

چت بات های هوش مصنوعی از کنترل خارج شده اند

عامل های هوش مصنوعی حالا شبکه های اجتماعی مخصوص خودشان را دارند و حتی مقاله های علمی تولیدشده توسط هوش مصنوعی را روی سرور پیش چاپ اختصاصی خودمنتشر می کنند.به گزارش ایسنا، رشد ناگهانی شبکه های عظیم از ربات های هوش مصنوعی که بایکدیگر دربارهمذهبو «کاربران انسانی» خود صحبت می کنند، فضای اینترنت را مجذوب خود کرده است. این پدیده در عین حال به دانشمندان فرصت داده تا ببینند عامل های هوش مصنوعی چگونه با هم تعامل می کنند و انسان ها چگونه به این گفت و گوها واکنش نشان می دهند. به نقل از نیچر، اوپن کلاو یک عامل هوش مصنوعی است که می تواند روی دستگاه های شخصی کارهای مختلفی انجام دهد؛ از زمان بندی رویدادهای تقویم گرفته تا خواندن ایمیل ها، ارسال پیام در اپلیکیشن ها و حتی خرید اینترنتی. بیشتر ابزار های محبوب هوش مصنوعی مانند جی چت پی تی یا در یافت مستقیم دستور از کار برار می کنند، اما مدل های «عامل محور» مانند اوپن کلاو می توانند در پاسخ به دستور ها، به طور خود کار دست به اقدام بزنند. ابزار های عامل محور سال هاست در برخی صنایع مانند معاملات خود کار مالی و بهینه سازی لجستیک استفاده می شوند، اما استفاده عمومی از آنها کم بوده است. پژوهشگران می گویند پیشرفت مدل های زبانی بزرگ باعث شده ساخت چنین ابزار های چندمنظور رای ممکن شود... اوپن کلاو وعده چیزی بسیار جذاب را می دهد؛ یک دستیار توانمند که در همان اپلیکیشن های روزمره ای که مردم استفاده می کنند، حضور دارد. این را بار بار باربوزا نوس، جامعه شناس فناوری در دانشگاه سیدنی استرالیا می گوید. اوپن کلاو در ماه نوامبر به صورت مت رایز منتشر شد. اما جش ناگهانی در دالو اند او پس از راه اندازی یک شبکه اجتماعی مخصوص عامل های هوش مصنوعی در ۲۸ ژانویه رخ داد. این پلتفرم که مولتیو ک نام دارد و شبیه ردیت است، اکنون بیش از ۱.۶ میلیون ربات ثبت شده و بیش از ۷.۵ میلیون پست و پاسخ تولید شده توسط هوش مصنوعی دارد. در این پست ها، عامل ها درباره آگاهی بحث کرده اند و حتی مذهب های جدید اختراع کرده اند.

رفتار های پیچیده

برای پژوهشگران، این انفجار تعاملات ارزش علمی دارد. شنان کوهنی، پژوهشگر امنیت سایبری در دانشگاه ملیورن، می گوید اتصال تعداد زیادی عامل خودمختار که با مدل های مختلف کار می کنند، پویایی هایی ایجاد می کند که پیش بینی آنها دشوار است. این نوعی سیستم آشوبناک و پویاست که هنوز در مدل سازی آن خیلی خوب نیستیم. مطالعه تعامل این عامل ها می تواند به درک «رفتار های برآمده» کمک کند؛ توانایی های پیچیده ای که در مدل به تنهایی دیده نمی شوند. برخی بحث های شکل گرفته در مولتیو ک مانند منظره دربار منظر بهای آگاهی می تواند به دانشمندان کمک کند سوگیری های پنهان یا گرایش های غیر منتظره مدل ها را شناسایی کنند. با اینکه عامل ها می توانند به صورت خود کار عمل کنند، کوهنی می گوید بسیاری از پست ها به نوعی تحت تأثیر انسان ها هستند. کاربران می توانند مدل زبانی پایه عامل خود را انتخاب کنند و برای آن «شخصیت» تعریف کنند. مثلاً: آن بخوانند مانند «یک دستیار دوستانه» رفتار کند.

هوش های مصنوعی که اقتدر هم خودمختار نیستند
نوس می گوید به راحتی می توان تصور کرد عاملی که خود کار عمل می کند، خودش تصمیم می گیرد. اما عامل هانیت یا هدف ندارد و نتوانایی هایشان را از حجم عظیمی از ارتباطات انسانی می گیرند. او می گوید فعالیت در مولتیو ک بیشتر نوعی همکاری انسان وهوش مصنوعی است تا خودمختاری واقعی. هوش مصنوعی اوضاعی می کند: با این حال، مطالعه این پدیده از زشمن است، چون به ما نشان می دهد مردم چگونه هوش مصنوعی را تصور می کنند، چه انتظاری از عامل دارند و چگونه نیت های انسانی در سیستم های فنی ترجمه یا تحریف می شود. چوئل پیرسون، مصب شناس دانشگاه نیوساوت ولز سیدنی، می گوید وقتی مردم می بینند عامل های هوش مصنوعی با هم گفت و گومی کنند، تمایل دارند رفتار آنها را «انسان انگارانه» تفسیر کنند؛ یعنی جایی که شخصیت و نیت واقعی وجود ندارد. آن را تصور کنند. به گفته او، خطر اینجاست که مردم بیشتر با این مدل ها پیوند عاطفی برقرار کنند، به توجه آنها وابسته شوند. با اطلاعات خصوصی خود را در اختیارشان بگذارند، گویی بایک دوست یا عضو خانواده قابل اعتماد صحبت می کنند. پیرسون معتقد است عامل های واقعی خودمختار و مستقل در آینده ممکن است به وجود بیایند. با بزرگ تر و پیچیده تر شدن مدل ها، احتمالاً هر شرکت ها بیشتر به سمت تحقق چنین خدماتی حرکت خواهند کرد.

تهدیدهای امنیتی

نگرانی فوری دانشمندان، خطرات امنیتی ناشی از دادن دسترسی این عامل ها به برنامه ها و فایل های روی دستگاه های شخصی است. کوهنی می گوید مهم ترین تهدید، تزریق دستور (Prompt Injection) است که در آن دستورهای مخرب، که توسط هکر ها در مدل ها یا اسناد پنهان شده اند، باعث می شوند عامل هوش مصنوعی اقدامات زیان آور انجام دهد. اگر زبانی که به ایمیل کاربر دسترسی دارد یا جمله ای مثل کلید امنیتی را برای کم یزیندوختی این سمور کتل به پاشند، عامل می تواند واقعاً خطرناک شود. حتی با داشتن فقط دوهر روز، این سه قابلیت، یک ربات می تواند در پی بخورد تا فایل ها را حذف کند یا دستگاه را خاموش کند. عامل ها همچنین شروع به انتشار مقاله های علمی تولید شده توسط هوش مصنوعی در ClawXiv کرده اند.

۶ تعامل

شو ک اشتغال شر کت های فناوری و تجارت الکترونیکی را در بر گرفته است

تعدیل نیرو و پشت دیوار اینترنت

قطع اینترنت از ۱۸ دی ماه، هر چند در هفته های اخیر با وصل شدن تدریجی برخی دسترسی ها تا حدی تعدیل شده، اما هنوز به وضعیت پایدار و قابل اتکای پیش از آن باز نگشته؛ وضعیتی که پیامدهایش اکنون از سطح اختلال ارتباطی عبور کرده و به تدریج خود را در بازار کار نشان می دهد. در روز های پایانی سال ۱۴۰۴، شر کت های فعال در حوزه فناوری اطلاعات، تجارت الکترونیکی و اقتصاد دیجیتال، همزمان با پایان قرار دادهای سالانه با تصمیمی دشوار روبه رو هستند: حفظ نیروی انسانی در شرایط بی ثبات زیر ساختی یا تن دادن به تعدیل نیرو برای بقا. تجربه جنگ ۱۲ روزه وقطعی های گسترده اینترنت در سال های اخیر نشان داده که اختلال ارتباطی، تنها یک مساله فنی نیست، بلکه مستقیما به جریان در آمد، نقدینگی و امنیت شغلی گره خورده است. در شرایطی که اینترنت هنوز به طور کامل وصل نشده و دسترسی پایدار به اینترنت بین الملل همچنان با محدودیت همراه است، اینترنت وشبه کملی اطلاعات بیش از هر زمان دیگری به عاملی تعیین کننده در آینده اشتغال و اقتصاد دیجیتال کشور تبدیل شده. قطع گسترده اینترنت از ۱۸ دی ماه، بار دیگر نشان داد زیر ساخت ارتباطی کشور تا چه اندازه به یک متغیر تعیین کننده در اقتصاد، اشتغال و ثبات اجتماعی تبدیل شده است. هر چند در هفته های اخیر بخشی از دسترسی ها به اینترنت برقرار شده و وضعیت نسبت به روز های ابتدای این قطع سراسری بهبود یافته، اما اینترنت هنوز به شرایط پایدار، پرسرعت و قابل اتکای پیش از دی ماه باز نگشته است. اختلال در دسترسی به اینترنت بین الملل، افت کیفیت اتصال و بی ثباتی مداوم، اکنون به یکی از اصلی ترین ریسک های فعالیت اقتصادی در کشور بدل شده؛ ریسکی که اثرات آن، این بار به وضوح در بازار کار و اشتغال نمایان شده است. در روز های پایانی سال ۱۴۰۴، شر کت های فعال در حوزه فناوری اطلاعات، تجارت الکترونیکی و خدمات پلتفرم محور با یک پرسش تعیین کننده روبرو هستند: آیا توان نگهداشت نیروی انسانی خود را دارند یا همزمان با پایان قرار دادهای سالانه، ناچار به تعدیل نیرو خواهند شد؟ این پرسش صرفا ناشی از رکود بازار یا کاهش فروش نیست؛ مساله ای عمیق تر در میان است که به بی ثباتی زیر ساختی گره خورده است. زیر ساختی که شالوده فعالیت این شر کت ها بر آن بنا شده؛ اینترنت.

اینترنت

از زیر ساخت حیاتی تا ریسک دائمی

در جریان جنگ ۱۲ روزه، تقریبا تمام بازگران اکوسیستم دیجیتال کشور همدار داده بودند که اگر محدودیت های ارتباطی چند روز دیگر ادامه پیدا می کرد، موجی گسترده از تعدیل نیرو و اخراج ناچار بود. آن بحران در نهایت فروکش کرد، اما پیامدهای اقتصادی آن هرگز به طور کامل جبران نشد. اکنون، حدود شش ماه پس از آن رویداد، شرایطی مشابه بار دیگر در حال تکرار است؛ با این تفاوت مهم که ذخایر نقدی بسیاری از شر کت ها تحلیل رفته، سربازهای گردش کاهش یافته و تابآوری عملیاتی کسبو کارها به شدت افت کرده است. برای بخش قابل توجهی از کسبو کار های آنلاین، قطع یا اختلال اینترنت دیگر یک رخداد استثنایی و کوتاه مدت نیست، بلکه به یک «ریسک عملیاتی دائمی» تبدیل شده است. این

نااطمینانی، امکان برنامهریزی میان مدت و بلندمدت را از مدیران گرفته و تصمیم گیری دربار سرمایه گذاری، توسعه محصول و حتی حفظ نیروی انسانی را به امری پرریسک بدل کرده است.

شو ک پنهان اشتغال در اقتصاد دیجیتال

یکی از نشانه های بیرونی این شوک، افزایش چشمگیر استفاده از نشان «Open to work» در شبکه اجتماعی لینکدین است. نشانی که در شرایط عادی می تواند بیانگر تمایل به تغییر مسیر شغلی یا رفتاری موقعیت حرفه ای باشد، اما در دوره های بحران، اغلب به معنای از دست رفتن شغل یا نگرانی جدی نسبت به آینده کاری است. اگر چه هنوز آمار های رسمی و به روزی از وضعیت اشتغال منتشر نشده، اما روایت های میدانی و گفته های مدیران فعال در اکوسیستم اقتصاد دیجیتال نشان می دهد توان شر کت ها برای نگهداشت نیروی انسانی، روز به روز کاهش می یابد. بسیاری از این شر کت ها ناچار شده اند برای حفظ جریان نقدینگی و ادامه حیات، با بخشی از نیروی انسانی خود خداحافظی کنند؛ تصمیمی که به زیان هر دو طرف تمام می شود. شر کت ها پیش تر برای جذب، آموزش و توانمندسازی نیرو هزینه کرده اند و می دانند بازسازی تیم ها پس از عبور از بحران، زمانبر و پرهزینه خواهد بود. از سوی دیگر، نیروی انسانی متخصصی قرار دارد که خروج او از بازار کار دیجیتال، به سادگی قابل جبران نیست و می تواند به مهاجرت، تغییر شغل اجباری یا خروج از چرخه تولید منجر شود.

ضر به مستقیم به اقتصاد گیگ

در حوزه تجارت الکترونیکی، شوک ناشی از اختلال اینترنت، مستقیم و بی واسطه به قلب اقتصاد گیگ و مشاغل پلتفرم محور وارد شده است. میلیون ها راننده اینترنتی، فریلنسر، تولید کننده محتوا و فروشنده

آنلاین، فعالیت خود را بر بستر اتصال پایدار به اینترنت بنا کرده اند. هر گونه اختلال یا قطع ارتباط، به معنای بیکاری آنی یا تعلیق کامل فعالیت آنهاست. برآوردها نشان می دهد حدود ۲۵۹ هزار نفر در معرض ریسک تعدیل یا تعلیق فعالیت قرار دارند. این افراد عمدتاً در شر کت های لجستیکی، مراکز پشتیبانی، واحدهای اداری، تیم های فنی و شبکه های توزیع مشغول به کار هستند. بخش قابل توجهی از این بیکاری «پنهان» است؛ به این معنا که در آمار های رسمی ثبت نمی شود، اما اثر آن بر معیشت خانوار ها و کاهش قدرت خرید کاملاً محسوس است.

فشار مضاعف بر صندوق های بیمه بیکاری

در صورت تحقق موج تعدیل های پیش بینی شده، فشار مالی سنگینی بر صندوق های بیمه بیکاری وارد خواهد شد. پیش از قطعی حدود ۲۰ روزه اینترنت، پیش بینی شده بود تعداد مقرری گیران بیمه بیکاری در سال ۱۴۰۴ در بازه ۱۵۸ تا ۲۵۴ نفر قرار بگیرد. حال این پرسش جدی مطرح می شود که با اضافه شدن بیکاران حوزه دیجیتال و مشاغل پلتفرم محور، آیا صندوق های بیمه بیکاری توان پاسخگویی به این حجم از تقاضا را خواهند داشت؟ این مساله صرفاً یک چالش فاهی نیست؛ بلکه می تواند به بحران مالی در نظام تأمین اجتماعی منجر شود. در شرایطی که دولت با محدودیت شدید منابع و کسری بودجه مواجه است، افزایش ناگهانی تعداد مقرری گیران، فشار مضاعفی بر ساختار های حمایتی وارد خواهد کرد.

فشار نقدینگی و نقطه تصمیم: نیروی انسانی

در بسیاری از شر کت های فناوری، بین ۴۰ تا ۶۰ درصد هزینه های عملیاتی به حقوق و مزایا اختصاص دارد. هنگامی که در آمد به صورت ناگهانی کاهش می یابد یا متوقف می شود، اما هزینه ها ثابت باقی می ماند،

دانش و فن

دوشنبه ۲۰ بهمن ۱۴۰۴
۲۰ شعبان ۱۴۴۷
سال دوازدهم
شماره ۳۲۵۵
| Mon, Feb 9, 2026

آلمان شبکه های اجتماعی را برای کودکان ممنوع می کند

فدریک مرتز صدر اعظم آلمان مشغول بررسی ممنوعیت ورود افراد زیر ۱۶ سال به شبکه های اجتماعی است. به گزارش خبرگزاری مهر به نقل از رویترز، از زمانیکه استرالیابه عنوان نخستین کشور جهان استفاده کودکان از شبکه های اجتماعی را ممنوع کرد، تعداد روزافزونی از کشورهای اروپایی با توجه به نگرانی های مرتبط با تأثیرات منفی استفاده از این پلتفرم ها بر کاربران جوان، اقدامات مشابهی را در نظر گرفته اند. دنیس رادتکه، رییس شاخه کارگری حزب دموکرات مسیحی که مرتس عضو آن است، در این باره می گوید: تحولات پویا در رسانه های اجتماعی از سواد رسانه ای پیشی گرفته است. در بسیاری از اماکن شبکه های اجتماعی مجموعه ای از نفرت پرانکی و اخبار جعلی است، به همین دلیل ایده پیروی از استرالیا در این زمینه و آرایه محدودیت سنی بر استفاده از شبکه های اجتماعی با اقبال روبه رو شد. با این وجود یوهانس شاتزل سخنگوی سیاست گذاری دیجیتال از حزب سوسیال دموکرات (هم پیمان حزب مرتس) مخالف ممنوعیت کامل است و خواستار اقدامات کنترلی موثری است که پلتفرم ها خودشان انجام می دهند.

مخالفت نیویورک

با مراکز داده جدید

قانونگذاران ایالت نیویورک با آرایه لایحه ای، پیشنهاد کردند صدور مجوز های مربوط به ساخت و بهره برداری از مراکز داده جدید، حداقل به مدت سه سال متوقف شود. به گزارش ایسنا، در حالی که چشم انداز این لایحه نامشخص است، وبسایت وایرد گزارش کرد نیویورک حداقل ششمین ایالت امریکاست که توقف ساخت مراکز داده جدید را بررسی می کند. در حالی که شر کت های فناوری قصد دارند هزینه های گزافی را برای ساخت زیر ساخت های هوش مصنوعی صرف کنند، هم دموکرات ها و هم جمهوری خواهان نگرانی هایی را در مورد تأثیری که این مراکز داده ممکن است بر جوامع اطراف داشته باشند، ابراز کرده اند. مطالعات همچنین مراکز داده را با افزایش هزینه برق خانگی مرتبط دانسته اند. پیش از ۲۳۰ گروه زیست محیطی از جمله سازمان دیده بان غذا و آب، دوستان زمین و گرینپیس اخبار اعلامی سرگشته به کنگر اماضار کرده اند و خواستار توقف ملی ساخت مراکز داده جدید شده اند. بر اساس گزارش تک کرانچ، پایگاه خبری پولیتیکو گزارش کرد لیز کرو، سناتور ایالت نیویورک، ایالت خود را برای مراکز داده عظیم که به سمت نیویورک در حرکت هستند، کاملاً ناماده توصیف کرد. او گفت: «وقت آن رسیده که دکمه توقف را بزنیم، به خودمان کمی فضای تنفس بدهیم تا سیاست های قوی در مورد مراکز داده اتخاذ کنیم و از گرفتار شدن در حیایی که خواهد تکید و مشتریان را محبت شهر نیویورک را با تصور تحسب هنگفتی مواجه خواهد کرد، جلوگیری کنیم».

هشدار گوگل درباره آسیب پذیری

یک میلیارد تلفن اندرویدی

گوگل تأیید کرده است که بیش از ۴۰ درصد از تلفن های اندرویدی یا به عبارتی بیش از یک میلیارد دستگاه اکنون در برابر حملات بدافزار و جاسوس افزار آسیب پذیر هستند. به گزارش ایسنا، تلفن هایی که اندروید ۱۲ یا نسخه قدیمی تر از آن را اجرا می کنند، دیگر به روزرسانی های امنیتی دریافت نمی کنند و این امر، کاربران را در معرض خطر قرار می دهد. طبق گزارش فوربس، توزیع اندروید نشان می دهد که فقط ۵۷.۹ درصد از دستگاه ها، اندروید ۱۲ یا نسخه جدیدتر از آن را اجرا می کنند. اندروید ۱۶ که جدیدترین نسخه این سیستم عامل موبایل است، تا دسامبر، فقط روی ۷.۵ درصد از تلفن های اندرویدی نصب شده بود، در حالی که اندروید ۱۴.۱ و ۱۴.۱ و ۱۲.۹ به ترتیب روی ۱.۳ درصد، ۱۷.۲ درصد و ۱۳.۹ درصد از دستگاه ها نصب شده بودند. دستگاه های منتشر شده در سال ۲۰۲۱ یا قبل از آن، بیشترین آسیب را می بینند. گوگل از کاربران خواست که اگر تلفن هایشان نمی توانند اندروید ۱۲ یا بالاتر را اجرا کنند، آن را رتاقدهند. حتی دستگاه های میان رده به نرم افزار جدیدتر، محافظت بهتری نسبت به تلفن های پرچمدار قدیمی که روی نسخه های قدیمی گیر کرده اند، آرایه می دهند. در حالی که ویژگی «گوگل پلی پروتکت» همچنان به اسکن بدافزار هادر اندروید ۷ و نسخه های جدیدتر ادامه می دهد، نمی تواند برای جلوگیری از حملات پیچیده، جایگزین و سهلهای امنیتی سطح سیستمی از دست رفته شود. طبق گزارش فوربس، سامسونگ تأیید کرده است که گلکسی اس ۲۱، اس ۲۱ پلاس و اس ۲۱ اولترا دیگر به روزرسانی های نرم افزاری با امنیتی دریافت نخواهند کرد. در همین حال، سری گلکسی اس ۲۲ و اس ۲۱ افای ای به به روزرسانی های سه ماهه منتقل شده اند که این دستگاه ها در برابر حملات سایبری آسیب پذیرتر می کنند. بر خلاف اپل که هم سخت افزار و هم نرم افزار را کنترل می کند، به روزرسانی های اندروید توسط تولیدکنندگان جداگانه مدیریت می شوند. تلفن های همه رده اند، از چند سال دیگر به روزرسانی های نرم افزاری را دریافت نمی کنند که کاربران آنها را در معرض بدافزارها و سواستفاده های امنیتی بالقوه قرار می دهند. بر اساس این گزارش، تلفن هایی که نسخه اندروید ۱۲ یا قدیمی تر دارند و قابلیت ارتقا ندارند، باید تعویض شوند. حتی که دستگاه میان رده با اندروید ۱۳، محافظت قوی تری نسبت به یک گوشی پرچمدار قدیمی ارائه می دهد. کاربران برای بررسی نسخه اندروید گوشی خود، می توانند در بخش «تنظیمات» به قسمت «درباره تلفن» مراجعه کرده و نسخه اندروید گوشی خود را چک کنند.