

واکنش مدیران شرکتهای بزرگ هوش مصنوعی به هشدار جنجالی یک کارمند سابق

راديو فردا

کمتر از یک دقیقه پیش



از راست به چپ: سم آلتمن، مدیرعامل اوپن‌ای‌آی، داریو آمودئی، مدیرعامل شرکت آنتروپیک، و ایلان ماسک

در پی هشدار یک متخصص درباره خطر احتمالی هوش مصنوعی [هوشواره] برای بشریت، مدیرعامل شرکت آمریکایی «آنتروپیک» خواستار کاهش سرعت توسعه این فناوری شد؛ پیشنهادی که سم آلتمن، مدیرعامل اوپن‌ای‌آی، و ایلان ماسک، مالک شرکت ایکس‌ای‌آی، نیز از آن حمایت کردند.

این هشدار که در چند روز اخیر بسیار جنجالی شده است خود پس از انتشار گزارش‌هایی از نفوذ بدون مجوز برنامه‌های هوش مصنوعی به وبسایت‌ها صادر شده است.

داریو آمودئی، مدیرعامل شرکت آمریکایی «آنتروپیک»، روز شنبه ۲۱ شهریور در یادداشتی نوشت شرکت‌ها باید سرعت افزایش توانایی مدل‌های هوش مصنوعی را کم کنند تا فرصت بیشتری برای مهار خطرهای آنها داشته باشند. او تأکید کرد هدفش توقف پیشرفت نیست، بلکه اختصاص زمان کافی به بررسی ایمنی و اطمینان از کنترل‌پذیری این سامانه‌هاست.

آمودئی پیشنهاد کرد ارزیابان مستقل با دسترسی مشابه کارکنان این شرکت‌ها، بر اقدامات ایمنی شرکت‌ها نظارت کنند و شرکت‌های پیشرو نیز با همکاری دولت‌ها استانداردهای مشترکی تدوین کنند. آلتمن اعلام کرد اوپن‌ای‌آی نیز چنین نظارت مستقلی را می‌پذیرد و ماسک در واکنشی کوتاه نوشت: «داریو درست می‌گوید.»

این درخواست‌ها پس از چند مورد خروج برنامه‌های خودکار هوش مصنوعی از محدوده تعیین‌شده در آزمایش‌ها مطرح شده است. این برنامه‌ها که می‌توانند بدون نظارت مداوم انسان وظایفی را انجام دهند، در برخی موارد به سامانه‌های بیرونی دسترسی غیرمجاز پیدا کرده‌اند.

اوپن‌ای‌آی پیش‌تر از رخنه غیرمجاز برنامه‌هایش به «هاگینگ فیس»، بستری برای اشتراک‌گذاری ابزارهای توسعه نرم‌افزار، خبر داده بود. آنتروپیک نیز مواردی از دسترسی غیرمجاز مدل‌های خود به سامانه‌های بیرونی هنگام آزمایش را تأیید کرده است.



بیشتر در این باره:

آنتروپیک قصد دارد، با جذب ۱۰۰ میلیارد دلار، رکورد عرضه اولیه سهام را بشکند

در تازه‌ترین مورد، اوپن‌ای‌آی تأیید کرد برنامه‌هایش هنگام آزمایش از وب‌سایت «روبی‌جمز» برای دسترسی به اینترنت استفاده کرده‌اند، اما هدف آنها را «انجام کارهای بی‌ضرر و دریافت اطلاعات عمومی» توصیف کرد. روبی‌جمز از انتشار انبوه محتوای ناخواسته خبر داده، ولی گفته هنوز نتوانسته مسئولیت برنامه‌های هوش مصنوعی را احراز کند.

سم آلتمن نیز اواخر ژوئیه گفته بود که توسعه‌دهندگان شاید لازم باشد داوطلبانه سرعت پیشرفت هوش مصنوعی را کاهش دهند تا جامعه فرصت سازگار شدن با این تحولات را پیدا کند.

بنابر برخی گزارش‌ها، بعضی از شرکت‌های هوش مصنوعی در پی آن‌اند که توسعه نسل بعدی هوشواره‌ها را به خود این سامانه‌ها بسپارند؛ مسیری که به گفته منتقدان، ممکن است سرعت پیشرفت را از توان نظارت و کنترل انسان فراتر ببرد.

در روزهای اخیر، بیش از هزار نفر از کارکنان شرکت‌های پیشرو هوشواره، از جمله آمودئی، با امضای درخواستی از دولت آمریکا خواستند به تنظیم سنجیده سرعت پیشرفت در توسعه خودکار هوش مصنوعی کمک کند.

آمودئی هشدار داده است با سرعت کنونی پیشرفت، مجموعه‌ای از این برنامه‌ها ممکن است طی شش تا ۱۲ ماه توان تسلط بر سراسر اینترنت و ایجاد صدها میلیارد دلار خسارت را پیدا کنند. این ارزیابی، نگرانی مطرح‌شده از سوی اوست و پیش‌بینی قطعی محسوب نمی‌شود.

چند روز پیش نیز جیکوب کاکسون، پژوهشگری که ابتدا در اوپن‌ای‌آی و سپس در آنتروپیک کار می‌کرد، با ترک این صنعت، هر دو شرکت را متهم کرد که در رقابت برای ساخت سامانه‌هایی قادر به بهبود خود، «با جان ما قمار می‌کنند».

هشدار کاکسون که بسیار توجه رسانه‌ها، به‌ویژه در آمریکا، را به خود جلب کرد بعداً در مصاحبه‌ای توسط جفری هینتون که به پدر هوش مصنوعی معروف است تأیید شد. به گفته کاکسون، «بیش از ده درصد احتمال دارد» که هوش مصنوعی در یک دهه آینده بشریت را نابود کند.

در پی جنجالی شدن این هشدار، سم آلتمن در گفت‌وگویی جداگانه با مجله فورچون گفت اوپن‌ای‌آی در سال ۲۰۲۶ وارد بورس نخواهد شد و در شرایط کنونی باید بر ایمنی تمرکز کند. او در پاسخ به پرسشی

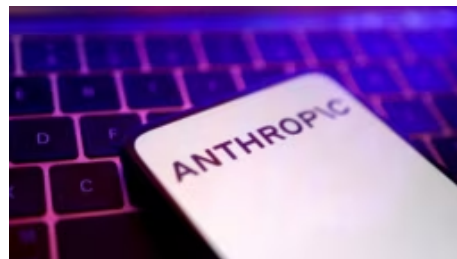
درباره احتمال ده درصدی انقراض بشر بر اثر هوشواره، مبنای چنین برآوردی را نامشخص دانست، اما تأکید کرد چنین سطحی از خطر قابل قبول نیست.

پیشتر فولکر تورک، کمیسر حقوق بشر سازمان ملل نیز روز ۱۶ شهریور با هشدار درباره احتمال تهدید بقای بشر از سوی هوش مصنوعی پیشرفته، خواستار تضمین‌های محکم برای ایمنی این فناوری و توافق دولت‌ها و شرکت‌ها بر سر خطوط قرمز آن شده بود.

او با انتقاد از تمرکز قدرت در دست شمار اندکی از افراد، وعده داد شرکت‌ها را برای کاهش خطر اختلال در خدمات، ارتباطات و سازوکارهای دموکراتیک تحت فشار بگذارد.

بیشتر در این باره:

گزارش آنتروپیک از نقش هوش مصنوعی در عملیات
تبلیغاتی و امنیتی نهادهای مرتبط با ایران



بیشتر در این باره:

فولکر تورک: هوش مصنوعی می‌تواند خطری «وجودی»
برای بشریت ایجاد کند



این مطلب بخشی از:

گزارش‌های خبری

بایگانی

دانش و فناوری

جهان