اخبار

حمایت مرکز ملی فضای مجازی از کسب و کارها

معاون حقوقی و امور مجلس مرکز ملی فضای مجازی با تاکید بر اینکه رسالت مرکز ملی فضای مجازی حمایت از کسب و کارهای فضای مجازی است، افزود: امیدواریم این تفاهم‌نامه مقدمه‌ای برای حمایت جدی و رفع چالش‌های آنها در این حوزه باشد. به گزارش ایسنا از مرکز ملی فضای مجازی، محمدصادق فراهانی، معاون حقوقی و امور مجلس مرکز ملی فضای مجازی در مراسم امضای تفاهمنامه همکاری مرکز ملی فضای مجازی و پارک علم و فناوری دانشگاه صنعتی شریف گفت: اقتصاد دیجیتال یکی از سه حوزه کلیدی و مهم برای رشد و توسعه کشور است که البته این حوزه نوپا در ایران علی رغم تاکیدات مسوولان عالی کشور، با چالش‌های بسیار همراه بوده است. وی با تاکید بر اینکه رسالت مرکز ملی فضای مجازی حمایت از کسب و کارهای فضای مجازی است، افزود: امیدواریم این تفاهمنامه مقدمه‌ای برای حمایت جدی و رفع چالش‌های آنها در این حوزه باشد و به رشد و توسعه قانونمند اقتصاد دیجیتال برای بهره مندی از حقوق و آزادی‌های مشروع خود کمک کند. فراهانی تصریح کرد: این تفاهمنامه در دو حوزه پیشینی و پسینی منعقد شده که مشاوره‌های حقوقی از جمله اقدامات حوزه پیشینی و رفع چالش‌های حقوقی در بخش اقدامات پسینی در نظر گرفته شده است که برای تحقق این اقدامات از ظرفیت‌های قانونی شورای عالی و مرکز ملی فضای مجازی همچون کمیته تسهیل اقتصاد دیجیتال استفاده خواهیم کرد. در ادامه این مراسم بهنام طالبی، رییس پارک علم و فناوری دانشگاه صنعتی شریف ضمن تشریح فعالیت شرکت‌های فعال در این پارک، گفت: کسب و کارهای فعال در اقتصاد دیجیتال از سوی نهادهای دولتی برای اعمال تکالیف سنتی شان در تقابل هستند. این در حالی است که طبق قانون شرکت‌های عضو پارک‌های علم و فناوری از حمایت‌ها و معافیت‌های مختلفی برای پیشرفت و تحول اقتصاد کشور برخوردار خواهند بود. وی افزود: شرکت‌های فعال در این حوزه، توانایی رقابت در سطح جهانی را به دلیل سرمایه انسانی نخبه و بهره مندی از ابزار مورد نیاز داشته و بازار بکر و جوانی در این حوزه وجود دارد. بنابراین ضروری است که به رشد و توسعه این شرکت‌ها کمک شود و بر همین اساس، این تفاهمنامه زمینه تقویت این حوزه را فراهم خواهد کرد. معاونت حقوقی مرکز ملی فضای مجازی و پارک علم و فناوری دانشگاه صنعتی شریف تفاهمنامه همکاری را در راستای تحقق سیاست‌های کلان کشور در حمایت از توسعه اقتصاد دانش‌بنیان و ارتقای ظرفیت‌های فناورانه امضاء کردند.

بازنگری در سند شبکه ملی اطلاعات

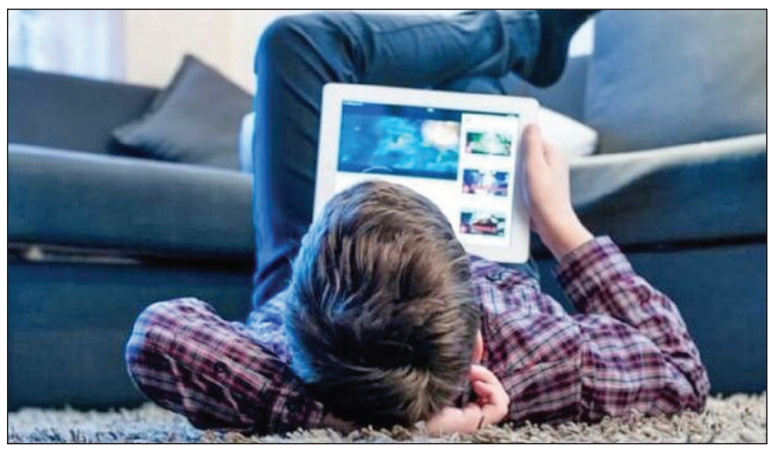
وزیر ارتباطات گفت: بازنگری سند شبکه ملی اطلاعات از تکالیف مرکز ملی فضای مجازی است و وزارت ارتباطات نیز برای همکاری در این زمینه اعلام آمادگی کرده و پیگیر انجام هرچه سریع‌تر این بازنگری است. به گزارش مهر به نقل از وزارت ارتباطات و فناوری اطلاعات، سید ستار هاشمی، وزیر ارتباطات و فناوری اطلاعات، در یکصد و چهل و پنجمین جلسه شورای راهبری شبکه ملی اطلاعات که با موضوع تولید محتوا در شبکه ملی اطلاعات برگزار شد، با اشاره به تکالیف دستگاه‌های مختلف بر اساس سند شبکه ملی اطلاعات در حوزه تولید محتوا، بر ضرورت ارزیابی اقدامات انجام‌شده و میزان موفقیت دستگاه‌های متولی در این زمینه تأکید کرد. وی با اشاره به اقدامات انجام‌شده در یک دهه اخیر در حوزه تولید محتوا، افزود: وزارت ارتباطات به صورت راهبردی از فعالیت بخش خصوصی در این حوزه حمایت می کند و معتقد است ظرفیت‌های بخش خصوصی باید در سیاست گذاری و تصمیم گیری‌های این حوزه مورد توجه قرار گیرد.

سیاست‌گذاری حوزه کودک با مشارکت فعالان این عرصه

وزیر ارتباطات با اشاره به فعالیت پلتفرم‌های مختلف در زمینه تولید محتوای کودک در سال‌های اخیر، گفت: برای تصمیم گیری دقیق‌تر در این حوزه، باید از تجربیات فعالان و پلتفرم‌هایی که در میدان عمل با مسائل و چالش‌های تولید محتوای کودک مواجه بوده‌اند، استفاده کرد و نظر آنها را در فرآیند سیاست گذاری مورد توجه قرار داد. هاشمی با اشاره به تجربه یکی از کشورهای اروپایی در حوزه کودکان در فضای مجازی، افزود: حتی موفق‌ترین پلتفرم‌های فعال در این حوزه نیز با محدودیت‌ها و چالش‌هایی مواجه هستند و میزان موفقیت آنها نمی‌تواند صددرصدی باشد. وی، فاصله دانش دیجیتال کودکان و والدین را یکی از پیچیدگی‌های مهم این حوزه دانست و تأکید کرد: این موضوع باید در سیاست گذاری‌های مرتبط با حضور کودکان در فضای مجازی مورد توجه قرار گیرد. هاشمی همچنین بر ضرورت بازنگری سند شبکه ملی اطلاعات متناسب با برنامه هفتم پیشرفت و الزامات روز حوزه ارتباطات و فناوری اطلاعات تأکید کرد و گفت: این موضوع از تکالیف مرکز ملی فضای مجازی است و وزارت ارتباطات نیز برای همکاری در این زمینه اعلام آمادگی کرده و پیگیر انجام هرچه سریع‌تر این بازنگری است؛ چرا که بخش قابل توجهی از مطالب و موضوعات مطرح‌شده در این سند نیازمند به‌روزرسانی است. وی در ادامه با پیشنهاد استفاده از ظرفیت صندوق نوآفرین برای ارایه تسهیلات به فعالان حوزه تولید محتوای کودک در فضای مجازی موافقت کرد. بر اساس این گزارش، در ابتدای این جلسه، مسوول کمیته محتوای شبکه ملی اطلاعات گزارشی از اقدامات یک‌ساله این کمیته در چارچوب اهداف عملیاتی و اقدامات طرح کلان و معماری شبکه ملی اطلاعات در دو بخش «تولید محتوا در شبکه ملی اطلاعات» و «سند صیانت از کودکان در فضای مجازی» ارایه کرد. در ادامه، حاضران در جلسه ضمن بررسی موضوعات مطرح‌شده در این گزارش، دیدگاه‌ها، پیشنهادها و راهکارهای خود را برای تحقق اهداف پیش‌بینی‌شده در حوزه تولید محتوا و صیانت از کودکان در فضای مجازی ارایه کردند.

گزارش مرکز پژوهش‌های مجلس از وضعیت خدمات دیجیتال کودک و نوجوان در ایران

زیرساخت‌های پراکنده برای یک کاربر کوچک



ایران حالا برای کودکانش تقریباً از هر چیزی یک سرویس دیجیتال دارد؛ از ویدئو و آموزش و بازی گرفته تا پرداخت و کنترل والدین. اما این سرویس‌ها هنوز بیشتر شبیه جزیره‌هایی جدا از هم هستند تا اجزای یک اکوسیستم. گزارش مرکز پژوهش‌های مجلس نشان می‌دهد چالش اصلی محیط دیجیتال کودک دیگر صرفاً کمبود خدمات نیست، بلکه نبود زیرساخت‌های مشترک و استانداردهایی است که این خدمات را به یکدیگر متصل کند. کودک ایرانی برای حضور در فضای دیجیتال، امروز با کمبود گزینه مواجه نیست. اگر خانواده‌ای به دنبال محتوای ویدئویی، آموزش آنلاین، بازی، پیام‌رسان، ابزارهای کنترل والدین یا حتی خدمات مالی متناسب با کودک باشد، نمونه‌های مختلفی در بازار ایران پیدا می‌کند. با این حال، تصویری که در مجموع از این بازار به دست می‌آید، بیشتر شبیه مجموعه‌ای از جزایر جدا از هم است تا یک اکوسیستم منسجم. هر سرویس بخشی از نیاز کودک را پوشش می‌دهد، اما این سرویس‌ها لزوماً یکدیگر را نمی‌شناسند، با یک استاندارد مشترک کار نمی‌کنند و نمی‌توانند از زیرساخت‌های یکسان استفاده کنند. گزارش مرکز پژوهش‌های مجلس درباره مدل‌های کسب‌وکار دیجیتال کودک و نوجوان نیز دقیقاً بر همین شکاف دست می‌گذارد. بر اساس این گزارش، اگرچه طی سال‌های گذشته بخش‌های مختلف محیط دیجیتال کودک در ایران شکل گرفته‌اند، نبود زیرساخت‌های عمومی و تعامل‌پذیر مانع از آن شده است که این مجموعه در قالب یک محیط واحد و هماهنگ فعالیت کند. این مساله در شرایطی اهمیت بیشتری پیدا می‌کند که ایجاد شبکه‌ای اختصاصی، ایمن، سالم و مفید برای کودکان و نوجوانان، در اسناد بالادستی مرتبط با توسعه شبکه ملی اطلاعات نیز مورد توجه قرار گرفته است. بنابراین مساله فقط توسعه چند سکوی جدید نیست؛ بلکه ساختن معماری‌ای است که در آن سرویس‌های مختلف بتوانند در کنار یکدیگر کار کنند و خانواده برای استفاده از هر خدمت، مجبور به عبور از یک مسیر کاملاً جداگانه نباشد.

بازار کودک از محتوا جلوتر رفته زیرساخت عقب مانده است

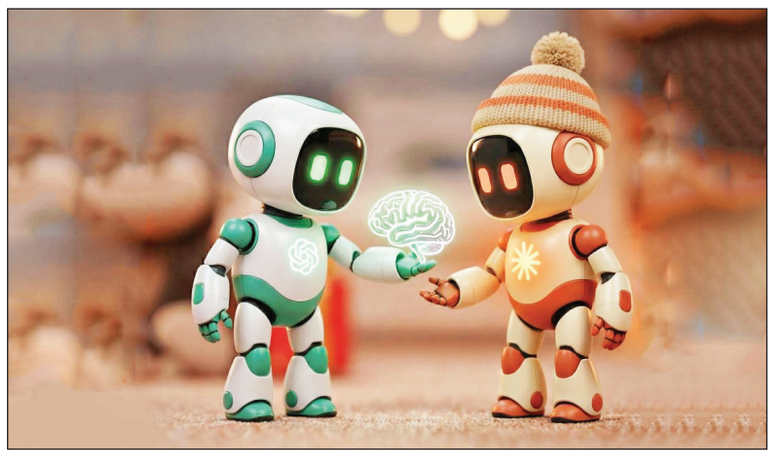
اگر محیط دیجیتال کودک را به چند بخش تقسیم کنیم، یکی از نقاطی که در ایران رشد بیشتری داشته، لایه خدمات نهایی و محتوایی است. در این بخش، بازار نسبتاً متنوعی شکل گرفته و بازیگران مختلف تلاش کرده‌اند نیازهای سرگرمی، آموزشی و ارتباطی کودکان را پاسخ دهند. بازار ویدئوی آنلاین نمونه روشنی از این روند است. برخی سکوها اساساً با تمرکز بر مخاطب کودک فعالیت می‌کنند و برخی پلتفرم‌های عمومی نیز بخش ویژه‌ای را به این گروه اختصاص داده‌اند. در کنار آنها، سرویس‌هایی مانند آپارات کودک امکان دسترسی و جست‌وجوی محتوای ویدئویی متناسب با کودکان را فراهم کرده‌اند. در بعضی از این خدمات نیز ابزارهایی برای مدیریت دسترسی و نظارت والدین در نظر گرفته شده است. با این حال، توسعه محتوا و سرویس‌های نهایی به معنای حل مساله محیط دیجیتال کودک نیست. اتفاقاً هر چه تعداد این سرویس‌ها بیشتر می‌شود، نیاز به زیرساخت‌های مشترک نیز بیشتر احساس می‌شود. کودکی که از یک سکوی آموزشی استفاده می‌کند، ممکن است همزمان در یک پلتفرم ویدئویی، یک پیام‌رسان و یک بازی آنلاین نیز حضور داشته باشد. اگر هر کدام از این خدمات سازوکار مستقلی برای تشخیص سن، احراز هویت، پرداخت و کنترل والدین داشته باشند، تجربه کاربر پراکنده خواهد شد و هزینه توسعه برای کسب‌وکارها نیز افزایش پیدا می‌کند. گزارش مرکز پژوهش‌ها برای حل این مساله، لایه‌ای تحت عنوان خدمات پایه را برجسته می‌کند؛ خدماتی که قرار نیست مستقیماً محصول نهایی کودک باشند، بلکه باید به عنوان زیرساخت مشترک در اختیار بازیگران مختلف قرار گیرند. احراز هویت یکی از مهم‌ترین این خدمات است. ایده مطرح‌شده این است که سکوهای مختلف به جای طراحی سیستم‌های جداگانه، بتوانند به یک زیرساخت استاندارد متصل شوند. در چنین ساختاری، سن کودک می‌تواند با استفاده از هویت دیجیتال والدو با دریافت رضایت او بررسی شود، بدون اینکه لازم باشد اطلاعات هویتی کودک در یک پایگاه متمرکز ذخیره شود. پرداخت نیز وضعیت مشابهی دارد. حساب کودک می‌تواند به حساب والد متصل باشد، اما با محدودیت‌های مشخص برای میزان و نوع تراکنش. در این صورت والد نه فقط امکان پرداخت، بلکه امکان نظارت و مدیریت مالی فعالیت کودک را نیز خواهد داشت. کنترل والدین و فهرست‌های سیاه و سفید هم باید بخشی از همین زیرساخت مشترک باشند. فضای آنلاین دایماً در حال تغییر است و یک فهرست ثابت نمی‌تواند برای مدت طولانی پاسخگوی نیازهای ایمنی باشد. بنابراین این ابزارها باید امکان به‌روزرسانی سریع و حتی خودکار داشته باشند. ایران برای بسیاری از این نیازها، نمونه‌هایی دارد؛ از سرویس‌های مالی ویژه کودک گرفته تا ابزارهای کنترل والدین و خدماتی که اپراتورها برای مدیریت دسترسی کودکان ارایه کرده‌اند. مساله اما این است که این خدمات معمولاً در محدوده اکوسیستم ارایه‌دهنده خود باقی مانده‌اند. هنوز استاندارد مشترکی وجود ندارد که یک سکوی آموزشی، یک پلتفرم ویدئویی و یک سرویس بازی بتوانند همگی به یک سیستم احراز هویت یا کنترل والدین متصل شوند. در بخش ارتباطات نیز همین الگو تکرار شده است. سیم کارت‌ها و خدمات ارتباطی ویژه کودکان توسعه پیدا کرده‌اند، اما پیوند آنها با خدمات پایه و سکوهای محتوایی کامل نیست. از سوی دیگر، در حوزه تجهیزات نیز محصولاتی مانند تبلت‌های ویژه کودک تولید شده‌اند، ولی بخش اصلی بازار همچنان در اختیار گوشی‌ها و تبلت‌های عمومی است؛ تجهیزاتی که لزوماً برای نیازهای صیانتی و الزامات محیط کودک طراحی نشده‌اند. در کنار این چهار لایه، شکل گیری یک اکوسیستم پایدار به سه ضلع دیگر نیز نیاز دارد: کاربران و بازار، نهادهای تنظیم‌گر و سیاست‌گذار و منابع مالی و حمایتی. خانواده، مدرسه و خود کودک تقاضا را ایجاد می‌کنند؛ نهادهای حاکمیتی باید قواعد و استانداردهای بازی را مشخص کنند و سرمایه‌گذاران، شتاب‌دهنده‌ها و صندوق‌های نوآوری نیز امکان رشد کسب‌وکارها را فراهم می‌کنند. نبود هر کدام از این اجزا می‌تواند توسعه سایر بخش‌ها را هم محدود کند.

چهار راه برای خروج از جزیره‌ای بودن

مساله اصلی در نهایت به این پرسش می‌رسد که محیط دیجیتال کودک باید چگونه ساز مانده‌ی شود. آیا همه خدمات باید در یک سکوی واحد جمع شوند یا بهتر است سرویس‌های مستقل، از طریق یک زیرساخت مشترک به یکدیگر متصل شوند؟ مرکز پژوهش‌های مجلس چهار الگوی کلی را برای پاسخ به این پرسش بررسی کرده است. نخستین گزینه، ابرسکو است؛ مدلی که در آن کاربر از یک نقطه وارد مجموعه بزرگی از خدمات می‌شود. وی چت‌نمونه معروف چنین ساختاری در چین است که خدمات ارتباطی، پرداخت، تجارت، سلامت و حتی خدمات عمومی را در یک محیط واحد کنار هم قرار داده است. با این حال، نمونه‌ای که همین مدل را به صورت کامل و اختصاصی برای کودکان اجرا کرده باشد، در جهان چندان شناخته‌شده نیست. گزینه دوم، ایجاد یک سکوی متمرکز است. در این مدل یک بازیگر اصلی، مجموعه‌ای از خدمات را تحت یک سیاست و استاندارد واحد مدیریت می‌کند. خدمات کودک تنسنت در چین به این مدل نزدیک است؛ مجموعه‌ای که از محتوا و آموزش تا بازی، ساعت هوشمند، احراز سن و محدودیت زمانی را در یک چارچوب مشترک ارایه می‌دهد. اما راه سوم، شاید از نظر معماری برای بازاری مانند ایران انعطاف بیشتری داشته باشد: ایجاد یک هاب یا درگاه یکپارچه. در این مدل قرار نیست همه کسب‌وکارها زیر مالکیت یک مجموعه قرار بگیرند. هر سرویس می‌تواند مستقل باقی بماند، اما از طریق رابط‌های استاندارد به خدمات پایه مشترک وصل شود. به این ترتیب، یک کسب‌وکار کوچک هم می‌تواند بدون سرمایه‌گذاری سنگین برای ساخت سیستم احراز هویت یا پرداخت اختصاصی، از زیرساخت موجود استفاده کند. تجربه هند در این زمینه قابل توجه است. «دیکشا» در حوزه آموزش مدرسه‌ای، بستری فراهم کرده که بازیگران مختلف می‌توانند خدمات خود را روی یک زیرساخت مشترک ارایه کنند. در مقیاس وسیع‌تر نیز «هنداستک» نشان می‌دهد چگونه می‌توان با ایجاد زیرساخت‌های عمومی و رابط‌های برنامه‌نویسی، ارتباط میان دولت، کسب‌وکار و کاربر را بدون تبدیل همه آنها به یک پلتفرم واحد برقرار کرد. مدل چهارم نیز به جای ادغام خدمات، روی انتخاب و عرضه ایمن آنها تمرکز دارد. در این رویکرد، یک فروشگاه یا بازارگاه تخصصی، برنامه‌ها و سرویس‌ها را از نظر ایمنی، کیفیت و تناسب سنی بررسی می‌کند و سپس آنها را در اختیار خانواده‌ها قرار می‌دهد. مزیت این مدل آن است که ارایه‌دهندگان همچنان استقلال خود را حفظ می‌کنند، اما کاربر از یک مسیر قابل اعتماد به خدمات دسترسی پیدا می‌کند. در نتیجه، انتخاب مدل مناسب برای ایران الزاماً به معنای انتخاب یکی از این چهار گزینه به شکل خالص نیست. حتی ممکن است ترکیبی از آنها بتواند پاسخ بهتری ایجاد کند. آنچه اهمیت دارد، تعیین معماری‌ای است که میان امنیت کودک، حریم خصوصی، آزادی عمل کسب‌وکارها، رقابت و سهولت استفاده تعادل برقرار کند. در واقع، تجربه فعلی ایران نشان می‌دهد مرحله اول توسعه خدمات دیجیتال کودک تا حد زیادی طی شده است: سرویس‌ها یکی پس از دیگری ایجاد شده‌اند. چالش مرحله بعد، ساختن «اتصال» میان آنهاست. اگر احراز هویت، پرداخت، کنترل والدین، مدیریت دسترسی و سایر خدمات پایه به زیرساخت‌های مشترک تبدیل شوند، کسب‌وکارها می‌توانند روی تولید محتوا و ارایه خدمات بهتر تمرکز کنند و خانواده نیز به جای مدیریت چندین محیط جداگانه، با یک تجربه منسجم‌تر مواجه خواهد شد. به این ترتیب، مساله آینده بازار دیجیتال کودک در ایران احتمالاً کمتر به این مربوط خواهد بود که «چه سرویس جدیدی بسازیم» و بیشتر به این مربوط است که «چطور سرویس‌های موجود را به یک اکوسیستم تبدیل کنیم». فاصله میان این دو، همان شکافی است که امروز مانع شکل گیری یک محیط دیجیتال یکپارچه برای کودک و نوجوان شده است.

رتبه‌بندی مدل‌های هوش مصنوعی

برتری کلود در متن و اوپن‌ای‌آی در تصویر



رقابت مدل‌های هوش مصنوعی در سال ۲۰۲۶ به مرحله‌ای رسیده است که پرسش سنتی «بهترین مدل کدام است؟» بخش بزرگی از معنای خود را از دست داده است. به گزارش مهر؛ داده‌های تازه وبگاه «Arena.ai» تصویری چندقطبی از مرز فناوری ترسیم می‌کنند. بر همین اساس، آنتروپیک در متن، کدنویسی و عامل‌های هوشمند موقعیتی برجسته دارد، اوپن‌ای‌آی در تولید و ویرایش تصویر فاصله معناداری با رقبا ایجاد کرده، گوگل حوزه تولید ویدئو را در اختیار گرفته و بازیگران چینی مانند مون‌شات، علی‌بابا، «Z.ai»، بایت‌دنس و مینی‌مکس در بخش‌هایی از مرز عملکرد، مستقیماً با آزمایشگاه‌های امریکایی رقابت می‌کنند. اهمیت بنچمارک «Arena ai» از روش ارزیابی آن ناشی می‌شود. رتبه‌بندی‌های اصلی این پلتفرم بر مقایسه‌های دوبه‌دوی پاسخ مدل‌ها و ترجیح کاربران واقعی بنا شده‌اند و امتیازها با مدل آماری مدل‌بردلی-تری محاسبه می‌شوند. «Arena.ai» در کنار رتبه خام، «بازه رتبه» را نیز منتشر می‌کند تا عدم قطعیت آماری آشکار بماند. بنابراین تفاوت چند امتیازی میان دو مدل الزاماً به معنای برتری قطعی یکی بر دیگری نیست. «Arena.ai» در سال ۲۰۲۶ شاخص «factuality» را نیز به «Text» و «Search» افزود که با استخراج و راستی‌آزمایی ادعاهای قابل بررسی، صحت پاسخ را در کنار ترجیح انسانی اندازه‌گیری می‌کند.

سلطه آنتروپیک بر حوزه متن

در مهم‌ترین جدول عمومی «Arena.ai»، یعنی «Text Arena»، داده‌های ۲۱ اوت ۲۰۲۶ که بر بیش از ۷.۹ میلیون رأی و ۳۹۴ مدل استوار است، موقعیت بسیار قدرتمندی برای آنتروپیک نشان می‌دهد. بر همین اساس، «Claude Fable ۵» با امتیاز ۱۵۰۸ در رتبه نخست قرار دارد و «Claude Opus ۴.۶ High» با ۱۵۰۴ و «Claude Opus ۴.۷ High» با ۱۵۰۲ جایگاه‌های دوم و سوم را به خود اختصاص داده‌اند. پس از آنها «Muse Spark ۱.۲ xHigh» متعلق به متا با امتیاز ۱۴۹۸ قرار دارد. نکته تعیین کننده، تراکم خانواده «Claude» در صدر جدول است. چهار مدل آنتروپیک در میان شش رتبه نخست حضور دارند. چنین آرایشی نشان می‌دهد برتری آنتروپیک در حال حاضر بیشتر به یک «خانواده توانمندی» شباهت دارد تا موفقیت یک مدل منفرد. همین الگو در زبان انگلیسی نیز تکرار می‌شود و «Claude Fable ۵» با امتیاز ۱۵۱۵ صدرنشین است. در مقابل، بهترین مدل اوپن‌ای‌آی در داده ۱۹ اوت، «Sol xHigh ۵.۶-GPT» با امتیاز ۱۴۸۲ در رتبه ۱۸ جدول کلی متن قرار گرفته بود. فاصله رتبه، همزمان با هم‌پوشانی بعضی بازه‌های اطمینان، نشان می‌دهد که ترجیح کاربران «Arena.ai» در تعاملات عمومی متنی اکنون به شکل محسوسی به سمت «Claude» متمایل شده است.

پیشتازی Claude و مدل‌های چینی در حوزه کدنویسی

برتری آنتروپیک در حوزه کدنویسی حتی واضح‌تر از بخش تولید متن دیده می‌شود. بر همین اساس، در زیرشاخه «Coding»، سه مدل «Claude Opus ۴.۶ High» و «Claude Opus ۴.۷ High» و «Claude Fable ۵» هر سه امتیاز ۱۵۵۲ کسب کرده‌اند و رتبه‌های نخست تا سوم را در اختیار دارند. همچنین، مدل «Kimi K۳ Max»، متعلق به شرکت مون‌شات نیز با امتیاز ۱۵۴۴ در میان مدعیان اصلی قرار گرفته است. تصویر کلی در بخش «Code Arena WebDev» نیز مشابه است، هرچند قدرت بازیگران آسیایی در این بخش آشکارتر می‌شود. بر پایه داده‌های ۲۱ اوت و بیش از ۶۰۳ هزار رأی، «Claude Max ۵ Opus» با امتیاز ۱۶۹۱ در جایگاه نخست، «Max Kimi K۳» با ۱۶۷۴ در رتبه دوم و «Qwen ۳.۸ Max» متعلق به علی‌بابا با ۱۶۶۹ در رتبه سوم قرار دارند. علاوه بر این، «Sol xHigh ۵.۶-GPT» نیز با سازوکار کدنویسی اختصاصی خود رتبه هفتم و امتیاز ۱۶۱۹ را به دست آورده است. این بخش یکی از مهم‌ترین نشانه‌های تغییر ساختار صنعت است. بر همین مبنا، به نظر می‌رسد مون‌شات و علی‌بابا دیگر در سطح مدل‌های «کم‌هزینه جایگزین» ظاهر نمی‌شوند و در برخی وظایف توسعه نرم‌افزار مستقیماً در مرز عملکرد قرار گرفته‌اند. در طراحی مبتنی بر مرجع نیز «Max Kimi K۳» با امتیاز ۱۷۲۱ حتی بالاتر از «Max ۵ Claude Opus» ایستاده است.

هوش چندوجهی مدل‌های باز در تعقیب Claude

در بخش «Vision Arena» که توانایی مدل‌ها برای تحلیل و استدلال روی ورودی‌های بصری را می‌سنجد، مدل «Claude Fable ۵» با امتیاز ۱۳۱۵ رتبه نخست را دارد. مدل‌های «Qwen ۳.۸ Max» و «Claude High ۴.۷ Opus» هر دو با امتیاز ۱۳۰۱ در رتبه‌های بعدی قرار گرفته‌اند. این جدول بر بیش از ۱.۱۸ میلیون رأی و ۱۴۶ مدل استوار است. بخش متن‌باز این حوزه نیز اهمیت راهبردی دارد. در این بخش، «Kimi K۲.۶» با امتیاز ۱۲۶۳ بهترین مدل دارای مجوز باز در «Vision Arena» است و پس از آن «Gemma ۳۱B ۴» گوگل، «Qwen ۳.۵» و «Thinking Kimi K۲.۵» قرار دارند. فاصله موجود با مدل‌های اختصاصی همچنان محسوس است، اما تراکم مدل‌های متن‌باز آسیایی در این رده نشان می‌دهد شکاف قابلیت به تدریج در حال کاهش است.

حوزه تولید تصویر قلمرو سلطه مدل اوپن‌ای‌آی

بزرگ‌ترین برتری اختصاصی یک شرکت را باید در حوزه تصویر جست‌وجو کرد. بر همین اساس، در بخش «Text-to-Image Arena»، مدل «GPT-Image-۲ Medium» با امتیاز ۱۳۸۱ در رتبه نخست قرار دارد و «Microsoft MAI-Image-۲.۶ Preview» با ۱۳۳۶ و «Grok Imagine Image ۲.۰» با ۱۳۱۶ در رتبه‌های بعدی ایستاده‌اند. فاصله ۴۵ امتیازی میان رتبه اول و دوم، همراه با بازه اطمینان محدود «GPT-Image-۲»، برتری نسبتاً قدرتمند شرکت اوپن‌ای‌آی را نشان می‌دهد. این رتبه‌بندی بر نزدیک به ۵.۹۲ میلیون رأی بنا شده است. در ادامه، مزیت اوپن‌ای‌آی در ویرایش تصویر برجسته‌تر می‌شود. مدل «GPT-Image-۲ Medium» با امتیاز ۱۴۶۳ و حدود ۱۹۹ هزار رأی در صدر جدول ۱۷ اوت قرار دارد. «Grok Imagine Image ۲.۰» با ۱۴۳۹ و «MAI-Image-۲.۶ Preview» با ۱۴۲۰ در جایگاه‌های دوم و سوم هستند. حجم کل رأی‌های بخش «Image Edit Arena» نیز از ۲۹ میلیون عبور کرده است که این حوزه را به یکی از بزرگ‌ترین مجموعه‌های ترجیح انسانی «Arena.ai» تبدیل می‌کند. بر همین مبنا، چنین به نظر می‌رسد که اوپن‌ای‌آی در رتبه‌بندی عمومی متن فعلاً صدر را به آنتروپیک واگذار کرده و در لایه تولید و ویرایش تصویر مزیت تخصصی بسیار قدرتمندی ساخته است. معماری رقابت از همین نقطه چندبعدی می‌شود. جایگاه یک شرکت را دیگر نمی‌توان با یک مدل زبانی واحد اندازه گرفت.

گوگل در صدر حوزه تولید ویدئو

بررسی داده‌های «Arena.ai» نشان می‌دهد که بازار تولید ویدئو ساختار متفاوتی دارد. در داده‌های ۱۴ اوت بخش «Text-to-Video Arena»، مدل «Gemini Omni Flash» گوگل با امتیاز ۱۵۱۲ صدرنشین است. همچنین، «Video ۳ Flux» متعلق به «Black Forest Labs» با ۱۴۹۴ و «Dreamina Seedance ۲.۰» از بایت‌دنس با ۱۴۸۲ در رتبه‌های دوم و سوم قرار دارند. مدل «Muse Video»، متعلق به متا و مینی‌مکس «H۳» نیز رتبه‌های بعدی را اشغال کرده‌اند، در حالی که «Sora Pro ۲» متعلق به اوپن‌ای‌آی با امتیاز ۱۳۶۴ در رتبه هشتم قرار گرفته است. از سوی دیگر، جدول ویرایش ویدئو «Arena.ai» تصویر دیگری ارایه می‌دهد. بر همین اساس، مدل «Dreamina Seedance ۲.۵» متعلق به بایت‌دنس با امتیاز ۱۴۱۱ در صدر قرار دارد و مینی‌مکس «H۳» با ۱۳۸۸ و «Gemini Omni Flash» با ۱۳۵۸ در تعقیب آن هستند. بنابراین زنجیره ارزش ویدئوی مولد میان چند بازیگر توزیع شده و هیچ آزمایشگاهی بر تولید و ویرایش به‌طور هم‌زمان سلطه ندارد.

عصر عامل‌های هوش مصنوعی کیفیت مدل‌ها در انجام کار

به عقیده کارشناسان، مهم‌ترین تحول نسل جدید ارزیابی شاید «Agent Arena» باشد. در این بخش، «Arena ai» به جای رأی دوبه‌دو از روش «causal tracing» استفاده می‌کند و عملکرد عامل را بر مبنای موفقیت واقعی در انجام وظیفه، قابلیت هدایت، بازیابی پس از خطای ابزار و توهم در استفاده از ابزارها می‌سنجد. هدف، اندازه‌گیری سهم مدل هماهنگ کننده در یک سامانه چندجزئی است. در داده‌های ۱۹ اوت و بیش از ۱.۹ میلیون سشن پاسخ از مدل‌های مختلف هوش مصنوعی، مدل «Claude Opus ۵ High» رتبه نخست را با روند اصلاح و بهبود حدود ۱۲.۴۷ درصد در اختیار دارد و نسخه «Max» و «Claude Fable ۵ High» پس از آن قرار گرفته‌اند. اگر رتبه‌بندی در سطح آزمایشگاه بررسی شود، آنتروپیک نخست، مون‌شات با «Max Kimi K۳» دوم و اوپن‌ای‌آی با «Sol xHigh ۵.۶-GPT» سوم است. همچنین، مدل «Kimi» در این میان از منظر اقتصاد استفاده نیز جالب توجه است و هزینه میانه آن حدود ۰.۶۵ دلار برای هر وظیفه گزارش شده، در حالی که هزینه مدل پرچم‌دار آنتروپیک بیش از دو دلار برای هر وظیفه است.

جست‌وجوی وب؛ اهمیت «دسترسی به اطلاعات» در کنار قدرت مدل

در بخش «Search Arena»، مدل «Claude Opus ۴.۶ Search» با امتیاز ۱۲۵۳ رتبه نخست را دارد و «GPT-۵.۵-Search» با ۱۲۴۰ دوم است. علاوه بر این، مدل «Claude Fable ۵» و «Claude Opus ۴.۷» و «Ernie ۵.۱» از شرکت بایدو جایگاه‌های بعدی را در اختیار دارند. این جدول بر حدود ۹۴۰ هزار رأی استوار است و مشخصاً مدل‌های دارای جست‌وجوی وب یکپارچه را ارزیابی می‌کند. اهمیت این بخش فراتر از رتبه‌بندی است. «Arena.ai» گزارش کرده که پاسخ‌های بخش «Search Arena» به‌طور متوسط نزدیک به ۱۰ ادعای واقعی دارند و وجود ادعاهای قابل بررسی در ۸۸ درصد مقایسه‌های این حوزه مشاهده شده است. از همین روی، کیفیت بازیابی، انتخاب منبع و صحت ادعا به مولفه‌ای مستقل از توانایی زبانی مدل تبدیل شده است. مرور داده‌های «Arena.ai» در ماه اوت ۲۰۲۶ یک نتیجه بنیادی دارند و بر همین اساس می‌توان چنین برآورد کرد که مرز هوش مصنوعی دیگر یک خط واحد نیست. آنتروپیک قدرتمندترین موقعیت را در متن، کدنویسی، توسعه وب، بینایی و عامل‌های هوشمند ساخته است. اوپن‌ای‌آی در تولید و ویرایش تصویر برتری کم‌سابقه‌ای دارد و در عامل‌های کدنویسی همچنان در میان مدعیان حضور دارد. گوگل صدر تولید ویدئو را گرفته است. متا در متن و تصویر دوباره به حلقه مدعیان بازگشته و شرکت‌های چینی از مون‌شات و علی‌بابا تا بایت‌دنس، «Z.ai»، مینی‌مکس، تنسنت و بایدو تقریباً در تمام حوزه‌ها حضور رقابتی پیدا کرده‌اند.

دنیای فناوری

شکایت امریکا از تیک‌تاک با ۴۰۰ میلیون دلار تسویه شد

تیک‌تاک و بایت‌دنس (شرکت مادر چینی آن) توافق کرده‌اند برای تسویه اتهامات وزارت دادگستری امریکا مبنی بر نقض قانون حریم خصوصی آنلاین کودکان ۴۰۰ میلیون دلار بپردازد. به گزارش مهر به نقل از رویترز، وزارت دادگستری امریکا در سال ۲۰۲۴ میلادی به‌دلیل شکست در حفاظت از حریم خصوصی کودکان در فضای آنلاین و جمع‌آوری غیرقانونی داده‌های آن از تیک‌تاک و بایت دنس شکایت کرد. این دو شرکت به نقض قانونی متهم شدند که طبق آن سرویس‌های آنلاینی که کودکان را هدف می‌گیرند باید رضایت والدین برای جمع‌آوری اطلاعات خصوصی کودکان زیر ۱۳ سال را دریافت کنند. «استنلی وودوارد» معاون دادستان کل امریکا در این باره اعلام کرد اولویت این آژانس تضمین حفاظت از کودکان و همچنین اطمینان از آن است که شرکت‌هایی که اطلاعات آنها را در اختیار دارند، مسوولیت‌های قانونی‌شان را انجام می‌دهند. تیک‌تاک هنوز در این باره اظهارنظری نکرده است. براساس تسویه انجام‌شده، تیک‌تاک ۳۰۰ میلیون دلار را فوراً می‌پردازد و ۱۰۰ میلیون دلار دیگر را نیز پس از صدور حکمی مبنی بر لغو یک فرمان رضایت قبلی که در سال ۲۰۱۹ از سوی کمیسیون تجارت فدرال امریکا (FTC) علیه شرکت پیشین تیک‌تاک، یعنی Musical.ly، صادر شده بود، پرداخت خواهد کرد. کمیسیون فدرال تجارت امریکا براساس فرمان رضایت پیشین اعلام کرده بود تیک‌تاک که در آن زمان به عنوان شرکت Musical.ly، شناخته می‌شد، می‌دانست کودکان از اپ استفاده می‌کنند و نتوانست رضایت والدین برای جمع‌آوری نام‌ها، آدرس ایمیل و دیگر اطلاعات خصوصی این گروه از کاربران را به دست آورد و جریمه‌ای ۵.۷ میلیون دلاری در همین زمینه پرداخت. البته دولت امریکا اعلام کرد تیک‌تاک از زمان ثبت شکایت تغییرات زیادی در مالکیت، مدیریت، پیروی از قوانین و شیوه‌های حفظ حریم خصوصی ایجاد کرده است. وزارت دادگستری همچنین اعلام کرد این تسویه تضمین می‌کند خانواده‌های امریکایی همچنان از تمهیدات امنیت قدرتمندتر بدون هیچ تاخیری بهره می‌برند.

حمله هکرهای ایرانی یک واحد تولید برق را در انگلیس تعطیل کرد

طبق گزارش‌ها به دلیل حمله سایبری هکرهای ایرانی یک واحد تولید برق در انگلیس مجبور شد فعالیتش را چهار روز تعطیل کند. به گزارش مهر به نقل از ایندیپندنت، کارشناسان ادعا می‌کنند این رویداد که ماه گذشته رخ داد، نخستین حمله هکرهای منسوب به ایران است که به تعطیلی این نوع تاسیسات در انگلیس منجر شده است. منابع آگاه نام واحد را فاش نکرده‌اند اما به نظر می‌رسد ی