتبه‌بندی BuucyBench به خروجی نزدیک‌تری نسبت به هدف رسید. این همان جایی است که مدل‌های ارزان‌تر و سریع‌تر می‌توانند برای کارهای روزمره، ساخت رابط‌ها و اتوماسیون‌های پرحجم واقعا ارزشمند باشند.

مُتَا Muse Glimmer

Muse Glimmer متا یک مدل ایجنتی با وزن‌های باز است که برخلاف مدل‌های بزرگ ابری، برای اجرای محلی طراحی شده است. یعنی به جای اینکه درخواست شما به سرورهای متا ارسال شود، می‌توانید مدل را روی سخت‌افزار خودتان اجرا و حتی آن را متناسب با نیازتان تنظیم کنید.

قابلیت‌ها

- **اجرای محلی:** کنترل بیشتری روی داده‌ها و نحوه استقرار مدل دارید.
 - **سخت‌افزار:** در حالت کامل حدود ۵۵ گیگابایت حافظه نیاز دارد.
 - **عملکرد:** با وجود اندازه کوچک‌تر، در بسیاری از آزمایش‌ها از مدل‌هایی مثل Gemma ۴ ۳۱B و Qwen ۳.۶ ۲۷B بهتر عمل کرده است.
- قرار نیست جای مدل‌های ابری بسیار بزرگ را بگیرد؛ بیشتر برای ایجنت‌های سریع، تخصصی و محلی ساخته شده است.

Nemotron ۳.۵ Lightning

Nemotron ۳.۵ Lightning انویدیا هم در همین دسته قرار می‌گیرد: یک مدل باز ۳۰ میلیارد پارامتری که هدفش اجرای سریع و کم‌هزینه وظایف تخصصی، به‌خصوص در ایجنت‌ها است.

تفاوت اصلی چیست؟

در آزمایش تولید SVG، عملکرد آن تقریباً نزدیک به Muse Glimmer بود؛ اما با هزینه‌ای حدود یک‌سوم سنت و در حدود ۴۰ ثانیه. بنابراین اگرچه این مدل‌های کوچک هنوز در بعضی کارهای بصری به پای مدل‌های بزرگ نمی‌رسند، مزیتشان جای دیگری است: «می‌توانند با هزینه کمتر و کنترل بیشتر، روی سیستم خودتان برای کارهای مشخص و تکراری اجرا شوند».

GPT-۵.۶ Sol Ultrafast و DeepSeek V۴ Pro، MAI Code

در این موج، چند مدل دیگر هم معرفی شده‌اند که هرکدام روی یک مزیت متفاوت تمرکز دارند:

- DeepSeek روی عملکرد رقابتی
- MAI Code روی کدنویسی

• OpenAI روی سرعت بسیار بالا

DeepSeek V4 Pro

DeepSeek V4 Pro نسخه نهایی مدل V4 Pro است و تمرکز اصلی آن روی انجام کارهای چندمرحله‌ای و استفاده از ابزارها است. DeepSeek می‌گوید قابلیت‌های Agent در این نسخه به‌طور قابل‌توجهی بهبود پیدا کرده‌اند و عملکرد آن در محیط‌های واقعی بهتر شده است.

چه چیزی بهتر شده است؟

ایجنت‌محورتر شده و در کارهایی که نیاز به چند مرحله، ابزار و تصمیم‌گیری پشت‌سرهم دارند، عملکرد بهتری دارد.

- برای کارهای روزمره ایجنت‌ها و محیط‌های تولیدی بهینه‌تر شده است.
- نسخه نهایی نسبت به نسخه پیش‌نمایش بهبودهای قابل‌توجهی در عملکرد Agent داشته است.
- برای برنامه‌نویسی، تحقیق و کارهای پیچیده چندمرحله‌ای مناسب‌تر است.

در واقع، V4 Pro بیشتر برای زمانی جذاب است که نمی‌خواهید فقط یک سوال بپرسید و پاسخ بگیرید؛ بلکه می‌خواهید مدل یک کار را از ابتدا تا انتها دنبال کند.

MAI-Code 1.1 Flash

MAI-Code 1.1 Flash یک مدل سبک و تخصصی برای برنامه‌نویسی است که مستقیماً در GitHub Copilot و VS Code استفاده می‌شود. هدفش این نیست بهترین مدل عمومی دنیا باشد؛ بلکه این است که برای کارهای کدنویسی، سریع و کم‌هزینه عمل کند.

چه چیزی بهتر شده است؟

- کیفیت کدنویسی بالاتر رفته است.
 - حدود ۲۵٪ بهره‌وری بیشتری از توکن‌ها دارد؛ یعنی برای انجام کار مشابه، توکن کمتری مصرف می‌کند.
 - هزینه آن حدود یک‌چهارم مدل قبلی شده است.
 - عملکرد در Terminal-Bench حدود ۲۲٪ بهتر شده است.
 - عملکرد در کارهای مرتبط با NET. حدود ۱۵٪ افزایش یافته است.
 - قابلیت دیداری هم به مدل اضافه شده است؛ بنابراین فقط متن و کد را نمی‌بیند و می‌تواند ورودی تصویری را نیز در فرایند کدنویسی در نظر بگیرد.
- به زبان ساده، Flash ۱.۱ MAI-Code برای کسی ساخته شده که می‌خواهد یک «کارمند کدنویسی سریع و ارزان» کنار دستش داشته باشد؛ نه یک مدل همه‌کاره برای همه نوع کار.

GPT-۵.۶ Sol Ultrafast

GPT-۵.۶ Sol Ultrafast در واقع یک مدل کاملاً جدید نیست؛ یک حالت اجرای سریع‌تر برای GPT-۵.۶ Sol است که با زیرساخت Cerebras ارائه می‌شود و هدف اصلی آن کاهش شدید زمان انتظار است.

چه چیزی تغییر کرده است؟

- سرعت تولید خروجی به حداکثر ۷۵۰ توکن در ثانیه می‌رسد.
 - این سرعت تا ۱۴ برابر حالت استاندارد گزارش شده است.
 - کیفیت و توانایی خود مدل قرار نیست کاهش پیدا کند؛ تغییر اصلی در زیرساخت اجرای مدل است.
- برای کارهایی مثل کدنویسی، تحقیق و برنامه‌هایی که پاسخ فوری در آن‌ها اهمیت

دارد، بسیار مناسب‌تر می‌شود.

تفاوت را می‌توان این‌طور تصور کرد: قبلاً ممکن بود از یک مدل بسیار قدرتمند بخواهید یک کار پیچیده را انجام دهد و چندین ثانیه یا بیشتر منتظر بمانید؛ Ultrafast تلاش می‌کند همان سطح از توانایی را با مکث بسیار کمتر در اختیارتان بگذارد.

آپدیت‌های سریع و کاربردی

ChatGPT و Codex برای لینوکس

اپلیکیشن دسکتاپ ChatGPT برای لینوکس در دسترس قرار گرفته و شامل ChatGPT، Codex و Cowork است. برای کاربران لینوکس که می‌خواهند محیط رسمی OpenAI را روی سیستم خود داشته باشند، این یک بهبود ساده اما کاربردی است.

Claude Code به صورت پیش‌فرض خودکارتر می‌شود

از ۱۴ آگوست ۲۰۲۶، جلسه‌های جدید Claude Code در پلن‌های Max، Pro و Team به‌طور پیش‌فرض در حالت Auto اجرا می‌شوند. یعنی برای کارهای عادی، لازم نیست هر قدم را به صورت دستی تایید کنید. Claude تنها زمانی توقف می‌کند که تشخیص دهد یک کار ممکن است خطرناک باشد.

هدف این است که یک درخواست بدهید، سیستم را رها کنید و بعد با پروژه تمام‌شده برگردید؛ نه اینکه برای تایید تک‌تک مرحله‌ها پشت سیستم بمانید.

Claude Cowork در Chrome کنترل بیشتری می‌گیرد

افزونه Chrome کلود حالا با Claude Cowork یکپارچه‌تر شده است. می‌تواند تب‌های باز را بررسی کند، از آن‌ها اسکرین‌شات بگیرد و سپس اطلاعات جمع‌شده را در Claude Cowork به یک گزارش تبدیل کند. این مدل از کار، یعنی جمع‌آوری زمینه از محیط مرورگر و بعد پردازش آن در یک محیط کاری، می‌تواند برای تحقیق و

گزارش‌نویسی واقعاً مفید باشد.

جلسه‌های مختلف Claude Code هم حالا قادرند با یکدیگر پیام ردوبدل کنند. بنابراین لازم نیست در یک جلسه جدید همه چیز را دوباره توضیح دهید. یک جلسه می‌تواند خلاصه وضعیت را برای جلسه دیگر بفرستد تا کار از همان نقطه ادامه پیدا کند.

مدل‌های ویدیویی و تصویری تازه

مدل‌های مولد تصویر و ویدیو دیگر فقط برای ساخت یک تصویر زیبا از روی متن نیستند. مدل‌های جدید بیشتر روی کنترل بهتر، حفظ ثبات شخصیت و صحنه، سرعت بالاتر و امکان اجرای محلی تمرکز دارند.

LTX ۲.۵؛ ساخت ویدیو روی کامپیوتر خودتان

LTX ۲.۵ یک مدل تولید ویدیو است؛ یعنی می‌توانید یک تصویر یا توضیح متنی به آن بدهید و از آن یک کلیپ ویدیویی بسازید. نکته مهم این است مدل با وزن‌های باز ارائه شده و امکان اجرای محلی روی سخت‌افزارهای قدرتمند را دارد. به زبان ساده، اگر یک تصویر از یک محصول دارید، می‌توانید از مدل بخواهید همان محصول را در یک صحنه متحرک نشان دهد؛ بدون اینکه برای تولید ویدیو مجبور باشید از یک سرویس ابری استفاده کنید.

Wan ۳.۰؛ ساخت کلیپ‌های طولانی‌تر

Wan ۳.۰ از Alibaba هم یک مدل تولید ویدیو است که برای ساخت کلیپ‌های نسبتاً طولانی طراحی شده است. مزیت اصلی آن این است می‌تواند صحنه‌های طولانی‌تر و پیچیده‌تری را نسبت به نسل‌های قبلی تولید کند.

برای مثال، به جای اینکه برای یک تبلیغ ۳۰ ثانیه‌ای چند ویدیوی کوتاه بسازید و آن‌ها را به هم بچسبانید، می‌توانید یک کلیپ طولانی‌تر را مستقیماً از مدل بگیرید. این نشان می‌دهد رقابت مدل‌های ویدیویی به سمت ویدیوهای طولانی‌تر و

قابل استفاده تر پیش می‌رود.

۲.۵-MAI-Image؛ فقط ساخت تصویر نیست

۲.۵-MAI-Image از شرکت مایکروسافت، یک مدل تولید و ویرایش تصویر است. یعنی هم می‌توانید بگویید «یک تصویر جدید بساز» و هم یک تصویر موجود را بدهید و دقیقاً مشخص کنید چه چیزی باید تغییر کند.

مثلاً می‌توانید یک عکس محصول را بدهید و بگویید پس‌زمینه را عوض کن، یک شیء را حذف کن یا رنگ بخشی از تصویر را تغییر بده؛ بدون اینکه کل تصویر از نو ساخته شود. مدل همچنین تلاش می‌کند چهره، هویت شخصیت، ترکیب‌بندی و نورپردازی اصلی را هنگام ویرایش حفظ کند.

پیشرفت اصلی ۲.۵-MAI چیست؟

یکی از پیشرفت‌های مهم این نسخه، درک بهتر متن داخل تصویر، تصاویر محصول و رعایت دقیق‌تر دستور کاربر است. همچنین در ویرایش‌های موضعی، مثلاً حذف یک اشیا یا تغییر یک بخش از عکس، کنترل بیشتری دارد.

این مدل در آزمون‌های انسانی Arena هم عملکرد بالایی داشته و در حال حاضر رتبه سوم Text-to-Image و رتبه دوم Image Editing را دارد. نکته مهم اینجاست رتبه‌بندی Arena، بر اساس انتخاب انسان‌ها بین خروجی‌های ناشناس انجام می‌شود؛ بنابراین این رتبه‌ها تا حدی نشان می‌دهند کدام تصاویر از نظر کاربران جذاب‌تر و مناسب‌تر بوده‌اند.

همه تغییرات چشمگیر نیستند؛ اما روند روشن است!

میان همه خبرها و معرفی‌های این هفته، چند تغییر مهم بیشتر از بقیه به چشم می‌آید. شاید هر مدل جدید یک انقلاب نباشد؛ اما وقتی این تغییرات را کنار هم می‌گذاریم، مسیر حرکت هوش مصنوعی کاملاً مشخص‌تر می‌شود:

تولید تصویر و ویدیو: مدل‌ها به سمت ساخت محتوای طولانی‌تر و منسجم‌تر

می‌روند.

تمرکز روی توسعه ایجنت‌ها به‌جای چت‌بات: حالا هوش مصنوعی باید بتواند خودش کار را دست بگیرد.

گزینه‌های اقتصادی‌تر: چون برای بسیاری از کارهای روزمره، نیازی به گران‌ترین و قدرتمندترین مدل نداریم.

مهم این است با هر آپدیت جدید، درگیر هیجان نشویم. هر مدل تازه‌ای قرار نیست دنیا را تغییر دهد؛ اما وقتی روندها را کنار هم می‌گذاریم، تصویر واضحی شکل می‌گیرد:

[message_box text_color="light]

- هوش مصنوعی از یک ابزار برای «تولید پاسخ» به‌سمت ابزاری برای «انجام کار» حرکت می‌کند.

- تولید محتوای تصویری و ویدیویی هر روز واقعی‌تر و کاربردی‌تر می‌شود.

- استفاده از مدل‌های قدرتمند ارزان‌تر شده و شفافیت و تشخیص اینکه محتوا توسط هوش مصنوعی ساخته شده یا انسان، اهمیت بیشتری پیدا می‌کند.

[message_box/]