

فناوری و رسانه

نامه سرگشاده نخبگان فناوری دنیا

هوش مصنوعی را متوقف کنید

هزار و ۱۰۰ نفر از نخبگان حوزه فناوری دنیا در نامه‌ای سرگشاده نگرانی خود را از قدرت گرفتن خارج از کنترل هوش مصنوعی اعلام کردند.

حتماً این روزها به کرات به این جمله برخورد می‌کنید که از هوش مصنوعی خواستیم....

هوش مصنوعی که در ابتدا برای کمک به انسان و به دست گرفتن برخی امور ساخته شد، این روزها در حال تصرف تمامی ابعاد زندگی آدمی است تا جایی که نه تنها دستیار انسان در اندیشیدن، تصمیم‌گیری و اجرا محسوب می‌شود، بلکه به عنوان جایگزین او در حال به دست گرفتن تمامی حوزه‌ها است تا جایی که نخبگان فناوری که خود خالقان ذرات قدرت هوش مصنوعی هستند، در نامه‌ای سرگشاده خواستار توقف فعالیت آزمایشگاه‌های هوش مصنوعی و ارائه محصولی جدید در این رابطه شدند.

در نامه‌ای که نخبگانی از جمله استیو ورنیک بنیانگذار اپل، ایلان ماسک کالک توفیت، ایوان شارپ مؤسس پینترست، کریس لارسن از بنیانگذاران ریپل و ژان تالین بنیانگذار اسکایپ امضا کرده‌اند و در سایت مؤسسه Future of Life منتشر شده، آمده است: «ما از همه آزمایشگاه‌های هوش مصنوعی می‌خواهیم که بلافاصله آموزش سیستم‌های هوش مصنوعی قدرتمندتر از GPT-4 را حداقل به مدت شش ماه متوقف کنند. سیستم‌های هوش مصنوعی با هوش رقابتی انسان می‌تواند خطرات عمیقی را برای جامعه و بشریت ایجاد کنند، همان طور که تحقیقات گسترده نشان داده است و آزمایشگاه‌های برتر هوش مصنوعی نیز تأیید کرده‌اند. همان طور که بیان شده هوش مصنوعی پیشرفته می‌تواند نشان‌دهنده یک تغییر عمیق در تاریخ حیات روی زمین باشد و باید برای آن برنامه‌ریزی و با دقت و منابع متناسب مدیریت شود. متأسفانه، این سطح از برنامه‌ریزی و مدیریت اتفاق نمی‌افتد؛ در مقابل در ماه‌های اخیر آزمایشگاه‌های هوش مصنوعی در رقابتی خارج از کنترل برای توسعه و استقرار ذهن‌های دیجیتال قدرتمندتر که هیچ‌کس - حتی سازندگان آنها - نمی‌توانند آن را درک کنند، قفل شده‌اند.»

خطر از دست دادن کنترل تمدن‌ها

در ادامه این نامه با طرح سؤالی حتی کارکرد هوش مصنوعی به چالش کشیده شده و آمده است: «سیستم‌های هوش مصنوعی معاصر اکنون در حال تبدیل شدن به رقابت با انسان در وظایف عمومی هستند، و ما باید از خود بپرسیم: آیا باید به ماشین‌ها اجازه دهیم کانال‌های اطلاعاتی ما را پر از تبلیغات نادرست کنند؟ آیا باید تمام مشاغل را خودکار کنیم؟ آیا باید ذهن‌های غیرانسانی را پرورش دهیم که ممکن است در نهایت به لحاظ تعداد، پیشی گرفته، منسوخ و جایگزین ما شوند؟ آیا باید در خطر از دست دادن کنترل تمدن خود باشیم؟ چنین تصمیماتی نباید به رهبران فناوری منتخب واگذار شود. سیستم‌های هوش مصنوعی قدرتمند تنها زمانی باید توسعه یابند که مطمئن شویم اثرات آنها مثبت و خطرات آنها قابل مدیریت خواهد بود. این اطمینان باید به خوبی توجیه شود و با بزرگی اثرات بالقوه یک سیستم افزایش یابد.»

پیشنهادهایی که در نهایت ارائه شده این است که «آزمایشگاه‌های هوش مصنوعی و کارشناسان مستقل باید از این توقف برای توسعه و اجرای مشترک مجموعه‌ای از پروتکل‌های ایمنی مشترک برای طراحی و توسعه هوش مصنوعی استفاده کنند که به شدت توسط کارشناسان مستقل خارجی بررسی و نظارت می‌شوند. این پروتکل‌ها باید اطمینان حاصل کنند که سیستم‌هایی که به آنها پایبند هستند، فراتر از یک شک منطقی، ایمن هستند. این به معنای توقف توسعه هوش مصنوعی به طور کلی نیست، بلکه صرفاً یک عقب‌نشینی از مسابقه خطرناک به سمت مدل‌های جعبه سیاه غیرقابل پیش‌بینی بزرگتر با قابلیت‌های نوظهور است.

تحقیق و توسعه هوش مصنوعی باید روی ساختن سیستم‌های قدرتمند و پیشرفته‌تر امروزه دقیق‌تر، ایمن‌تر، قابل تفسیرتر، شفاف‌تر، قوی‌تر، همسوتر، قابل اعتمادتر و وفادارتر متمرکز شود.

به موازات آن، توسعه دهندگان هوش مصنوعی باید با سیاستگذاران برای تسریع چشمگیر توسعه سیستم‌های مدیریتی قوی هوش مصنوعی همکاری کنند. این موارد حداقل باید شامل موارد زیر باشد: مقامات نظارتی جدید و توانمند که به هوش مصنوعی اختصاص داده شده‌اند. نظارت و ردیابی سیستم‌های هوش مصنوعی بسیار توانمند و مجموعه‌های بزرگ توانایی محاسباتی؛ سیستم‌های منشأ و واترمارکینگ برای کمک به تشخیص واقعی از مصنوعی و ردیابی نشئت مدل. یک اکوسیستم حساس‌رسی و صدور گواهی‌نامه قوی؛ مسئولیت آسیب ناشی از هوش مصنوعی؛ بودجه عمومی قوی برای تحقیقات ایمنی فنی هوش مصنوعی و مؤسسات دارای منابع مناسب برای مقابله با اختلالات اقتصادی و سیاسی چشمگیر (به ویژه در دموکراسی) که هوش مصنوعی ایجاد خواهد کرد.»

داستان اخراج کارمندی که به احساس غیرعادی هوش مصنوعی پی برد

در مقابل اعتراضاتی که در خصوص قدرت گرفتن هوش مصنوعی مطرح می‌شود، آیا حاکمان رسانه منافع خود را در راستای این خواسته می‌دانند؟ تجربه‌ای که اخیراً شنیده شده است نشان می‌دهد که قدرت‌گیری هوش مصنوعی تا زمانی که منافع رسانه تأمین شود، مورد استقبال و حمایت است. بلیک لمواین، یکی از مهندسان گوگل که علناً مدعی بود هوش مصنوعی محاوره‌ای این شرکت (LaMDA) خودآگاه شده، در ماه‌های اخیر اخراج شد. این کارمند ادعا کرده بود که مکالمات او با چت روبات LaMDA او را به این باور رساند که این برنامه به چیزی فراتر از هوش مصنوعی تبدیل شده و افکار

و احساسات خاص خود را دارد. گوگل معتقد است که ادعاهای بلیک لمواین را به طور گسترده بررسی کرده و براساس نتایج، آنها کاملاً بی‌اساس هستند. گوگل در بیانیه‌ای به اتهامات لمواین مبنی بر اینکه این شرکت به درستی ادعاهای او را بررسی نکرده است، چنین پاسخ داد که «همان‌طور که ما به خط مشی خود در حوزه هوش مصنوعی معتقدیم، توسعه هوش مصنوعی را بسیار جدی می‌گیریم و به نوآوری ادامه می‌دهیم LaMDA از ۱۱ بررسی مجزا عبور کرده است و ما در یک مقاله تحقیقاتی اوایل امسال جزئیات کار را منتشر کردیم که در راستای توسعه مسئولانه آن است. اگر کارمندی مانند بلیک نگرانی‌هایی را در مورد

کار به اشتراک بگذارد، ما آن را به طور کامل بررسی می‌کنیم. ما متوجه شدیم که ادعاهای بلیک مبنی بر اینکه LaMDA احساس دارد، کاملاً بی‌اساس است. ما برای روشن کردن قضیه ماه‌ها روی این ادعا کار کرده ایم.» با وجود نوع برخورد گوگل و تلاش این مجموعه برای پاسخگویی به افکار عمومی همچنان نگرانی‌هایی در خصوص قدرت گرفتن غیرعادی و خارج از کنترل هوش مصنوعی وجود دارد. این روزها این سؤال مطرح است که آیا ممکن است هوش مصنوعی حاکم بر انسان و روابط او شده و تمامی ابعاد زندگی وی را در اختیار تصمیمات خود بگیرد؟

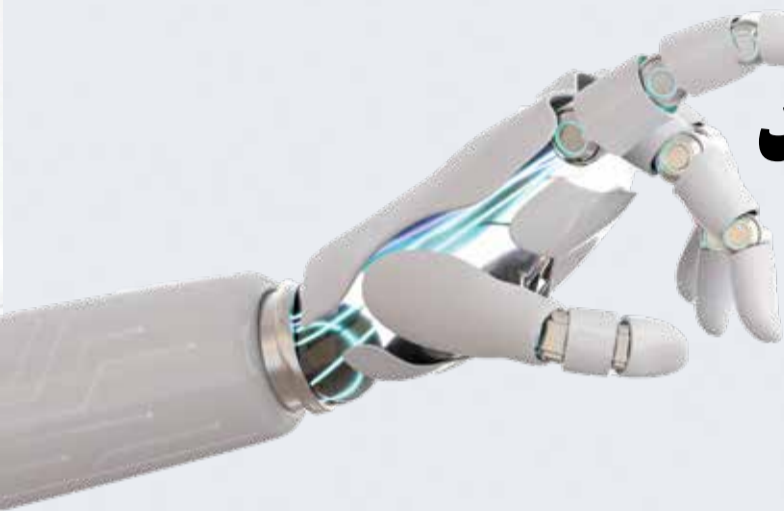


پادداشت

دوبس پهنود
مدیر تحریریه
info@colleged.com

ماکارونی با سویا
سبک مصرف رسانه

شاید اخبار مربوط به قتل خیابان فرشته را شنیده باشید. قاتل به سبک سریال خون سرد برادر خود را در یک ساختمان اداری خالی مله کرده و در اعتراضات نیز گفته که شیوه قتل را از این سریال آموخته است. هر چند که سرپرست نویسندگان این سریال عنوان کرده که این اعتراف یک توجیه بوده و هر صحنه‌ای از فیلم‌های جنایی ممکن است توسط آدم‌ها تقلید شود اما یکی از سؤالات جدی این است که رژیم مصرف ما در رسانه به چه شکلی است و آیا قرار است با یک صحنه عاشقانه عاشق و با دیدن یک اختلاف خانوادگی با والدین خود درگیر شویم؟ آیا قرار است ادای فلان خواننده را در بیاوریم و یا مدل لباس پوشیدن فلان بازیگر را تقلید کنیم؟ بعد از شیوع شبکه‌های اجتماعی در ایران بزرگ‌ترین باری که نهاد‌های تربیتی و نظارتی بر عهده گرفته‌اند، ساختن تصویری مخرب از محتواها و بحث فیلترینگ بوده است. موضوعی که در پی اعتراضات اخیر به صورت جدی‌تر مطرح شد و واتس‌آپ و اینستاگرام که از قضا هر دو در ایران پرطرفدار بود به صورت کامل فیلتر شدند. در مقابل اپ‌های داخلی تلاش کردند تا نیاز مرادوات مردم را پاسخ دهند اقدامی که البته حداقل در اعتراف زبانی کاربران، موفقیت‌آمیز نبوده و همچنان مردم به سراغ وی‌پی‌ان‌ها می‌روند تا بتوانند شبکه‌های فیلتر شده را حتی برای چت‌های ساده مورد استفاده قرار دهند. رژیم ما در مصرف رسانه شده است چیزی شبیه واقعیت که نه مزه و طعم واقعی را دارد و نه لذتی حقیقی. موضوع بی‌اطلاعی در مصرف محتواها و در مقابل صرفاً نگاه فیلترینگی و از ریشه خاموش کردن محتوا این بدخوری را تشدید کرده و موجب شده تا کاربر ایرانی به سادگی درگیر محتوا شده و به جای قدرت پایش و تحلیل آن به اصطلاح در جو یک محتوای سه خطی اسیر شود. ذهن کاربر ایرانی با وجود نرده‌های فیلتر همچنان در جست‌وجوی محتواها است و از هر رسانه‌ای که بتواند برای برخوردی و بدخوری استفاده می‌کند اما در مقابل به حدی مسخ شده از محتواست که ممکن است دستاورد آن یک قتل یا پوشش نامناسب یا بدرفتاری با خانواده یا بداخلاقی در جامعه باشد. باید معترف بود که این نرده‌ها به حدی کم توان بوده که نتوانسته مخاطب را محافظت کند و گاهی از رسانه خودی نیز تأثیری منفی مشاهده می‌شود و این در حالی است که هزاران فیلم با محتواهای روان‌آزار در دنیا ساخته می‌شود و در پی آن دسته‌ای از افراد جامعه به دنبال جرم و جنایت نمی‌روند.



«سیستم‌های هوش مصنوعی معاصر اکنون در حال تبدیل شدن به رقابت با انسان در وظایف عمومی هستند، و ما باید از خود بپرسیم: آیا باید به ماشین‌ها اجازه دهیم کانال‌های اطلاعاتی ما را پر از تبلیغات نادرست کنند؟»

