جهان برای حرکت در جغرافیای سیاسی GenAI آماده می‌شود

صدای پای هوش مصنوعی مولّد را می‌شنویم؟

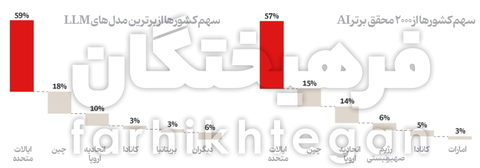
علی محجوبی
مترجم

ایران همواره از کشورهایی بوده که دانشمندان آن برای علم در جهان، آورده‌های خوبی داشته‌اند. پیشرو بودن ما در تعداد نه چندان کمی از حوزه‌های علمی برکسی پوشیده نیست ولی عضویت به تنهایی در «باشگاه فناوران جهان» برای ما کافی نیست. تبدیل بودن دراین باشگاه رقابتی، آینده‌ای جز حذف و مصرف‌کننده صرف نشدن ندارد. ما اکنون در عصری هستیم که فنون، سریع‌تر از هر زمانی حرکت می‌کنند و البته فناوری بازو عصر ما، مانند شیمی و پزشکی و علوم مهندسی کلاسیک نیست که هر دستاورد، سال‌ها زمان نیاز داشته باشد. فناوری بازو عصر ما به شکلی است که با چند روز کدنویسی در صفحه یک رایانه شخصی ارزان‌قیمت، ممکن است آورده‌ای به وجود آورد که برای به دست آوردن آن، دولت‌ها و حتی شرکت‌ها به جنایات متعدد دست بزنند. با این حال، نگرانی مهمی که وجود دارد، این است که درحالی ما تازه می‌خواهیم به شکل جدی به این موضوع وارد شویم که دنیا از آن عبور کرده و آینده منافع ملی کشورها و منافع سازمانی شرکت‌ها، به هوش مصنوعی مولد گره خورده. ظاهراً ما عزم اعتقاد داریم در مسابقه بین لاکچشت و خرگوش هستیم، درحالی‌که مسابقه حقیقی، چیزی است بین هواپیما و فضاپیما! از این رو، امید وجود دارد که مقاله زیر، نگاهی واقع‌بینانه نسبت به آنچه در اطراف ما در حال وقوع است به خوانندگان هدف آن، ارائه دهد. این گزارش تحلیلی بازگردان‌شده از گروه مشاوران بوستون (BCG)، بر عرضه هوش مصنوعی مولد متمرکز است که نقشی اساسی در شکل‌دهی به پویایی‌های جغراسیاسی این فناوری خواهد داشت. اما در سوی دیگر عرضه، پذیرش واقعی این فناوری نیز اهمیت بالایی دارد. این گزارش بر آن است که با توجه به اثرات فوق‌العاده مهمی که هوش مصنوعی مولد (Generative AI یا مخفف آن، GenAI) بر جهان خواهد گذاشت- از ایجاد مراکز داده تا تقویت تولیدات علمی دانشگاهی و افزایش بازدهی فعالیت‌های دولتی و تجاری- تمام کشورهایی که از حداقلی‌ترین منابع لازم برای ورود به این عرصه بهره‌مندند، برای رقابت با کشورهای قدرتمندتر و تأمین امنیت ملی و منافع خود، عمیقاً بر این حوزه متمرکز شده‌اند. درسی که این گزارش می‌تواند برای ایران به عنوان کشوری که از منابع علمی و علیه‌رغم درروزنا و بسیار خوبی برخوردار است داشته باشد، درواقع لزوم اتصال و هم‌افزایی بخش‌های مختلف کشور و تغییر در نگرش مدیران به فناوری‌های پیش‌برنده‌ای نظیر هوش مصنوعی مولد است. این اتفاق که بنا بر گزارش تفصیلی زیر، در کشورهای عضو اتحادیه اروپا رخ داده، عرصه را برای حفاظت راحت‌تر از امنیت ملی آن‌ها در آینده، هموارتر کرده است. مدیران باید بتوانند این پویایی را درک کرده و خود و مجموعه‌های خود را در جغرافیای سیاسی چالش برانگیز GenAI مدایست کنند و آگاه باشند که دیگر رویکردهای سنتی (از میزان دسترسی کشورها به نیمه‌رسانا‌های پیشرفته تا لزوم وجود مقررات مطلوب) برای این دوران کافی نخواهند بود و سرمایه، استعداد، مالکیت فکری (IP)، داده، انرژی و توان محاسباتی، اگر از آن سوار مهم‌تر نباشند، کم‌اهمیت‌تر نیز نخواهند بود. در ادامه، گزارش BCG را مطالعه خواهید کرد.

پیش‌فرض جهان «ساختِ آمریکا» است

ایالات متحده و چین در حال حاضر تنها بازیگرانی هستند که به بخش‌های گسترده‌ای از زنجیره ارزش هوش مصنوعی مولد دسترسی پایدار و کنترل کامل دارند. این دو کشور بزرگ‌ترین منابع استعدادهای هوش مصنوعی را در اختیار دانشه، برخی از غنی‌ترین بوم‌سازگان‌های داده‌ای جهان و ظرفیت زیرساختی گسترده‌ای در مراکز داده را دارا هستند و در دسترسی به سرمایه نیز پیشروانند. با این مورد، ایالات متحده برتری چشمگیری در GenAI داشته که بر پایه دهه‌ها رهبری در زمینه هوش مصنوعی بنا شده است. نزدیک به ۷۰ درصد از مدل‌های برجسته هوش مصنوعی جهان از سال ۱۹۵۰ تاکنون، توسط شرکت‌ها یا مؤسسات آکادمیک مستقر در ایالات متحده و چین در حال حاضر تنها بازیگرانی هستند که در بخش‌های گسترده‌ای از زنجیره ارزش هوش مصنوعی مولد دسترسی پایدار و کنترل کامل دارند. این دو کشور بزرگ‌ترین منابع استعدادهای هوش مصنوعی را در اختیار دانشه، برخی از غنی‌ترین بوم‌سازگان‌های داده‌ای جهان و ظرفیت زیرساختی گسترده‌ای در مراکز داده را دارا هستند و در دسترسی به سرمایه نیز پیشروانند. با این مورد، ایالات متحده برتری چشمگیری در GenAI داشته که بر پایه دهه‌ها رهبری در زمینه هوش مصنوعی بنا شده است. نزدیک به ۷۰ درصد از مدل‌های برجسته هوش مصنوعی جهان از سال ۱۹۵۰ تاکنون، توسط شرکت‌ها یا مؤسسات آکادمیک مستقر در ایالات متحده و با در همکاری با آن‌ها توسعه یافته‌اند و ۵۷ درصد از برترین مدل‌های زبان بزرگ نیز در این کشور شکوفا شده‌اند. حدود ۶۰ درصد از ۲۰۰۰ محقق برجسته هوش مصنوعی در جهان ساکن ایالات متحده هستند و این کشور بین سال‌های ۲۰۲۲ تا ۲۰۲۴، تقریباً یک چهارم تمام متخصصان هوش مصنوعی مهاجرت کرده در سطح جهانی را به خود جذب کرده است. در مجموع، ذخیره استعداد هوش مصنوعی ایالات متحده به نزدیک نیم میلیون نفر رسیده که بزرگ‌ترین مجموعه در جهان است. از سوی دیگر، پژوهشگران آمریکایی در زمینه هوش مصنوعی از زمان آغاز این رشته، ۳۵ درصد از تأثیرگذارترین مقالات علمی جهان را تألیف کرده‌اند.

پیشرو بودن ایالات متحده در تعداد دانشمندان حوزه AI و مدل‌های برتر GenAI



منبع: Stanford's HELM Leaderboard, BCG Henderson Institute

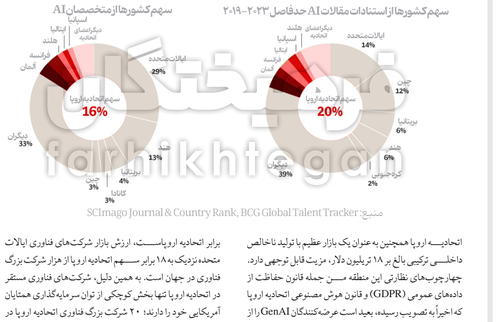
استارتاپ‌های GenAI مستقر در ایالات متحده نیز بی‌سابقه‌ترین میزان سرمایه‌گذاری خصوصی را جذب کرده‌اند؛ از سال ۲۰۱۹ تاکنون درمجموع ۶۵ میلیارد دلار. همچنین، پیش‌بینی می‌شود هزینه‌های سرمایه‌ای مرتبط با هوش مصنوعی در شرکت‌های فناوری تثبیت‌شده مانند آلفابت، آمازون، متا و مایکروسافت در سال ۲۰۲۴ از مرز ۲۰۰ میلیارد دلار فراتر رود. با ظرفیت محاسباتی عظیم، رویکرد جامع (از طراحی سخت‌افزار تا توسعه نرم‌افزار)، سرمایه فراوان، دسترسی جهانی، بهترین استعدادهای جهان و همکاری نزدیک با درون‌سازمانی با توسعه‌دهندگان مدل‌ها، این شرکت‌ها به شکلی بی‌مانند برای حفظ رهبری ایالات متحده در هوش مصنوعی مولد در آینده نزدیک مجهز شده‌اند. علاوه بر استعداد و سرمایه، ایالات متحده زیرساخت‌هایی برای پشتیبانی از استنتاج در مقیاس وسیع دارد؛ یعنی تأمین مداوم GenAI برای کاربران مدل‌ها. این کشور، بزرگ‌ترین قطب مراکز داده در جهان است و انتظار می‌رود ظرفیت آن در سال ۲۰۲۴ حدود ۴۵ گیگاوات باشد. ایالات متحده با دسترسی مطمئن به سخت‌افزارهای پیشرفته (از طریق شرکت آمریکایی انویدیا و روابط جغراسیاسی قوی با شرکت تایوانی TSMC)، توانایی گسترش مراکز داده خود را دارد، هرچند که زمان طولانی اتصال به شبکه برق ممکن است چالش‌برانگیز باشد.

افق روشن برای قدرت‌های میانه

داستان رقابت ایالات متحده و چین، برساخت «دو ابرقدرت» را تقویت کرده و به طور قانع‌کننده‌ای این دیدگاه را ارائه می‌دهد. اما در سایر نقاط جهان نیز حرکت‌های مهمی در جریان است. به عنوان نمونه، اتحادیه اروپا از تکامل متقابل توانمندی‌ها میان کشورهای عضو خبر می‌دهد که هر کدام، نقاط قوت ویژه‌ای دارند. در منطقه خلیج فارس، سرمایه متمرکز و چابک به همراه انرژی ارزان، حتی بدون وجود حوزه فناوری قوی و مستقر، پیشرفت را سرعت می‌بخشند. کره‌جنوبی و ژاپن نیز هر دو از بخش‌های فناوری قدرتمند و سرمایه کافی برای گسترش برخوردارند.

برای برخی کشورها، آنچه در سطح فردی خارج از دسترس به نظر می‌رسد، در سطح گروهی ممکن می‌شود. اتحادیه اروپا از تکامل متقابل عوامل کلیدی در کشورهای عضو و مقیاسی که به عنوان یک بازار واحد می‌تواند ارائه دهد، بهره‌مند است. این اتحادیه در حال حاضر میزبان بوم‌سازگان نوپای استارتاپ‌های هوش مصنوعی مولد است. شرکت فرانسوی میسترال (Mistral) تاکنون هفت مدل زبان بزرگ از برترین مدل‌های جهان را توسعه داده و ۲ و ۱ میلیارد دلار سرمایه جذب کرده است. استارتاپ آلمانی الف‌آلفا (Aleph Alpha) نیز مدل‌های بنیادین قدرتمندی تولید کرده است (دو مورد از برترین مدل‌های زبان بزرگ جهان)، هرچند اخیراً تمرکز خود را بر کاربردهای خاص صنعتی تغییر داده است. استارتاپ‌های کوچک‌تر نیز سرمایه‌گذاری‌های خوبی جذب کرده‌اند. شرکت کیوتای (Kyutai) که در سال ۲۰۲۳ تأسیس شد، حدود ۳۵۰ میلیون دلار سرمایه دریافت کرده و مدل MoshiAI را برای پردازش پیشرفته گفتار عرضه کرده است. شرکت اچ (H) نیز با جذب ۲۲۰ میلیون دلار سرمایه اولیه، در حال توسعه عوامل هوش مصنوعی برای افزایش بهره‌وری است. در همین حین، پوساید (Poolside) توانسته ۶۲۶ میلیون دلار برای توسعه یک مدل پیشرو در کدنویسی جذب کند و بلک فارست لبز (Black Forest Labs) نیز حدود ۱۵۰ میلیون دلار برای توسعه مدل بنیادین تبدیل متن به تصویر خود جذب کرده است. اگرچه این شرکت‌ها از نظر اندازه و منابع مالی در مقایسه با همتایان آمریکایی و چینی خود کوچک‌ترند، اما بر پایه استعدادی قوی بنا شده‌اند. اتحادیه اروپا دارای یک مجموعه گسترده و در حال رشد از متخصصان هوش مصنوعی است که تنها در فرانسه و آلمان به بیش از ۱۰۰ هزار نفر می‌رسد.

مزیت اروپا در ترکیبی بودن توان اعضای آن



منبع: SCImago Journal & Country Rank, BCG Global Talent Tracker

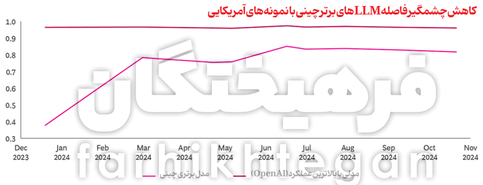
اتحادیه اروپا همچنین به عنوان یک بازار عظیم با تولید ناخالص داخلی ترکیبی بالغ بر ۱۸ تریلیون دلار، مزیت قابل توجهی دارد. چهارچوب‌های نظارتی این منطقه مسن جمله قانون حفاظت از داده‌های عمومی (GDPR) و قانون هوش مصنوعی اتحادیه اروپا که اخیراً به تصویب رسیده، بعید است عرضه‌کنندگان GenAI را از فعالیت در این منطقه منصرف کنند. در واقع این قوانین می‌توانند تقاضا برای هوش مصنوعی توسعه‌یافته و مبتنی‌بر داده در اتحادیه اروپا را افزایش دهند؛ چراکه ممکن است قابل اعتمادتر و محافظت‌کننده‌تر از داده‌های کاربران تلقی شوند. اتحادیه اروپا همچنین دارای سومین ظرفیت بزرگ مراکز داده در جهان- پس از ایالات متحده و چین- است که در سال ۲۰۲۴ به ۹ گیگاوات می‌رسد. هرچند قیمت برق در این منطقه بالاست، اما تنوع قابل توجهی وجود دارد؛ هزینه برق صنعتی در فرانسه ۲۰ درصد کمتر و در سوئد و فنلاند نزدیک به ۶۰ درصد کمتر از آلمان است. علاوه بر این، شرکت‌هایی که قصد دارند مدل‌های هوش مصنوعی مولد خود را به بازار اتحادیه اروپا عرضه کنند، ممکن است به دلایل نظارتی، چاره‌ای جز گسترش مراکز داده خود در این منطقه نداشته باشند. بر اساس این نقاط قوت، شاهد تکامل متقابل در عمل هستیم؛ به عنوان مثال، مدل‌های میسترال در فرانسه طراحی می‌شوند، روی ابررایانه‌های ایتالیایی آموزش می‌بینند و به مشتریانی معمولاً خارج از مراکز داده سوئدی ارائه می‌شوند. همچنین، این مدل‌ها از چیپ‌هایی که توسط تجهیزات لیتوگرافی EUV که انحصاراً توسط شرکت هلندی ASML تولید می‌شوند، قدرت می‌گیرند. البته بزرگ‌ترین چالش اتحادیه اروپا، تأمین سرمایه‌گذاری‌های لازم برای همگام شدن با روند گسترش مدل‌ها و توسعه زیرساخت‌های مراکز داده (از جمله ارتقای شبکه برق) است. تارنجخه سرمایه‌گذاری ناکافی در بخش فناوری، توانایی اتحادیه اروپا را برای مقیاس‌دهی به شرکت‌های برتر فناوری محدود کرده است. در حالی که تولید ناخالص داخلی ایالات متحده ۱/۵ برابر اتحادیه اروپا است، ارزش بازار شرکت‌های فناوری ایالات متحده نزدیک به ۱۸ برابر سهم اتحادیه اروپا از هزار شرکت بزرگ فناوری در جهان است. به همین دلیل، شرکت‌های فناوری مستقر در اتحادیه اروپا تنها بخش کوچکی از توان سرمایه‌گذاری همتایان آمریکایی خود را دارند؛ ۲۰ شرکت بزرگ فناوری اتحادیه اروپا در سال ۲۰۲۲ مجموعاً ۴۰ میلیارد دلار در تحقیق و توسعه هزینه کردند، در حالی که این رقم برای ۲۰ شرکت بزرگ فناوری ایالات متحده به ۲۰۰ میلیارد دلار رسید. طبیعی است که استارتاپ‌های GenAI در اتحادیه اروپا از سال ۲۰۱۹ تاکنون تنها ۳ میلیارد دلار سرمایه‌گذاری جذب کرده‌اند که معادل ۵ درصد از سرمایه‌گذاری خصوصی استارتاپ‌های مشابه در ایالات متحده است. با این حال، نشانه‌هایی از حرکت مثبت دیده می‌شود. گزارش سال ۲۰۲۲ رقابت‌پذیری اروپا (که توسط نخست وزیر سابق ایتالیا تهیه شده و به افتخار وی «گزارش دراگی» نام دارد) و اعلامیه‌های اخیر بانک سرمایه‌گذاری اروپا (EIB) و کمیسیون اروپا، همگی بر نیاز به یک بوم‌سازگان فناوری قوی‌تر، به همراه تأمین مالی لازم برای رشد استارتاپ‌ها در داخل این اتحادیه، تأکید دارند. منابعی وجود دارند که اتحادیه اروپا می‌تواند از آن‌ها استفاده کند، از جمله هزینه‌های عمومی تحقیق و توسعه در کشورهای عضو که میانگین آن حد فاصل سال‌های ۲۰۱۶ تا ۲۰۲۲ حدود ۴۰ میلیارد دلار در سال بوده است. اما استفاده از این منابع، نیازمند تلاشی هماهنگ در محیطی پیچیده با تصمیم‌گیری‌های جمعی و توزیع شده است. پیشینه‌هایی برای چنین همکاری‌هایی وجود دارد، به ویژه در تلاش‌های مشترکی مانند ایجاد کنسرسیوم هواپیمایی اروپا یا همان ایرباس. اما هماهنگی در فرانسوی مرزها کافی دشوار بوده است، همانند پروژه Gaia X که هدف آن، توسعه زیرساخت ابری قابل اعتماد و حاکمیتی برای اتحادیه اروپا بود. هنوز مشخص نیست که آیا شتاب کنونی سیاست‌گذاری به اقداماتی به موقع بدل خواهد شد یا خیر.

«ساختِ چین» هر روز بهتر از دیروز

نشانه‌هایی مبنی بر نزدیک شدن چین به ایالات متحده در برخی از عوامل کلیدی وجود دارد؛ چراکه این کشور پیشرفت سریعی در تولید مدل‌های زبان بزرگ برتر داشته است. دوشرکت چینی، علی‌بابا و استارتاپ هوش مصنوعی مولد 01.AI پیش از یک چهارم از مدل‌های متن‌باز برتر جهان را توسعه داده‌اند. غول‌های فناوری مستقر در چین مانند بایدو (Baidu) و تنسنت (Tencent) نیز مدل‌های قدرتمندی عرضه کرده‌اند. درحالی‌که نسل جدیدی از استارتاپ‌های هوش مصنوعی مولد، موسوم به «ببرهای هوش مصنوعی» شامل MoonshotAI، BaichuanAI، ZhipuAI و MiniMaxAI به سرعت در حال پیشرفتند. مدل‌های برتر چینی طی سال گذشته به شکل قابل توجهی فاصله خود را با نمونه‌های پیشرفته جهانی کاهش داده‌اند و در شاخص‌های مربوط به زبان چینی عملاً این فاصله کاملاً از بین رفته است.

بوم‌سازگان GenAI چین از دسترسی به مدل‌های بنیادین متن‌باز یا با وزن‌باز خارجی بهره‌مند شده است. به عنوان مثال، خانواده مدل‌های Yi متعلق به 01.AI، بر اساس مدل لاما (Llama) از شرکت متا طراحی شده‌اند. دولت چین قصد خود را برای تضمین عرضه کافی مدل‌های متن‌باز داخلی که نقش مشابهی را در آینده ایفا کنند، اعلام کرده و از شرکت‌های کوچک‌تری که توانمندی‌های فنی خود را ارتقا می‌دهند، حمایت می‌کند. اگرچه چین همچنان از ایالات متحده عقب است، اما از منابع قابل توجهی در حوزه استعدادهای هوش مصنوعی و مؤسسات تحقیقاتی برجسته مانند دانشگاه‌های سینگ‌هوا و شانگ‌های جیاتونگ برخوردار است. هر چهار «ببر هوش مصنوعی» توسط اعضای هیئت‌علمی یا فارغ‌التحصیلان دانشگاه سینگ‌هوا تأسیس شده‌اند. توانمندی چین در حوزه استعدادهای هوش مصنوعی در پیشتازی این کشور در زمینه ثبت اختراعات مشهود است؛ بین سال‌های ۲۰۱۹ تا ۲۰۲۳، بیش از ۷۶۰۰۰ ثبت اختراع هوش مصنوعی توسط چین در سازمان جهانی مالکیت فکری (WIPO) به ثبت رسیده است. رقمی تقریباً چهار برابر ایالات متحده. چین همچنین زیرساخت‌های گسترده مراکز داده با ظرفیت تقریبی ۲۰ گیگاوات دارد و از سرمایه قابل توجهی از طریق بودجه تحقیق و توسعه عمومی بهره‌مند است که بین سال‌های ۲۰۱۶ تا ۲۰۲۳ به طور میانگین، ۵۰ میلیارد دلار در سال بوده است. سرمایه‌گذاری‌های خطرپذیر دولتی نیز تعهد چین به تقویت نوآوری هوش مصنوعی از طریق سرمایه‌گذاری عمومی را نشان می‌دهد. صندوق‌های دولتی خطرپذیر نزدیک به یک‌چهارم کل سرمایه خود را در شرکت‌های هوش مصنوعی سرمایه‌گذاری کرده‌اند و از سال ۲۰۰۰ تاکنون، بیش از ۱۸۰ میلیارد دلار به این حوزه اختصاص داده‌اند که حدود ۱۰۰ میلیارد دلار آن بین سال‌های ۲۰۱۹ تا ۲۰۲۳ بوده است. علاوه بر این، سرمایه‌گذاران خصوصی، چه داخلی و چه خارجی، پیشرفت ببرهای هوش مصنوعی را به شکل فعالی پیش برده‌اند. این شرکت‌ها به همراه 01.AI، طی سال‌های اخیر، بیش از ۶ میلیارد دلار جذب سرمایه کرده‌اند، ازجمله از ابرمقیاس‌گرهای چینی مانند علی‌بابا. محدودیت‌های فعلی چین در دسترسی به تراشه‌های پیشرفته برای آموزش و استنتاج مدل‌های هوش مصنوعی، احتمالاً پیشرفت را به تأخیر می‌اندازد اما متوقف نمی‌کند. گزارش‌ها حاکی از آن است که بسیاری از شرکت‌های چینی همچنان به تراشه‌ها دسترسی دارند و انویدیا مدل‌هایی طراحی کرده که مشمول محدودیت‌های تجاری ایالات متحده نیستند. شرکت‌های چینی در سال ۲۰۲۳، پنج میلیارد دلار تراشه از این نوع خریداری کرده‌اند. علاوه بر این، چین به طور جدی در حال سرمایه‌گذاری در ظرفیت تولید تراشه‌های داخلی خود است. دولت چین حدود ۴۰ میلیارد دلار در این زمینه تعهد سرمایه‌گذاری داده و شرکت هوآوی نیز اخیراً، تراشه Ascend 910 خود را که برای صرف نظر از اینکه هوآوی بتواند با انویدیا در مرز فناوری سخت‌افزار رقابت کند یا خیر (و پس از آن در حوزه نرم‌افزاری)، این شرکت به تازگی به نقطه‌عطفی دست یافته و توانسته یک مدل زبان بزرگ پرقدرت را از شرکت iFlytek به نام Xinghuo4 را به طور کامل بر سکوی Ascend خود آموزش دهد. البته شرکت‌های چینی، جاه‌طلبی‌هایی فراتر از بازار داخلی دارند. به عنوان مثال، علی‌بابا در حال گسترش مراکز داده خود در مالزی، فیلیپین، تایلند، کره‌جنوبی و مکزیک است و در حال افزایش دامنه دسترسی خانواده مدل‌های مولد Qwen خود است.

کاهش چشمگیر فاصله LLM‌های برترچینی با نمونه‌های آمریکایی

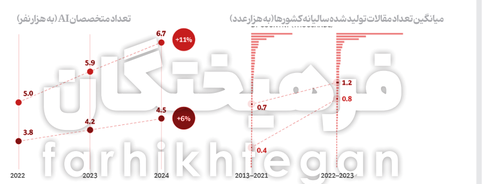


منبع: Stanford's HELM Leaderboard, BCG Henderson Institute

مج‌اندازی عربستان و امارات

به دلیل هزینه بالای ایجاد یک بوم‌سازگان قوی GenAI، برخی کشورها می‌توانند از مالکیت فکری عمومی استفاده کنند و به صورت راهبردی در توسعه استعدادها و توان محاسباتی سرمایه‌گذاری کنند. این کشورها می‌توانند با تکیه بر دسترسی قوی به منابع مالی، توسعه مدل‌ها را بیرون‌سپاری کرده و همچنان مالکیت و اجرای گسترده آن‌ها را در اختیار داشته باشند. این رویکرد برای اقتصادهایی مانند عربستان سعودی و امارات متحده عربی که دسترسی متمرکز به منابع عظیم مالی مانند صندوق‌های ثروت ملی دارند، مناسب است. در منطقه خلیج فارس، عربستان سعودی و امارات متحده عربی، سرمایه‌گذاری‌های قابل توجهی برای تنوع‌بخشی به اقتصادهای خود و تسریع فناوری در منطقه انجام داده‌اند. تعهد بی‌سابقه سرمایه‌گذاری مستقیم در هوش مصنوعی کاملاً مشهود است. امارات، صندوق سرمایه‌گذاری خطرپذیر ۱۰ میلیارد دلاری برای هوش مصنوعی راه‌اندازی کرده است که بیش از هشت برابر مجموع سرمایه دریافتی تا به امروز توسط شرکت‌های مستقر امارات است. عربستان سعودی نیز برنامه دارد تا ۴۰ میلیارد دلار در توسعه هوش مصنوعی سرمایه‌گذاری کند که این مبلغ علاوه بر برنامه جدید سعودی‌ها برای سرمایه‌گذاری ۱۰۰ میلیارد دلار در گسترش مراکز داده (پروژه ترنسدنس) و صندوق فناوری ۱۰۰ میلیارد دلاری دیگر آن‌هاست (با نام پروژه آلاک). این سرمایه‌گذاری‌ها از منابع بسیار بزرگ‌تری تأمین می‌شوند؛ از میان ۲۰ صندوق ثروت ملی بزرگ جهان، پنج صندوق در امارات مستقرند که مجموعاً حدود ۲ تریلیون دلار دارایی تحت مدیریت دارند، در حالی که صندوق سرمایه‌گذاری عمومی عربستان (PIF) حدود ۹۲۰ میلیارد دلار را مدیریت می‌کند. هر دو کشور به شدت در ایجاد بوم‌سازگان تحقیق و توسعه هوش مصنوعی از طریق دانشگاه‌ها و استارتاپ‌های پشتیبانی شده توسط صندوق‌های ثروت ملی، مانند G42 در امارات، سرمایه‌گذاری کرده‌اند. از سال ۲۰۲۲، تعداد متخصصان هوش مصنوعی در امارات و عربستان سعودی، به ترتیب ۳۶ و ۱۷ درصد افزایش یافته است و مهاجرت خالص استعدادهای هوش مصنوعی به این کشورها نیز به ترتیب ۴۰ و ۷۰ درصد رشد را نشان می‌دهد.

استعدادهای جدید عامل پیشرفت عربستان و امارات در تحقیقات AI



منبع: SCImago Journal & Country Rank, BCG Global Talent Tracker

زیرساخت مراکز داده در منطقه خلیج فارس در حال حاضر در مقیاس مطلق کوچک است و ظرفیت هر کدام در عربستان سعودی و امارات حدود ۰/۲ گیگاوات است. این ظرفیت به سرعت در حال رشد است و بخشی از آن به دلیل هزینه‌های پایین انرژی در منطقه است. هزینه انرژی در این منطقه به طور میانگین ۳۰ تا ۵۰ درصد کمتر از ایالات متحده است. اگرچه نیز تسهیل محدودیت‌های صادرات تراشه‌های پیشرفته به امارات- که دولت عربستان انتظار دارد در آینده شامل این کشور نیز شود- به دسترسی پایدار به سخت‌افزارهای بهینه برای وظایف هوش مصنوعی کمک خواهد کرد، با این حال، سرعت واقعی گسترش مراکز داده همچنان مسئله‌ای با پایان باز است. پیشرفت‌های حاصله تاکنون قابل‌توجه بوده است. آرامکوی عربستان، با بر گزارش‌ها، بزرگ‌ترین مدل صنعتی GenAI جهان را تولید کرده است و سازمان داده و هوش مصنوعی عربستان نیز مدل زبان بزرگ ALLaM را ارائه داده که بزرگ‌ترین نسخه‌های آن بر اساس Llama-2 از شرکت متا ساخته شده‌اند. دو مدل فالکون (Falcon) از مؤسسه نوآوری فناوری امارات، همچنان در میان مدل‌های برتر جهان قرار دارند و G42 امارات، خانواده مدل‌های عربی جیس (Jais) را توسعه داده است که توانایی دسترسی به مجموعه‌ای قدرتمند از داده عربی برای آموزش مدل‌ها را نشان می‌دهد. سؤال اصلی این است که آیا عربستان سعودی و امارات می‌توانند خط تولید استعدادهای هوش مصنوعی لازم را برای ایجاد بوم‌سازگان خودکفای توسعه مدل حفظ کنند؟ طبق رقم رشد چشمگیر در این زمینه با حدود ۵ الی ۷ هزار متخصص هوش مصنوعی در امارات و عربستان سعودی به ترتیب، استخر استعدادهای این منطقه در مقایسه با کشورهایی مانند آلمان (۵۵ هزار) یا فرانسه (۵۰ هزار) همچنان کوچک است. از همه مهم‌تر، برای اینکه عربستان سعودی و امارات به تأمین‌کنندگان سوآور GenAI تبدیل شوند، باید به بازار های صادرات فناوری بزرگی دست یابند تا نیازهای سرمایه‌ای فزاینده توسعه مدل‌های هوش مصنوعی را توجیه کنند. به نظر می‌رسد G42 امارات دقیقاً در حال انجام این کار با مدل زبان بزرگ نانادا (Nanda) است که برای مصرف‌کنندگان هندی‌زبان در هندوستان طراحی شده است.

توکیو و سئول کجا هستند؟

برخی کشورها قابلیت‌هایی در زمینه توانمندسازهای اصلی هوش مصنوعی مولد دارند، اما از مقیاس لازم برخوردار نیستند. برای این کشورها که شامل کره جنوبی و ژاپن می‌شود، سرمایه‌گذاری برای دستیابی به مقیاس می‌تواند رقابت‌پذیری آن‌ها را حداقل در سطح داخلی یا منطقه‌ای افزایش دهد. کره جنوبی و ژاپن همچنین دارای بازارهای داخلی بزرگ و منابع مالی قابل توجهی برای تأمین سرمایه‌گذاری‌های مورد نیازند. مسیر این کشورها با توجه به چالش‌هایی که پیش رو دارند، کمتر قابل‌پیش‌بینی است، اما قوت آن‌ها در برخی توانمندسازی‌های کلیدی نشان می‌دهد که هنوز زود است آن‌ها را کنار گذاشت.

هر دو کشور دارای بوم‌سازگان‌های فناوری قوی و متمرکزی‌اند که توسط شرکت‌های بزرگ سرمایه‌گذار در تحقیق و توسعه (R&D) هدایت می‌شوند. هزینه سالانه تحقیق و توسعه ۲۰ شرکت بزرگ فناوری در کره جنوبی و ژاپن به ترتیب ۲۸ و ۲۶ میلیارد دلار است که بیشتر از دیگر قدرت‌های میانی GenAI مانند آلمان (۱۴ میلیارد دلار) و فرانسه (۹ میلیارد دلار)، است. این تمرکز بر تحقیق و توسعه به عملکرد قوی در ثبت اختراعات منجر شده است؛ کره جنوبی و ژاپن از سال ۲۰۱۳ به ترتیب در جایگاه اول و سوم در میان قدرت‌های میانی در زمینه ثبت اختراعات هوش مصنوعی قرار دارند. این دو کشور همچنین از موقعیت مهمی در زنجیره ارزش سخت‌افزار بهره‌مندند که به آن‌ها دسترسی قابل‌اعتماد به تراشه‌های باکیفیت را می‌دهد. کره حدود ۳۰ درصد سهم بازار در تولید تراشه‌های پیشرفته (کمتر از ۱۰ نانومتر) را دارد، هرچند در زمینه تراشه‌های پیشرو از انویدیا عقب است و انتظار می‌رود سهم بازار آن در سال‌های آینده کاهش یابد. هر دو کشور همچنین از صادرکنندگان پیشرو در زمینه مواد حیاتی برای ساخت تراشه‌اند؛ کره جنوبی در تولید ایندیم و ژاپن در زمینه‌هایی چون پردازش فوتورزیست، تجهیزات حذف مواد و تمیز کردن، تجهیزات خودکارسازی تولید و آرسنیک.

اگرچه تاکنون هیچ‌کدام از این کشورها مدل‌های زبان بزرگی در سطح جهانی تولید نکرده‌اند، اما تلاش‌هایی در زمینه توسعه مدل‌های پایه در آن‌ها قابل مشاهده است. در کره جنوبی، شرکت‌هایی مانند سامسونگ و نابور (Naver) مدل‌های زبان بزرگ خود را توسعه داده‌اند و شرکت مخابراتی KT Corp نیز با استفاده از تراشه‌های ساخت داخل، مدل‌های خود را ساخته است. نابور، به‌عنوان بزرگ‌ترین موتور جست‌وجوی اینترنتی در کشور، از دسترسی به داده‌های فراوان سود می‌برد (کره جنوبی یکی از چهار کشوری است که گوگل در آن بازار جست‌وجوی اینترنت را تحت سلطه ندارد.) در ژاپن، همکاری میان دانشگاه‌ها و شرکت‌های خصوصی اخیراً منجر به انتشار مدل منبع‌باز Fugaku-LLM شده است که در معیارهای زبان ژاپنی، عملکرد قدرتمندی نشان داده و با استفاده از ابررایانه رایکن فوگاکو (Riken Fugako) آموزش دیده است. سایرین مانند شرکت راکوتن (Rakuten)، از مدل‌های منبع باز خارجی مانند لاما و می‌سترال استفاده کرده‌اند که برای زبان ژاپنی بهینه‌سازی شده‌اند.

برای توجیه سرمایه‌گذاری‌های لازم، کره جنوبی و ژاپن باید بازارهای هدف کافی برای هوش مصنوعی مولد خود تضمین کنند. کره جنوبی در این زمینه نسبت به ژاپن که اقتصادش تقریباً ۲/۵ برابر بزرگ‌تر است، مزیت کمتری دارد. اما به نظر می‌رسد توسعه‌دهندگان کره‌ای به‌طور تهاجمی‌تری در جهت تأمین بازارهای صادراتی برای هوش مصنوعی مولد حرکت می‌کنند. برای مثال نابور با عربستان سعودی برای توسعه مشترک یک مدل زبان بزرگ عربی همکاری می‌کند و شرکت‌های دیگر کره‌ای، مدل‌های تخصصی برای بازارهای تایلند، ویتنام و مالزی می‌سازند.

ژاپن دارای بازار بزرگ‌تر و زیرساخت مراکز داده بیشتری است که ظرفیت آن در سال ۲۰۲۴ به ۹/۱ گیگاوات می‌رسد. این ظرفیت با آلمان برابر است و بسیار جلوتر از کره جنوبی با ۰/۹ گیگاوات است. هرچند بوم‌های ساخت‌وساز می‌توانند ظرفیت کره جنوبی را در سال‌های آینده دو برابر کنند و از قیمت انرژی نسبتاً پایین این کشور بهره ببرند اما هنوز مشخص نیست آیا توسعه‌دهندگان مدل‌های زبانی قادر به رقابت به‌عنوان ذینفعان تأمین‌کنندگان GenAI با شرکت‌های بزرگ آمریکایی خواهند بود یا خیر؛ شرکت‌هایی که بسیاری از آن‌ها سرمایه‌گذاری برای گسترش حضور خود در ژاپن را افزایش داده‌اند. تاکنون هیچ قهرمان آشکاری برای توسعه مدل‌های هوش مصنوعی مولد داخلی در ژاپن ظاهر نشده است.

در کره جنوبی، جایی که شرکت‌های نابور