به نام خدا

آینده هوش مصنوعی: مسیر پیش‌رو

(چه اتفاقی قراراست رخ دهد؟!)

نویسنده:

سعید جوی‌زاده

آینده هوش مصنوعی: مسیر پیش‌رو

نویسنده: سعید جوی‌زاده

ناشر: انتشارات پارسیس

نوبت و تاریخ چاپ: اول/ ۱۴۰۴

طراح جلد و صفحه‌آرا: تیم طراحی پارسیس

قطع: وزیری / ۳۱۶ صفحه

تیراژ: ۱۰۰۰ نسخه

شابک: ۰-۰-۸۸۸۳۷-۶۲۲-۹۷۸

قیمت: ۰۰٫۰۰۰ تومان

دفتر مرکزی: تهران ،بلوار کشاورز،کبکانیان،کوچه ۵ ،پلاک ۱

اهواز- کیانپارس خیابان ۸ غربی فاز سوم پلاک ۱۱۸

تلفن: ۰۹۳۸۲۲۵۲۷۷۴

وب سایت: **www.biabook.com** - پست الکترونیک: **saeedjavizadeh@gmail.com**

پیشگفتار

در آستانه عصر هوش مصنوعی

بشریت همواره در آستانه تحولات عظیم تکنولوژیکی قرار داشته است؛ از اختراع ابزارآلات سنگی تا انقلاب کشاورزی، از ماشین بخار تا اینترنت، هر جهش فناورانه بنیان‌های اجتماعی، اقتصادی و سیاسی ما را دگرگون کرده است. کتاب «آینده هوش مصنوعی» اثر سعید جوی‌زاده، این سفر تکاملی را در پرتو جدیدترین و شاید قدرتمندترین نیروی تحول‌آفرین زمانه ما، یعنی هوش مصنوعی، مورد کاوش قرار می‌دهد. این کتاب، نه تنها به بررسی پتانسیل‌های بی‌نظیر هوش مصنوعی برای پیشرفت بشری می‌پردازد، بلکه چالش‌ها، مخاطرات و پرسش‌های بنیادینی را نیز مطرح می‌کند که آینده ما را شکل خواهند داد. در عصری که هوش مصنوعی نه فقط یک ابزار جدید برای گسترش توانایی‌های ما، بلکه آینه‌ای برای انعکاس ماهیت انسان، اراده آزاد و مسیرهای آینده جامعه ماست، این کتاب یک نقشه راه ضروری برای فهم این دگرگونی بی‌سابقه ارائه می‌دهد.

فصل اول: آینده هوش مصنوعی: قرار است چه اتفاقی رخ دهد؟

این فصل با معرفی هوش مصنوعی به عنوان نیرویی تحول‌آفرین، به بررسی نقطه عطفی می‌پردازد که در اواخر سال ۲۰۲۲ و اوایل ۲۰۲۳ با ظهور مدل‌های زبانی بزرگ (LLMs) مانند ChatGPTایجاد شد. این رویداد، که بشریت را وارد گفت‌وگویی بی‌سابقه با هوش مصنوعی کرد، نه تنها یک نقطه عطف در توسعه هوش مصنوعی بود، بلکه آغازگر بحث‌های گسترده‌ای درباره پتانسیل، خطرات و جایگاه آن در زندگی انسان شد. نویسنده هوش مصنوعی را با انقلاب صنعتی مقایسه کرده و پتانسیل آن را برای تقویت عاملیت انسانی تا سطح »ابرعاملیت (Superagency) « نشان می‌دهد. فصل اول همچنین به مبانی فنی و محدودیت‌های LLMها، از جمله پدیده

»توهمات (hallucinations) «و »جعبه سیاه(black box) «، و همچنین پیامدهای عمیق هوش مصنوعی بر عاملیت انسانی، شغل، حریم خصوصی و ماهیت واقعیت در عصر دیجیتال می‌پردازد. در نهایت، بر اهمیت همکاری جهانی و مشارکت گسترده عمومی برای توسعه مسئولانه و فراگیر هوش مصنوعی تأکید می‌شود.

فصل دوم: دانش بزرگ: از سایه اورول تا عصر هوش مصنوعی و ساختارهای اعتماد دیجیتال

این فصل با نگاه پیشگویانه جورج اورول در رمان «۱۹۸۴» آغاز می‌شود که تصویری هراس‌انگیز از نظارت دولتی و خردکننده ارائه داد و بنیان نگرانی‌های عمومی درباره فناوری را گذاشت. در ادامه، به ظهور کامپیوترهای بزرگ (مین‌فریم‌ها) در دهه ۱۹۶۰ و واکنش‌های دوگانه نسبت به آن – از وحشت از »جامعه عریان «و کنترل کامل اطلاعات تا رویای »عصر روشنگری جدید – «پرداخته می‌شود. نویسنده سیر تحول دیدگاه‌ها نسبت به جمع‌آوری داده‌ها، از شکست طرح »مرکز داده فدرال «به دلیل نگرانی‌های حریم خصوصی در دهه ۱۹۶۰، تا گسترش کامپیوترهای شخصی و اینترنت که به‌جای سرکوب، به توانمندسازی فردی و گسترش هویت عمومی منجر شد، را بررسی می‌کند. پلتفرم‌هایی مانند لینکدین نشان دادند که چگونه اعتماد مبتنی بر هویت عمومی و شبکه‌های ارتباطی می‌تواند به خلق ارزش جمعی و »دانش بزرگ «منجر شود. در نهایت، تأکید می‌شود که هوش مصنوعی، با وجود خطرات احتمالی، ابزاری حیاتی برای تبدیل »داده‌های بزرگ (Big Data) «به »دانش بزرگ (Big Knowledge) «و افزایش توانمندی‌های انسانی برای ورود به عصری نوین است.

فصل سوم: چه اتفاق خوبی ممکن است رخ دهد؟

این فصل به بررسی دو رویکرد متضاد در مواجهه با پیشرفت‌های تکنولوژیک، یعنی »راه‌حل‌گرایی (solutionism) «و »مشکل‌گرایی(problemism) «، می‌پردازد. در حالی که

انتقادات سازنده از فناوری می‌تواند ارزشمند باشد، تمرکز صرف بر خطرات احتمالی، فرصت‌های بی‌شماری را که ممکن است از نوآوری حاصل شود، نادیده می‌گیرد. نویسنده استدلال می‌کند که هوش مصنوعی پتانسیل فوق‌العاده‌ای برای حل مبرم‌ترین چالش‌های جهانی از جمله تولید انرژی پایدار، مراقبت‌های بهداشتی، آموزش و امنیت سایبری دارد. تمرکز اصلی این فصل بر پتانسیل تحول‌آفرین هوش مصنوعی در حوزه سلامت روان است، جایی که بحران فزاینده‌ای ناشی از کمبود متخصصان، دسترسی محدود و ناکارآمدی سیستماتیک وجود دارد. پلتفرم‌هایی مانند »کوکو «و تحقیقات بر روی مدل‌های هوش مصنوعی مانند Woebot و Wysa نشان داده‌اند که چگونه هوش مصنوعی می‌تواند دسترسی به درمان‌های سلامت روان را به‌طور چشمگیری گسترش داده، آن‌ها را شخصی‌سازی کرده و کارایی‌شان را افزایش دهد، تا مراقبتی »به سبک اسپاتیفای و نتفلیکس «ارائه دهد. در نهایت، با ترویج دیدگاهی موسوم به »چه اتفاق خوبی ممکن است رخ دهد؟ «و مفهوم »فوق‌بشری«، این فصل استدلال می‌کند که تعامل با هوش مصنوعی می‌تواند نه تنها مشکلات را حل کند، بلکه ظرفیت انسانی برای همدلی و مهربانی را نیز افزایش داده و به یک جامعه بهتر و انسانی‌تر منجر شود.

فصل چهارم: پیروزی مشاعات خصوصی: خلق ارزش و پتانسیل بی‌نهایت در عصر هوش مصنوعی

این فصل به بررسی عمیق پدیده »مشاعات خصوصی (private commons) «می‌پردازد. با وجود انتقادهایی مانند »سرمایه‌داری نظارتی (surveillance capitalism) «شوژانا زوبوف و تمرکز ثروت در دستان شرکت‌های بزرگ فناوری، نویسنده استدلال می‌کند که این روایت اغلب بخش عظیمی از ارزش خلق‌شده برای کاربران و جامعه را نادیده می‌گیرد .مشاعات خصوصی، پلتفرم‌های دیجیتالی هستند که تحت اداره یا مالکیت خصوصی قرار دارند، اما کاربران را به عنوان »تولیدکننده «و »ناظر «درگیر می‌کنند و خدمات رایگان یا تقریباً رایگان ارائه می‌دهند که به مثابه زیرساخت‌ها و خدمات اجتماعی خصوصی عمل می‌کنند. مفهوم »مازاد مصرف‌کننده «

(consumer surplus)نشان می‌دهد که کاربران از خدمات رایگان یا کم‌هزینه دیجیتال، ارزشی به‌مراتب بیشتر از آنچه شرکت‌ها کسب می‌کنند، دریافت می‌نمایند. این فصل تأکید می‌کند که داده‌ها، برخلاف منابع فیزیکی، »غیررقابتی (non-rivalrous) «هستند و با افزایش استفاده، نه تنها کاهش نمی‌یابند بلکه ارزششان بیشتر می‌شود و پلتفرم‌ها را از دام »تراژدی مشاعات «سنتی رها می‌سازد. با ورود هوش مصنوعی و به‌ویژه مدل‌های چندوجهی(multimodal models) ، قابلیت‌های این مشاعات خصوصی به سطحی بی‌سابقه ارتقا یافته، امکانات جدیدی برای تعامل فراهم می‌کند و نویدبخش «هوش شبکه‌ای جهانی» و فراوانی گسترده اجتماعی است.

فصل پنجم: آزمون، محک و اعتماد در هوش مصنوعی: فراتر از مسابقه تسلیحاتی، به سوی پیشرفتی شفاف و مسئولانه

این فصل روایت رسانه‌ای »مسابقه تسلیحاتی هوش مصنوعی «را به چالش می‌کشد و استدلال می‌کند که توسعه هوش مصنوعی بیش از آنکه شبیه به مسابقه تسلیحاتی باشد، به »مسابقه فضایی «دهه میانی قرن بیستم شباهت دارد. توسعه هوش مصنوعی یک فرآیند مبتنی بر داده، با مشارکت گسترده، شفافیت و مهم‌تر از همه، آزمون و محک مداوم است. از ریشه‌های تاریخی آزمون تورینگ تا تکامل معیارهای سنجش عملکرد) محک‌ها(، توسعه هوش مصنوعی همواره بر ارزیابی دقیق و داده‌محور بنا شده است .محک‌ها نه تنها ابزارهایی برای مقایسه و ارتقای فنی هستند، بلکه نقشی پویا در ایجاد هنجارهای شفافیت و مسئولیت‌پذیری ایفا می‌کنند که فراتر از مقررات سنتی و ایستا است. فصل همچنین چالش‌هایی مانند »مثال‌های خصمانه « (adversarial examples)و »شکنندگی مدل (model brittleness) «را بررسی می‌کند. در نهایت، پلتفرم »چت‌بات آرنا (Chatbot Arena) «به عنوان الگویی برای حکمرانی مردمی معرفی می‌شود، که در آن اعتماد از طریق شفافیت و بازخورد جمعی کاربران شکل می‌گیرد.

فصل ششم: نوآوری به مثابه ایمنی: تعادل میان توسعه سریع و مدیریت مخاطرات در عصر هوش مصنوعی

این فصل به بررسی تعامل مفاهیم نوآوری و ایمنی در توسعه فناوری می‌پردازد و استدلال می‌کند که نوآوری نه تنها با ایمنی در تضاد نیست، بلکه خود بهترین راهبرد برای افزایش ایمنی است. نویسنده دو دیدگاه غالب در این زمینه، یعنی »اصل احتیاطی (precautionary principle)« که بر کنترل پیشگیرانه تأکید دارد، و »نوآوری بدون اجازه (innovation without permission)« که فضایی برای آزمایش و تطبیق فراهم می‌کند، را مقایسه می‌کند. با ارائه مثال‌های تاریخی از صنعت خودروسازی (مانند استارت برقی) و توسعه اینترنت (مانند ماده ۲۳۰ قانون ارتباطات)، نشان داده می‌شود که رویکردهای توسعه تکرارشونده و بازخورد محور، به‌طور مؤثرتری به ایجاد محصولات ایمن‌تر و قابل اعتمادتر منجر می‌شوند. این فصل همچنین بر این ایده تأکید می‌کند که »تهدید وجودی وضعیت موجود (existential threat of the status quo)« – یعنی تعویق نوآوری – می‌تواند منجر به تداوم آسیب‌های موجود شود.

فصل هفتم: جی‌پی‌اس اطلاعاتی: نقشه راه هوش مصنوعی

این فصل هوش مصنوعی را به عنوان یک »جی‌پی‌اس اطلاعاتی (Information GPS)« معرفی می‌کند. با آغاز از ریشه‌های سامانه موقعیت‌یاب جهانی (GPS) که از یک پروژه نظامی به یک ابزار عمومی ضروری تبدیل شد، نویسنده نحوه مواجهه دولت‌ها با فناوری‌های تحول‌آفرین را واکاوی می‌کند. سپس، به تشریح شباهت‌ها و تفاوت‌های بنیادین میان LLMها و GPS می‌پردازد و نشان می‌دهد که چگونه LLMها به عنوان ابزارهایی برای تجزیه و تحلیل و نقشه‌برداری از جریان‌های زبانی، توانایی‌های ناوبری اطلاعاتی را در اختیار افراد قرار می‌دهند. تأکید اصلی بر دموکراتیک کردن دسترسی به دانش و افزایش توانمندی فردی با LLMهاست. این فصل کاربردهای چندوجهی هوش مصنوعی را برای ارتقای مهارت‌ها، پر کردن شکاف‌های

دانش و کمک به افراد دارای معلولیت، از جمله از طریق هوش مصنوعی عامل (agentic AI) که می‌تواند وظایف پیچیده را با حداقل نظارت انسانی انجام دهد، بررسی می‌کند. اهمیت تعامل مداوم و مبتنی بر گفتگو با سیستم‌های هوش مصنوعی برای بهبود دقت و کاهش سوگیری‌ها نیز مورد بحث قرار می‌گیرد.

فصل هشتم: کد به مثابه قانون: حکمرانی و کنترل در عصر هوش مصنوعی

این فصل به بررسی عمیق مفهوم محوری »قانون کد است (Code is Law) «می‌پردازد که توسط لارنس لسیگ مطرح شد. لسیگ استدلال کرد که در اینترنت اولیه، معماری (کد) نقش نامتناسبی در تنظیم رفتار انسان ایفا می‌کرد. نویسنده نشان می‌دهد که چگونه تجاری‌سازی اینترنت و نیازهای متعاقب آن برای احراز هویت، به تغییر معماری دیجیتال به سمت هویت، تمرکزگرایی و کنترل بیشتر منجر شد. این فصل مفهوم »کنترل کامل (full control) «را بررسی می‌کند که توسط فناوری‌های پیشرفته امکان‌پذیر شده است، جایی که کد و هوش مصنوعی در خودروها، لوازم خانگی و زیرساخت‌های شهری نفوذ کرده و بر زندگی روزمره حکم می‌رانند. تفاوت‌های اساسی میان سازوکارهای سنتی اجرای قانون (مبتنی بر انتخاب و رعایت داوطلبانه) و کنترل معماری (مبتنی بر کد) و پدیده »ضدقراردادها (anti-contracts) «و قراردادهای هوشمند مبتنی بر بلاک‌چین نیز مورد بحث قرار می‌گیرد. در نهایت، بر نقش حیاتی »رضایت حکومت‌شوندگان (consent of the governed) «در مشروعیت‌بخشی به سیستم‌های هوش مصنوعی و ادغام ایمن و مؤثر آن‌ها در جامعه تأکید می‌شود.

فصل نهم: آزادی: اجماع ملی و اراده جمعی در عصر هوش مصنوعی

این فصل به بررسی مفهوم پویای آزادی می‌پردازد و استدلال می‌کند که آزادی نه یک اصل ثابت، بلکه مفهومی در حال تحول است که به‌شدت تحت تأثیر پیشرفت‌های تکنولوژیکی، چارچوب‌های نظارتی و نیازهای جمعی قرار دارد. با الهام از تاریخچه صنعت اتومبیل‌رانی، این

فصل نشان می‌دهد که چگونه نوآوری، با وجود گسترش اولیه خودمختاری فردی، به‌تدریج نیازمند مقررات و زیرساخت‌های مشترک شد تا به افزایش ایمنی، کارایی و دسترسی گسترده‌تر همراه گردد. این فصل همچنین به بررسی نمونه‌هایی مانند اختراع چاپ و مدیریت بحران کووید-۱۹ توسط کره جنوبی می‌پردازد تا نشان دهد که چگونه فناوری‌های انقلابی، با وجود گسترش اولیه آزادی بیان، به‌تدریج نیازمند چارچوب‌های نظارتی جدیدی برای مدیریت ریسک‌ها و چالش‌های اجتماعی شدند. مفهوم »خودمختاری شبکه‌ای (networked autonomy)« در اینجا مطرح می‌شود که آزادی واقعی اغلب در یک سیستم وابسته و پیوسته که در آن اقدامات جمعی و زیرساخت‌های مشترک آن را تقویت می‌کنند، بهتر شکوفا می‌شود.

فصل دهم: آینده در آمریکا: نقشه راه هوش مصنوعی برای تقویت ملت

این فصل به بررسی عمیق پیامدهای تحول‌آفرین هوش مصنوعی بر آینده ملت‌ها و شهروندان می‌پردازد و به‌طور خاص، موقعیت ایالات متحده را در این چشم‌انداز جدید مورد ارزیابی قرار می‌دهد. با الهام از درس‌های تاریخی جنبش لودیت‌ها در انگلستان و داستان حزب دونر در آمریکا، نویسنده نشان می‌دهد که مقاومت یا احتیاط مفرط در برابر نوآوری می‌تواند به انزوای ملی، رکود اقتصادی و از دست دادن عاملیت فردی منجر شود، در حالی که پذیرش استراتژیک و مسئولانه فناوری می‌تواند مسیر را برای رفاه بی‌سابقه، پیشرفت اجتماعی و تقویت هویت ملی هموار کند. مفهوم »هوش مصنوعی ملی (Sovereign AI)« مطرح می‌شود که بر اهمیت حیاتی کنترل کشورها بر داده‌ها و فناوری هوش مصنوعی خود برای حفظ حریم خصوصی، امنیت داده و منافع ملی تأکید دارد. در نهایت، با تأکید بر نقش هوش مصنوعی در تحول خدمات دولتی به »دولت ۲.۰ (Government 2.0)« و تقویت مشارکت شهروندی، این فصل چشم‌اندازی آینده‌نگرانه ارائه می‌دهد که در آن هوش مصنوعی نه تنها به عنوان یک ابزار، بلکه به عنوان یک نیروی تحول‌آفرین برای خیر جمعی عمل می‌کند و به افراد امکان دستیابی به »فراعاملیت (Superagency)« را می‌دهد.

این کتاب از طریق تحلیل عمیق و مثال‌های تاریخی و معاصر، خواننده را به یک گفت‌وگوی مستمر و مشارکتی درباره آینده هوش مصنوعی دعوت می‌کند. این یک فراخوان است برای رهبران، سیاست‌گذاران، توسعه‌دهندگان و عموم مردم تا با ذهنیتی آینده‌نگر، مسئولانه و مبتنی بر اجماع ملی، مسیر را برای آینده‌ای روشن و توانمندساز با هوش مصنوعی هموار سازند.

سعید جوی زاده

فهرست

فصل ۱

بشریت وارد گفتگو شد...

"هوش مصنوعی نه تنها ابزار جدیدی برای گسترش توانایی‌های ماست، بلکه آینه‌ای است که در آن ماهیت انسان، اراده آزاد و مسیرهای آینده جامعه خود را منعکس می‌کنیم." – سعید جوی زاده

اهداف اصلی فصل:

- بررسی چگونگی ظهور هوش مصنوعی به عنوان نیروی تحول‌آفرین در بحبوحه تحولات اقتصادی و فناوری اواخر سال ۲۰۲۲.
- تحلیل تأثیرات فوری و گسترده مدل‌های زبانی بزرگ، به ویژهChatGPT ، بر صنعت فناوری و جامعه جهانی.
- توضیح مبانی فنی و محدودیت‌های ذاتی مدل‌های زبانی بزرگ، از جمله پدیده‌هایی مانند "توهمات" و چالش‌های "جعبه سیاه."
- کاوش در پیامدهای عمیق هوش مصنوعی بر عاملیت انسانی، شغل‌ها، حریم خصوصی و ماهیت واقعیت در عصر دیجیتال.
- تبیین چشم‌اندازهای مختلف و گفتمان‌های جاری پیرامون توسعه و تنظیم هوش مصنوعی، و ارائه راهبردهایی برای توسعه مسئولانه و همه‌شمول.

چکیده

این فصل به بررسی نقطه عطفی می‌پردازد که هوش مصنوعی در اواخر سال ۲۰۲۲ و اوایل ۲۰۲۳، با ظهور مدل‌های زبانی بزرگ مانندChatGPT ، در صحنه جهانی ایجاد کرد. در حالی که جهان درگیر پیامدهای اقتصادی پس از همه‌گیری و چالش‌های صنعت فناوری بود، معرفی ناگهانی این سیستم‌های هوش مصنوعی تحول عظیمی را رقم زد. این فصل ابتدا زمینه اقتصادی و فناوری آن دوره را تشریح می‌کند، سپس به استقبال بی‌سابقه از ChatGPT و واکنش‌های صنعت فناوری می‌پردازد. در ادامه، ساختار درونی مدل‌های زبانی بزرگ، نحوه عملکرد و محدودیت‌های آن‌ها، از جمله تولید اطلاعات نادرست ("توهمات") و چالش‌های شفافیت، تشریح می‌شود. بخش مهمی از فصل به بررسی تأثیرات هوش مصنوعی بر عاملیت انسانی می‌پردازد و آن را با انقلاب صنعتی مقایسه می‌کند تا پتانسیل هوش مصنوعی را به عنوان یک "ابزار" برای افزایش توانایی‌های انسان نشان دهد. در نهایت، رویکردهای مختلف توسعه هوش مصنوعی و نیاز به مشارکت گسترده عمومی برای شکل‌دهی به آینده‌ای که عاملیت فردی و جمعی را تقویت می‌کند، مورد بحث قرار می‌گیرد، با تأکید بر اینکه همکاری جهانی برای عبور از این دوران تحول‌آفرین ضروری است.

مقدمه

اواخر سال ۲۰۲۲، دوره‌ای سرشار از تناقضات و تحولات عمیق جهانی بود. در حالی که جوامع مختلف در تلاش برای بازیابی از پیامدهای همه‌گیری کووید-۱۹ بودند و نگرانی‌ها از تورم سر به فلک می‌کشید و قیمت مواد غذایی در اوج تاریخی خود قرار داشت، نشانه‌هایی از بازگشت به زندگی عادی نیز مشهود بود. رشد مشاغل از تخمین‌ها فراتر می‌رفت و شور و هیجان اجتماعی، مانند تقاضای بی‌سابقه برای بلیت‌های کنسرت، اوج گرفته بود. در نقطه مقابل، صنعت فناوری با چالش‌های جدی روبرو بود؛ کاهش درآمدهای تبلیغاتی، تغییر در احساسات سرمایه‌گذاران و الگوهای در حال تحول تعامل کاربران به رکود منجر شد. شرکت‌های بزرگی چون متا و آمازون اقدام به اخراج‌های گسترده و بی‌سابقه‌ای کردند، و ورشکستگی صرافی ارز دیجیتال FTX به اتهام کلاهبرداری، کل اکوسیستم ارزهای دیجیتال را متزلزل ساخت. این حوادث، روایت غالب دهه ۲۰۱۰ را که قدرت "بیگ تک" را مطلق و توانایی آن‌ها در کنترل بازارها و دستکاری رفتار مصرف‌کننده را بی‌چون و چرا می‌دانست، به چالش کشید. تلاش‌های چند میلیارد دلاری متا در زمینه متاورس نتوانسته بود کاربران زیادی را جذب کند و سرمایه‌گذاری‌های عظیم در بلاک‌چین و ارزهای دیجیتال نیز هنوز نتوانسته بودند این فناوری‌ها را به بخشی جدایی‌ناپذیر از زندگی روزمره تبدیل کنند، آنگونه که وب، جستجو، ایمیل و شبکه‌های اجتماعی پیشتر شده بودند. در این بستر از بی‌ثباتی و ناامیدی در صنعت فناوری، رویدادی به وقوع پیوست که به طور غیرمنتظره‌ای مسیر آینده دیجیتال را تغییر داد. در آخرین روز نوامبر ۲۰۲۲، آزمایشگاه تحقیقاتی OpenAI، بدون هیچگونه اطلاع قبلی یا تبلیغاتی، محصول جدید خود را با نام ChatGPT معرفی کرد و به این ترتیب، بشریت وارد گفتگوی بی‌سابقه‌ای با هوش مصنوعی شد. این رویداد نه تنها نقطه‌ی عطفی در توسعه هوش مصنوعی بود، بلکه آغازگر بحث‌های گسترده‌ای درباره پتانسیل، خطرات و جایگاه آن در زندگی انسان شد.

طلوع عصر جدید: تحولات اقتصادی و فناوری در اواخر ۲۰۲۲

در اواخر سال ۲۰۲۲، جهان شاهد یک دوره‌ی گذار پیچیده بود که با دو جریان اصلی مشخص می‌شد: بازیابی از پیامدهای همه‌گیری و بحران‌های مالی. پس از سال‌ها مبارزه با کووید-۱۹، نگرانی‌های عمومی به سرعت از سلامت به سمت مسائل اقتصادی، به ویژه تورم، معطوف شد. قیمت مواد غذایی در اوج خود قرار داشت و فشارهای اقتصادی بر خانوارها و کسب‌وکارها مشهود بود. با این حال، نشانه‌هایی از بازگشت به "عادی‌سازی" نیز در افق دیده می‌شد؛ آمارهای مربوط به رشد مشاغل و افزایش میانگین دستمزد ساعتی در ماه نوامبر به طور قابل توجهی از پیش‌بینی‌ها فراتر رفت و فعالیت‌های اجتماعی نیز با شور و هیجان از سر گرفته شد. تقاضای بی‌سابقه برای رویدادهایی نظیر تور کنسرت تیلور سویفت، که منجر به فروپاشی سیستم‌های شرکت‌های بلیت‌فروشی شد، نمایانگر این بازگشت به زندگی عمومی بود.

در نقطه مقابل این شور و هیجان اجتماعی، صنعت فناوری با چالش‌های بی‌سابقه‌ای دست و پنجه نرم می‌کرد. کاهش درآمدهای حاصل از تبلیغات دیجیتال، تغییر در رویکردها و احساسات سرمایه‌گذاران، و الگوهای در حال تحول تعامل کاربران، همگی به رکود و نااطمینانی در این بخش دامن زدند. این وضعیت منجر به موجی از تعدیل نیروهای گسترده در شرکت‌های بزرگ فناوری شد. به عنوان مثال، شرکت متا، که قبلاً با نام فیس‌بوک شناخته می‌شد، در تاریخ ۹ نوامبر، ۱۱ هزار نفر از کارکنان خود را اخراج کرد که بزرگترین کاهش نیروی کار در تاریخ این شرکت بود. تنها دو روز بعد، صرافی ارز دیجیتال FTX مستقر در باهاما اعلام ورشکستگی کرد و اتهامات گسترده‌ای از کلاهبرداری و سوءاستفاده از وجوه مشتریان به سرعت مطرح شد که میلیاردها دلار از ارزش بازار ارزهای دیجیتال را محو کرد. در ادامه این اخبار ناگوار، در ۱۴ نوامبر گزارش شد که آمازون نیز همانند متا، اقدام به اخراج‌های بی‌سابقه‌ای کرده و قرار است دست‌کم ۱۰ هزار کارمند خود را در هفته‌های آتی کنار بگذارد.

این پیامدهای نامطلوب تا حدی نتیجه‌ی یک تصحیح طبیعی در بازار بودند. سال‌های همه‌گیری با محرک‌های دولتی و تقاضای سرکوب‌شده مصرف‌کننده، باعث افزایش بی‌سابقه

استخدام، درآمد و ارزش بازاری شرکت‌های فناوری شده بود. اما این رویدادها همچنین به عنوان پاسخی واضح به این روایت غالب عمل کردند که "بیگ تک" قدرت مطلق برای کنترل بازارها و دستکاری رفتار مصرف‌کننده را به اراده خود دارد. در طول دهه ۲۰۱۰، این روایت به یک باور عمومی تبدیل شده بود که شرکت‌هایی مانند آلفابت (گوگل) و متا، با استفاده از طراحی‌های اعتیادآور و تکنیک‌های ربایش توجه، هنر تیره حداکثرسازی تعامل و به دام انداختن میلیون‌ها کاربر در "جهنم‌های دیجیتال" از جمله اسکرول‌های بی‌پایان، توییت‌های خشمگین، پست‌های نامربوط، دنبال کردن زنجیره‌ای اطلاعات، و شرمساری گروهی را به کمال رسانده‌اند.

با این حال، حتی تاکتیک‌های ظاهراً مقاومت‌ناپذیر متا نیز نتوانسته بودند بسیاری از مردم را به متاورس، پروژه چند میلیارد دلاری این شرکت برای ایجاد یک دنیای مجازی وسیع سه‌بعدی برای کار، بازی و تعامل اجتماعی، جذب کنند. در حالی که سرمایه‌گذاران خطرپذیر سیلیکون ولی میلیاردها دلار را به استارتاپ‌های بلاک‌چین و پروژه‌های ارز دیجیتال سرازیر کرده بودند، توسعه‌دهندگان این فناوری‌ها هنوز نتوانسته بودند رمز موفقیت برای تبدیل مالی غیرمتمرکز به یک جزء ضروری برای کاربران عادی را کشف کنند؛ برخلاف آنچه در موج‌های قبلی نوآوری دیجیتال برای وب، جستجو، ایمیل، محاسبات موبایلی، پیام‌رسانی متنی و شبکه‌های اجتماعی به سرعت اتفاق افتاده بود. در این شرایط از ناامیدی و چالش، نقطه عطفی جدید در تاریخ فناوری شکل گرفت.

رونمایی از ChatGPT و تأثیر فوری آن

در حالی که ماه نوامبر برای صنعت فناوری به یک ماه پر از اخبار ناگوار تبدیل شده بود، در آخرین روز آن ماه، OpenAI، یک آزمایشگاه تحقیقاتی مستقر در سان فرانسیسکو با حدود ۳۷۵ کارمند، محصول جدید خود را بدون هیچ‌گونه اطلاع قبلی و بدون تبلیغات گسترده‌ای رونمایی کرد. ساعت ۱۰:۰۲ صبح آن روز، حساب رسمی X.com این آزمایشگاه با انتشار توییتی اعلام کرد: "تلاش کنید باChatGPT ، سیستم هوش مصنوعی جدید ما که برای گفتگو بهینه‌سازی شده

است، صحبت کنید. بازخورد شما به ما در بهبود آن کمک خواهد کردسام آلتمن، هم‌بنیان‌گذار و مدیرعامل OpenAI، در اولین توییت خود درباره ChatGPT، رویکردی محتاطانه داشت و نوشت: «به نظر من رابط‌های زبانی بسیار مهم خواهند شد. با کامپیوتر (صوتی یا متنی) صحبت کنید و آنچه را که می‌خواهید، برای تعریف‌های پیچیده‌تر از "خواستن"، به دست آورید! این یک نمایش اولیه از آنچه ممکن است است (هنوز محدودیت‌های زیادی دارد؛ این یک نسخه تحقیقاتی است).»

همانطور که از نامش پیدا بود، ChatGPTیک چت‌بات بود؛ ابزاری که معمولاً در ذهن بسیاری از مردم جایگاه ویژه‌ای نداشت و کمتر کسی آن را کاندیدای موفقیت‌های بزرگ بعدی می‌دانست. در طول تاریخ وب، چت‌بات‌ها عمدتاً در زمینه‌های خدمات مشتری به کار گرفته می‌شدند و عملکرد آن‌ها اغلب چنان ضعیف بود که یک فرد معمولی که به دنبال راهنمایی برای بازگرداندن یک خرید آنلاین بود، ترجیح می‌داد بیست دقیقه در صف انتظار تلفنی بماند تا با کسی صحبت کند که صدایش از یک مرکز تماس در ته یک استخر روی مریخ می‌آمد.

اما ChatGPT متفاوت بود، بسیار متفاوت. این تفاوت بلافاصله آشکار شد ChatGPT .به شکلی چشمگیر دانش‌آموز، به طرز خیره‌کننده‌ای همه‌کاره و به صورت متقاعدکننده‌ای شبیه به انسان بود. این سیستم می‌توانست توضیحی آسان‌فهم از مکانیک کوانتوم ارائه دهد. می‌توانست غزلیاتی درباره شاخص قیمت مصرف‌کننده بسازد، اگر چنین چیزی خواسته می‌شد. می‌توانست به اشکال‌زدایی – یا شاید هم ایجاد اشکال – در کدهای پایتون کمک کند ChatGPT .همیشه حقایق را درست نمی‌گفت، اما حتی اشتباهاتش، که معمولاً "توهمات" نامیده می‌شدند، موجب شگفتی و کنجکاوی می‌شدند.

بدون هیچگونه سرمایه‌گذاری در بازاریابی، ChatGPTتنها در پنج روز اولین میلیون کاربر خود را جذب کرد. به نظر می‌رسید که تقریباً دو میلیون از آن یک میلیون نفر، روزنامه‌نگار بودند که تجربیات خود را گزارش می‌دادند – و در این زمان بود که علاقه به ChatGPT واقعاً اوج گرفت.

تنها در دو ماه، این سیستم به ۱۰۰ میلیون کاربر دست یافت و آنچنان هیجان، آرزو و ترس از دست دادن (FOMO) را در صنعت فناوری ایجاد کرد که می‌توانست از کمیسیون تجارت فدرال نشان تقدیر دریافت کند.

این رویداد بلافاصله تأثیرات گسترده‌ای بر غول‌های فناوری گذاشت. ساندار پیچای، مدیرعامل آلفابت، یک هشدار "کد قرمز" به تمامی کارمندان گوگل ارسال کرد و اعلام نمود که هوش مصنوعی اکنون اولویت اصلی کل شرکت است. مایکروسافت، که سه سال قبل در OpenAI سرمایه‌گذاری کرده بود، شروع به اشاره به "همکاران (copilots)" بیشتر از یک کتابچه راهنمای آموزش خلبانی کرد. مارک زاکربرگ اعلام کرد که متا یک گروه جدید محصول هوش مصنوعی مولد در سطح بالا ایجاد کرده تا تلاش‌های این شرکت را در این زمینه "تسریع" کند. شرکت‌های تازه‌وارد و نوپا مانندAnthropic ، Midjourney، Hugging Faceو Replika نیز هوش مصنوعی را به گونه‌ای به جلو هل دادند که غول‌ها را در تلاش برای همگام شدن با خودشان قرار دادند. حتی مقالات تحقیقاتی با عناوین پیچیده نیز می‌توانستند به صورت غیررسمی در X.com وایرال شوند، که نشان از تغییر فضای گفتمان علمی و عمومی داشت.

گفتمان در حال تحول پیرامون هوش مصنوعی: از نگرانی‌های ضدانحصار تا دغدغه‌های وجودی در مواجهه با این شرایط جدید، احساسات عمومی و تخصصی نسبت به هوش مصنوعی ۱۸۰ درجه تغییر کرد. سال‌ها پیش از انتشارChatGPT ، بزرگترین منتقدان "بیگ تک" اصرار داشتند که اقدامات ضدانحصار برای تزریق انرژی رقابتی جدید به این بخش فناوری پویای آمریکا ضروری است. اما اکنون همان منتقدان شروع به گفتن این نکته کردند که نوآوری با سرعت بیش از حد در حال وقوع است و از کنترل خارج شده است. در مارس ۲۰۲۳، یک سازمان غیرانتفاعی به نام "مؤسسه آینده حیات" نامه‌ای سرگشاده منتشر کرد که در آن از "همه آزمایشگاه‌های هوش مصنوعی خواسته شد تا فوراً حداقل به مدت ۶ ماه آموزش سیستم‌های هوش مصنوعی قدرتمندتر از GPT-4 را متوقف کنند." بیش از ۳۳ هزار نفر، از جمله بسیاری از رهبران و فناوران

صنعت هوش مصنوعی، این نامه را امضا کردند. حال و هوای آن‌ها ناگوار و فوریتشان محسوس بود. کمیته قضایی سنا نیز به این موضوع توجه کرد و چندین جلسه استماع درباره نظارت بر هوش مصنوعی را در طول سال برگزار کرد.

با این حال، شش ماه گذشت، و سپس شش ماه دیگر. توسعه‌دهندگان به کار خود ادامه دادند و نوآوری‌ها متوقف نشدند .OpenAI به انتشار به‌روزرسانی‌های GPT-4 ، مدل بنیادینی که زیربنای ChatGPT است و در حال دستیابی به عملکردی پیشرفته در وظایف پیچیده و حل مسئله بود و قابلیت تجزیه و تحلیل تصاویر و ارائه بازخورد در مورد آن‌ها را کسب می‌کرد، ادامه داد Claude 2 .از شرکت Anthropic نیز به سطوح جدیدی از دقت واقعی دست یافت و طول زمینه خود را گسترش داد، به این معنی که می‌توانست تا حدود ۷۵ هزار کلمه را پردازش و زمینه را در متون ورودی حفظ کند. اگر کسی می‌خواست نسخه کامل و بدون ویرایش "جنگ دنیاها" نوشته اچ.جی. ولز را در بیست نکته کلیدی خلاصه کند، Claude می‌توانست این کار را انجام دهد. این پیشرفت‌ها، در عین حال که هیجان‌انگیز بودند، دغدغه‌های موجود را نیز تشدید می‌کردند و نشان می‌دادند که سرعت نوآوری چقدر بالا است.

درک مدل‌های زبانی بزرگ: ساختار، عملکرد و محدودیت‌ها

همچنان بسیاری از چالش‌های مداوم باقی مانده بودند. عوامل گفتگومحور مانند ChatGPT و Claude بر پایه مدل‌های زبانی بزرگ (LLMs) ساخته شده‌اند؛ نوع خاصی از ساختارهای یادگیری ماشینی که برای وظایف پردازش زبان طراحی شده‌اند LLMs .مانند GPT-4 ، زبان را با استفاده از معماری شبکه عصبی پردازش و تولید می‌کنند. در این معماری، لایه‌های متعدد گره‌ها (nodes) یک آبشار پیچیده از محاسبات به هم پیوسته را انجام می‌دهند. هر گره در یک لایه، ورودی را از لایه قبلی دریافت می‌کند، عملیات ریاضی را اعمال کرده و نتیجه را به لایه بعدی منتقل می‌کند. پارامترها، که در این فرآیند حیاتی هستند، قدرت ارتباطات بین گره‌ها را تعیین

می‌کنند.

در فرآیندی که به عنوان پیش‌آموزش (pretraining) شناخته می‌شود، LLMsبا اسکن کردن حجم وسیعی از متن، ارتباطات و همبستگی‌ها را بین توکن‌ها – کلمات یا قطعات کلمات – یاد می‌گیرند. در یکLLM ، هر پارامتر مانند یک دکمه تنظیم عمل می‌کند و در بزرگترین مدل‌های امروزی، صدها میلیارد از این پارامترها وجود دارد. از طریق یک فرآیند تکراری تنظیم این پارامترها در تمام گره‌های شبکه، مدل ارتباطات بین توکن‌ها را در داده‌های آموزشی خود تقویت یا کاهش می‌دهد و شروع به شناسایی و تکرار الگوهای پیچیده در زبان می‌کند.

این به معنای آن است که LLMs هرگز یک واقعیت را نمی‌دانند یا یک مفهوم را به شکلی که ما درک می‌کنیم، نمی‌فهمند. در عوض، هر بار که شما با یک سوال به یک LLM فرمان می‌دهید یا از آن می‌خواهید کاری انجام دهد، صرفاً از آن می‌خواهید پیش‌بینی کند که کدام توکن‌ها به احتمال زیاد به دنبال توکن‌های تشکیل‌دهنده فرمان شما به صورت متناسب با متن ظاهر خواهند شد. و آن‌ها همیشه پیش زبان اغراق‌آمیز برای تأثیر بیانی استفاده می‌کند نه اینکه درخواست دستور پخت اسب را داشته باشد. اما حتی زمانی که به نظر می‌رسد مدل‌ها دارای استدلال انسانی مشترک هستند، این‌گونه نیست. در عوض، آن‌ها در حال انجام پیش‌بینی‌های آماری محتمل در مورد الگوهای زبان هستند. این بدان معناست که گاهی اوقات اشتباه می‌کنند و ممکن است به طور غیرقابل پیش‌بینی عمل کنند. هنگامی که مدل‌ها اطلاعات نادرست یا نتایج گمراه‌کننده تولید می‌کنند که به درستی حقایق، الگوها یا ارتباط نظر منطقی ناسازگار یا نامفهوم هستند.

علاوه بر این، وابستگی یک مدل به داده‌ها و اندازه‌گیری ممکن است ظاهری از عینیت یا بی‌طرفی به آن بدهد، اما این مدل نه عینی است و نه بی‌طرف. در عوض، توسط توسعه‌دهندگان و موسسات انسانی ساخته شده که در مورد جمع‌آوری داده‌ها، نحوه پردازش آن‌ها، هدف یا عملکرد خاصی که مدل برای آن بهینه‌سازی می‌شود، نحوه بهترین همسو کردن این عملکرد با

ارزش‌ها و مقاصد انسانی و غیره تصمیم‌گیری می‌کنند. اگر عملکرد ضعیفی دارند.

مسئله دیگر، روش‌های اغلب مبهم عملکرد مدل‌های زبانی بزرگ است، که به عنوان پدیده "جعبه سیاه" شناخته می‌شود. این اتفاق زمانی می‌افتد که شبکه‌های عصبی پیچیده که صدها میلیارد نمونه متنی را با جزئیات بسیار ظریف پردازش می‌کنند، الگوهایی را شناسایی می‌کنند که ناظران انسانی در تشخیص آن‌ها مشکل دارند – و توضیح خروجی‌های مدل یا ردیابی فرآیند تصمیم‌گیری آن را دشوار یا حتی غیرممکن می‌سازد. در حالی که توسعه‌دهندگان از تکنیک‌های مختلفماهای نسبتاً ساده دچار گره شوند. گاهی اوقات خروجی‌های متعصبانه و در موارد دیگر خروجی‌های نامربوط به متن تولید می‌کنند.

شکاکان معتقدند که این وضعیت هرگز به طور کامل تغییر نخواهد کرد. حتی با ساخت سوپرکامپیوترهای بزرگتر برای آموزش مدل‌هایشان با تریلیون‌ها پارامتر و گسترش مجموعه‌های داده آموزشی برای شامل ورودی‌های چندوجهی مانند تصاویر، ویدئوها و داده‌های ساختاریافته، افزایش عملکرد کند شده است. خطاها همچنان پابرجا هستند. منتقدان می‌گویند که آن‌ها هرگز به "جام مقدس" هوش مصنوعی – هوش عمومی مصنوعی – (AGI) دست نخواهند یافت؛ جایی که مدل‌ها قادر به اعمال دانش از یک حوزه یا بافت به موقعیت‌های کاملاً متفاوت، سازگاری با چالش‌های جدید با انعطاف‌پذیری انسانی، استدلال انتزاعی در زمینه‌های متنوع، و تولید ایده‌ها و راه‌حل‌های اصلی – بدون برنامه‌ریزی صریح برای هر وظیفه – می‌شوند.

پیامدهای گسترده اجتماعی: عدم قطعیت و عاملیت انسانی

در طول سال ۲۰۲۳، به لطفLLMs ، هوش مصنوعی بزرگترین داستان در فناوری بود. بسیاری از ناظران معتقد بودند که این مدل‌های جدید در آستانه تغییر همه چیز، به شکلی مثبت هستند. بسیاری دیگر باور داشتند که آن‌ها در آستانه تغییر همه چیز، به شیوه‌ای فاجعه‌بار و منفی هستند. هنوز هم برخی دیگر معتقد بودند که همه چیز یکسان باقی خواهد ماند، فقط با تمرکز بیشتر قدرت، سود و آینده جهان در دست چند بازیگر بزرگ فناوری.

اما تا تابستان ۲۰۲۴، یک تغییر طعنه‌آمیز رخ داد. در حالی که منتقدان هوش مصنوعی مولد قبلاً خواستار یک توقف شش ماهه در توسعه سیستم‌های قدرتمندتر از GPT-4 ، به دلیل ترس از خطرات بالقوه فاجعه‌بار بودند، اکنون سوالاتی مطرح می‌شد که چرا تحویل نسل بعدی مدل‌های پیشگامانه اینقدر طول می‌کشد. و همراه با این سوالات، شک و تردیدهای مداوم و فزاینده‌ای در مورد اینکه LLMs در نهایت چقدر می‌توانند توانمندتر شوند، وجود داشت. بنابراین، آنچه زمانی به عنوان "دشمن عمومی شماره ۱" معرفی می‌شد، اکنون یک "شکست" تلقی می‌شد. عباراتی مانند "هیجان هوش مصنوعی"، "حباب هوش مصنوعی" و "کوهان‌های ناامیدی" شروع به نفوذ در عناوین رسانه‌ها کردند.

اما نویسنده این متن پیش از این نیز تغییرات سرگیجه‌آوری در انتظارات را تجربه کرده بود – اما در جهت مخالف. در سال ۲۰۱۵، زمانی که برای اولین بار درگیر OpenAI شد، این ایده که هوش مصنوعی در نهایت می‌تواند درک و استدلال انسانی را به دست آورد یا به طور معتبری شبیه‌سازی کند، در حاشیه دانش رایج قرار داشت. حتی در سیلیکون ولی، این چشم‌انداز یک شانس بسیار کم در نظر گرفته می‌شد. این یکی از دلایلی بود که OpenAI خود را به عنوان یک سازمان غیرانتفاعی تأسیس کرد. شرکت‌های سرمایه‌گذاری خطرپذیر که به دنبال بازگشت سرمایه در چارچوب زمانی سنتی پنج تا ده ساله بودند، بعید بود که به چنین تلاش سفته‌بازانه و بلندمدتی متعهد شوند.

اگر در سال ۲۰۲۴ چالش‌های بزرگی وجود داشت، در سال ۲۰۱۵ نیز چالش‌های بزرگی وجود داشت. در سال‌های ۲۰۱۸ و ۲۰۲۰ نیز چالش‌های بزرگی وجود داشت. اما به نحوی، OpenAIو بقیه جامعه توسعه‌دهندگان هوش مصنوعی همیشه توانسته‌اند تکنیک‌ها و دستاوردهای جدیدی را کشف کنند که موج بعدی بهبود را امکان‌پذیر می‌ساخت. البته، اگر افق زمانی بلندمدت باشد، خوش‌بین بودن آسان است. و این دیدگاه، بلندمدت است. در واقع، باور بر این است که ما هنوز در مراحل بسیار ابتدایی این فاز جدید از کشف و رشد انسانی قرار داریم. سوپرکامپیوترها حتی

قدرتمندتر خواهند شد. توسعه‌دهندگان به نوشتن الگوریتم‌های کارآمدتر ادامه خواهند داد. برای غلبه بر برخی از محدودیت‌هایی که LLMs را مشخص می‌کند، آن‌ها معماری‌ها و تکنیک‌های جدیدی را ابداع خواهند کرد و رویکردهای متفاوتی مانند یادگیری چندوجهی و هوش مصنوعی نوروسمبولیک – سیستم‌هایی که شبکه‌های عصبی را با استدلال سمبولیک مبتنی بر قوانین و منطق صریح و تعریف‌شده توسط انسان ادغام می‌کنند – را در خود جای خواهند داد.

تمام این‌ها بدان معناست که ما هنوز در مراحل ابتدایی رویارویی وجودی هستیم که این سیستم‌ها برمی‌انگیزند، زیرا تلاش می‌کنیم تا درک کنیم که معرفی اشکال جدید و نه کاملاً قابل پیش‌بینی از هوش در جهان ما به راستی چه معنایی دارد. ماشینی که می‌تواند مانند یک انسان فکر کند – به صورت استراتژیک، انتزاعی و حتی خلاقانه، با سرعت و مقیاس یک کامپیوتر – به وضوح انقلابی خواهد بود. اگر هر کودک روی کره زمین به طور ناگهانی به یک معلم هوش مصنوعی دسترسی داشته باشد که به اندازه لئوناردو داوینچی باهوش و به اندازه یک شخصیت کارتونی معروف همدل باشد، چه می‌شود؟ اگر میلیاردها نفر در سراسر جهان به طور ناگهانی در هر زمان یک مشاور بهداشتی بسیار آگاه و قابل اعتماد در جیب خود داشته باشند، چه اتفاقی می‌افتد؟

البته، همه بر جنبه‌های مثبت بالقوه هوش مصنوعی تمرکز نمی‌کنند – به ویژه در ایالات متحده، جایی که مردم به طور مداوم نگرانی‌های بالایی در مورد این فناوری ابراز می‌کنند. بر اساس یک نظرسنجی در سال ۲۰۲۲ که توسط Ipsos ، یک شرکت تحلیلگر جهانی انجام شد، "فقط ۳۵ درصد از آمریکایی‌های مورد بررسی (که یکی از پایین‌ترین ارقام در میان کشورهای مورد بررسی است) موافق بودند که محصولات و خدمات با استفاده از هوش مصنوعی مزایای بیشتری نسبت به معایب دارند." یک نظرسنجی مشابه از مرکز تحقیقات Pew نشان داد که تنها ۱۵ درصد از بزرگسالان ایالات متحده گفتند که "بیشتر هیجان‌زده هستند تا نگران استفاده فزاینده از هوش مصنوعی در زندگی روزمره." در مطالعه دیگری، از دانشگاه Monmouth ، ۵۶ درصد گفتند که

"ماشین‌های هوش مصنوعی کیفیت کلی زندگی انسان‌ها را کاهش خواهند داد".

چنین نگرانی‌هایی کاملاً قابل درک هستند. ما در حال تجربه یک لحظه دگرگون‌کننده جهان هستیم که عدم قطعیت قابل توجهی ایجاد می‌کند. این سیستم‌ها دقیقاً چقدر خوب خواهند شد و با چه سرعتی؟ چه نوع مشاغلی برای انسان‌ها باقی خواهد ماند در حالی که هوش مصنوعی به پیشرفت خود ادامه می‌دهد؟ چه اتفاقی برای اعتماد و گفتمان عمومی، که از قبل تحت محاصره است، می‌افتد در حالی که فناوری‌های هوش مصنوعی تولید شبیه‌سازی‌های متقاعدکننده واقعیت را در مقیاس وسیع‌تر و ارزان‌تر می‌کنند؟ چه اتفاقی برای حریم خصوصی و خودمختاری فردی در دنیایی بسیار مجهز می‌افتد، جایی که میلیاردها سیستم، دستگاه و ربات می‌توانند عملکرد انسانی را مطابقت دهند یا از آن فراتر روند و می‌توانند اقداماتی انجام دهند که ممکن است انتخاب‌ها و خواسته‌های ما را نقض کند؟ آیا می‌توانیم کنترل زندگی خود را حفظ کنیم و سرنوشت خود را با موفقیت ترسیم کنیم؟

هوش مصنوعی به عنوان کاتالیزور "ابراژانس": یک چشم‌انداز تاریخی

این سوال آخر است که تمرکز اصلی این کتاب را شکل می‌دهد: آیا می‌توانیم کنترل زندگی خود را حفظ کنیم و سرنوشت خود را با موفقیت ترسیم کنیم؟ در نهایت، سوالات مربوط به جابجایی شغلی، سوالات مربوط به عاملیت انسانی فردی هستند: آیا ابزارهای اقتصادی برای حمایت از خود و فرصت‌هایی برای درگیر شدن در فعالیت‌هایی که برایم معنادار است، خواهم داشت؟ سوالات مربوط به اطلاعات غلط و گمراه‌کننده، سوالات مربوط به عاملیت انسانی فردی هستند: چگونه بدانم به چه کسی و چه چیزی اعتماد کنم در حالی که تصمیماتی می‌گیرم که بر زندگی من تأثیر می‌گذارند؟ سوالات مربوط به حریم خصوصی، سوالات مربوط به عاملیت انسانی فردی هستند: چگونه یکپارچگی هویت خود و نحوه شناخته شدنم در جهان را حفظ کنم و حس واقعی خود را حفظ کنم؟

عاملیت انسانی یک مفهوم بنیادی در فلسفه، جامعه‌شناسی و روان‌شناسی است. این مفهوم بیان می‌کند که شما به عنوان یک فرد، توانایی انتخاب‌های خود، عمل مستقل و در نتیجه

اعمال نفوذ بر زندگی خود را دارید. در حالی که ممکن است معتقد باشید که شرایط و وضعیت بیرونی نقش مهمی در نتایجی که تجربه می‌کنید ایفا می‌کنند، این حس عاملیت شماست که شما را وادار به شکل‌گیری نیت‌ها، تعیین اهداف و انجام اقدامات برای دستیابی به آن نتایج می‌کند. بنابراین، حس عاملیت می‌تواند زندگی شما را با هدف و معنا غنی سازد.

همانطور که سیستم‌های هوش مصنوعی تکامل می‌یابند، ظرفیت آن‌ها برای یادگیری خودگردان، حل مسئله و اجرای یک سری وظایف پیچیده بدون نظارت مداوم انسانی در حال افزایش است. در حالی که یک خودروی خودران به همان شکلی که انسان‌ها از عاملیت خود آگاه هستند، آگاه نیست، اما توانایی تصمیم‌گیری مستقل، انجام اقدامات و پیگیری اهداف در حوزه‌های عملیاتی خود را دارد. با گذشت زمان، این بدان معناست که تنوع فزاینده‌ای از سیستم‌ها، دستگاه‌ها و ماشین‌ها در حوزه‌هایی که به طور سنتی توسط عاملیت انسانی اداره می‌شدند، نفوذ خواهند کرد – به روش‌هایی که انسان‌ها اغلب ممکن است آن را ناخوشایند بدانند. و حتی در مواردی که ما از چنین انتقال شناختی استقبال می‌کنیم، مسائل دیگری مطرح می‌شوند: اگر از طریق وابستگی بیش از حد به عاملیت و قابلیت‌های ماشین، مهارت‌ها و عاملیت خود ما به مرور زمان تحلیل روند، چه می‌شود؟ اگر سیستم‌هایی که قرار است به نفع ما کار کنند – و نتایجی را که ما تأیید می‌کنیم ارائه دهند – در نهایت رفتارها و انتخاب‌های ما را به روش‌هایی شکل دهند که صریحاً به آن‌ها رضایت نداده‌ایم، چه می‌شود؟

در واقع، این داستان بشریت تا به امروز است. به عنوان "Homo techne" (انسان فن‌ورز)، ما با ظرفیت و تعهدمان به ایجاد روش‌های جدید بودن در جهان از طریق ابزارسازی تعریف می‌شویم. این چیزی است که ما را از سایر موجودات زنده روی کره زمین، حتی ابزارسازان ابتدایی، متمایز می‌کند. در حالی که نسل به نسل شامپانزه‌ها ممکن است از ابزارهای سنگی برای شکستن آجیل استفاده کنند، ما به طور مداوم فناوری‌های جدیدی را ایجاد می‌کنیم که سپس از آن‌ها برای اعمال عاملیت خود به روش‌های جدید و فزاینده‌ای مولد استفاده می‌کنیم. از ابزارهای سنگی گرفته تا تلفن‌های هوشمند، این چرخه فضیلت‌مند نوآوری، معنای انسان بودن را در طول زمان

گسترش و ترکیب می‌کند.

اکنون فرصتی برای توسعه "ابرابزارهای" جدید داریم. ابزارهایی که می‌توانند عاملیت ما را آنقدر افزایش دهند که در حال حاضر در میان چیزی شبیه به انقلاب صنعتی قرار داریم. پیامدهای این ادعا گسترده است، اما یک راه برای فکر کردن در مورد آن این است: هوش و انرژی، در کنار هم، عاملیت انسانی و در نتیجه پیشرفت انسانی را به حرکت درمی‌آورند. هوش به ما ظرفیت سنجش گزینه‌ها و تصور و برنامه‌ریزی برای سناریوهای بالقوه مختلف را می‌دهد. انرژی ما را قادر می‌سازد تا سپس بر اساس آنچه آرزو داریم به آن دست یابیم، اقدام کنیم. هرچه بیشتر هوش و انرژی را به نفع خود به کار گیریم، ظرفیت ما برای تحقق بخشیدن به چیزها، به صورت فردی و جمعی، بیشتر می‌شود.

زبان گفتاری، استفاده کنترل‌شده از آتش، چرخ و زبان نوشتاری همگی فناوری‌های محوری بودند که اجداد ما برای تقویت و گسترش هوش و انرژی انسانی از آن‌ها استفاده کردند. با ادامه افزودن نوآوری‌های جدید به این ترکیب، این فرآیند ادامه می‌یابد. در اوایل قرن بیستم، اتومبیل‌های مجهز به موتورهای احتراق داخلی، میلیون‌ها نفر را قادر ساختند تا به طور معمول به آنچه چند دهه قبل، شاهکارهای فراانسانی تلقی می‌شد، دست یابند. در دهه‌های اخیر، کامپیوترهای شخصی، اینترنت و تلفن‌های هوشمند نیز به همین ترتیب در تقویت و گسترش هوش انسانی تحول‌آفرین بودند.

هوش مصنوعی جهش بزرگ بعدی ما را امکان‌پذیر خواهد ساخت. برخلاف نوآوری‌هایی مانند کتاب‌ها یا ویدئوهای آموزشی در یوتیوب، هوش مصنوعی صرفاً راهی برای تولید و توزیع دانش نیست، هرچند که این نیز ارزشمند است. از آنجا که یک هوش مصنوعی ظرفیت "عاملیت‌مند" بودن را دارد، می‌تواند اهداف را تعیین کرده و به تنهایی برای دستیابی به آن‌ها اقدام کند، شما می‌توانید هوش مصنوعی را به دو روش متمایز به کار گیرید. در برخی موارد، ممکن است بخواهید از نزدیک با یک هوش مصنوعی کار کنید – مانند زمانی که در حال یادگیری یک زبان جدید یا تمرین مهارت‌های ذهن‌آگاهی هستید. در موارد دیگر، مانند بهینه‌سازی مصرف انرژی خانه شما

بر اساس قیمت‌های لحظه‌ای انرژی و پیش‌بینی‌های آب و هوا، ممکن است ترجیح دهید که هوش مصنوعی خود به تنهایی این کار را انجام دهد.

در هر صورت، هوش مصنوعی عاملیت شما را افزایش می‌دهد، زیرا به شما کمک می‌کند اقداماتی را انجام دهید که برای رسیدن به نتایج دلخواه شما طراحی شده‌اند. و در هر صورت، چیزی جدید و تحول‌آفرین در حال وقوع است. برای اولین بار در تاریخ، هوش مصنوعی ترکیبی (synthetic intelligence)، نه فقط دانش، به همان انعطاف‌پذیری هوش مصنوعی ترکیبی از زمان ظهور نیروی بخار در قرن ۱۸، قابل استقرار شده است. خود هوش اکنون یک ابزار است – یک موتور مقیاس‌پذیر، بسیار قابل تنظیم و خودتقویت‌کننده برای پیشرفت. اگر به درستی از آن استفاده کنیم، می‌توانیم به وضعیت جدیدی از "ابراژانس (superagency) "دست یابیم. این اتفاق زمانی می‌افتد که توده بحرانی از افراد، که شخصاً توسط هوش مصنوعی توانمند شده‌اند، شروع به فعالیت در سطوحی کنند که در سراسر جامعه ترکیب می‌شوند. به عبارت دیگر، این فقط این نیست که برخی افراد به لطف هوش مصنوعی آگاه‌تر و مجهزتر می‌شوند. همه این گونه خواهند شد، حتی آن‌هایی که به ندرت یا هرگز مستقیماً از هوش مصنوعی استفاده نمی‌کنند. زیرا ناگهان پزشک شما می‌تواند شکایات مبهم و به ظاهر نامربوط شما را با دقت هوش مصنوعی تشخیص دهد. مکانیک خودروی شما دقیقاً می‌داند که آن صدای عجیب از صندوق عقب ماشین شما در هنگام شتاب‌گیری از چراغ راهنمایی در یک روز گرم به چه معناست. حتی خودپردازها، پارکومترها و دستگاه‌های فروش خودکار، نابغه‌های چندزبانه هستند که نیازهای شما را فوراً درک کرده و با ترجیحات شما سازگار می‌شوند. این دنیای ابراژانس است.

قیاس انقلاب صنعتی: طرحی برای پیشرفت

پیش از عصر صنعتی، بهره‌وری در سراسر جامعه به شدت محدود بود، زیرا انرژی کالایی کمیاب محسوب می‌شد. شخم زدن مزارع، حفر تونل‌ها، کار با قرقره‌ها – همه اینها نیازمند نیروی کار گسترده انسانی یا حیوانی بود که مقیاس، کارایی و به طور کلی شکوفایی انسانی را محدود می‌کرد،

زیرا انسان و اسب هر دو عملاً به عنوان حیوانات بارکش عمل می‌کردند. آسیاب‌های بادی و آسیاب‌های آبی امکان تولید کارآمدتر غلات و منسوجات را در برخی مناطق، البته با اجازه طبیعت، فراهم می‌کردند. اگر در نزدیکی رودخانه متولد می‌شدید، زندگی ممکن بود کمی مرفه تر باشد و فرصت‌های بیشتری برای بیان خود به شما ارائه دهد. در غیر این صورت، امید است که از شخم زدن لذت می‌بردید.

اما در اواسط قرن هجدهم، با ظهور نیروی بخار، اوضاع شروع به تغییر کرد. در نگاه به گذشته، ما اغلب سریعاً انقلاب صنعتی را از طریق بدترین اثرات ناشی از آن مشاهده می‌کنیم. شهرهایی که با سوزاندن زغال‌سنگ برای نیروی بخار، سیاه می‌شدند. کارهای طاقت‌فرسا در کارخانه‌های خطرناک که کارگران در آنها حقوق کمی داشتند یا اصلاً حقوقی نداشتند. کار کودکان. الگوهای زندگی منظم‌تر و کمتر جمعی. تراکم جمعیت در شرایط اسکان‌های نامناسب و بی‌هویتی شهری در همان زمان. حتی می‌توان گفت که آن دوران، دوره‌ای خشونت‌آمیز و غیرانسانی بود. ساعت‌های کارخانه و چراغ‌های خیابان، ریتم‌های ذاتی طبیعت را از بین بردند. از مردم اغلب انتظار می‌رفت که بیشتر شبیه ماشین‌ها عمل کنند. معاملات مصرف‌کننده در دنیایی که به طور فزاینده‌ای تجاری شده بود، فرصت‌ها برای مبادله متقابل و اعتماد شخصی را محدود کرد. رشد و رفاه در مقیاس جهانی نابرابر بود، با تقاضاهای جدید برای منابع طبیعی که روابط استعماری از قبل استثمارگرانه را تشدید کرده و نابرابری‌های کاملاً جدیدی را ایجاد می‌کرد.

اما اینها تنها حقایق مربوط به نیروی بخار و انقلاب صنعتی نبودند. در مجموع، در طول قرون، انرژی ترکیبی و صنعتی‌سازی نیروهای رادیکالی برای انسانی‌تر کردن بوده‌اند. مکانیزاسیون با نیروی بخار فرصت‌هایی برای دستیابی به مقیاس در سطوح بی‌سابقه ایجاد کرد. پیگیری مقیاس نیازمند سطوح جدیدی از همکاری و مشارکت بود. پس از دستیابی به آن، مقیاس سطوح جدیدی از فراوانی و تنوع را تولید کرد. جامعه‌ای که بیشتر بر فناوری متکی بود، به جمعیتی باسوادتر نیز نیاز داشت. افراد بیشتر با تنوع بیشتری از مهارت‌ها و تخصص‌ها منجر به نوآوری‌های اضافی شد که به نوبه خود الهام‌بخش امکانات و فرصت‌های جدید شد. کارایی‌های جدید در

تولید غذا، انواع کالاها و اطلاعات در نهایت به ایجاد پیشرفت اجتماعی کمک کرد که اغلب با رفاه و فراوانی همراه است: قوانین عادلانه‌تر، تحرک اقتصادی و فرهنگی بیشتر، سیستم‌های رفاه اجتماعی فزاینده قوی‌تر و تأکید روزافزون بر حقوق و آزادی‌های فردی. زمانی بود که تقریباً تمام تلاش انسانی بر روی کشت غذا و کشتار دام متمرکز بود. با تقویت نیروی کار فیزیکی با نیروی کار ترکیبی، موتور بخار امکانات انسان بودن را به روش‌هایی با دامنه تصاعدی گسترش داد. این یک مسیر رو به جلو به سوی وجودی غنی‌تر، متنوع‌تر، با کرامت‌تر و انسانی‌تر فراهم کرد.

اکنون تمام این‌ها می‌تواند دوباره با هوش مصنوعی تکرار شود. ما می‌توانیم یک جهش تصاعدی رو به جلو داشته باشیم، با هوش مصنوعی ترکیبی که همان تأثیری را در قرن بیست و یکم و فراتر از آن خواهد داشت که انرژی ترکیبی در قرن هجدهم شروع به ایجاد آن کرد. همانطور که انقلاب صنعتی فرصت‌های جدیدی برای همکاری و ظرفیت‌های جدیدی برای نوآوری، خلاقیت و بهره‌وری ایجاد کرد، این انقلاب شناختی نیز چنین خواهد کرد. با وقوع این تحول، آموزش و توسعه مهارت‌ها حتی حیاتی‌تر خواهند شد و به جمعیتی آگاه‌تر و توانمندتر منجر خواهند شد. و در نهایت هوش مصنوعی ترکیبی می‌تواند پتانسیل انسانی و عاملیت انسانی را به همان شکلی که انرژی ترکیبی انجام داده است، گسترش دهد. این مسیری به سوی وجودی رضایت‌بخش‌تر و انسانی‌تر است.

دموکراتیزه کردن هوش مصنوعی: فلسفه OpenAI و استقرار تکراری

پس بهترین راه برای پیگیری این چشم‌انداز مثبت از هوش فراوان متمرکز بر عاملیت انسانی فردی چیست؟ هنگامی که OpenAI در سال ۲۰۱۵ راه‌اندازی شد، هیچ محصول یا خدمات خاصی در ذهن نداشت، چه رسد به یک مدل کسب‌وکار یا استراتژی ورود به بازار. در عوض، همانطور که دو هم‌بنیان‌گذار OpenAI ، گرگ براکمن و ایلیا سوتسکور، در مقاله‌ای که هنگام راه‌اندازی منتشر شد، گفتند، مأموریت گسترده OpenAI این بود که "هوش دیجیتال را به شیوه‌ای پیش ببرد که به احتمال زیاد برای بشریت به طور کلی مفید باشد".

برای دستیابی به این هدف بلندپروازانه اما نسبتاً انتزاعی، براکمن و سوتسکور پیشنهاد کردند که OpenAI از یک دستورالعمل پیروی کند که اثبات شد تحول‌آفرین است: "ما معتقدیم هوش مصنوعی باید گسترشی از اراده‌های انسانی فردی باشد و با روح آزادی، به طور گسترده و یکنواخت توزیع شود." چرا این دیدگاه اینقدر محوری بود؟ حتی در سال ۲۰۱۵، برخی از نسخه‌های هوش مصنوعی در واقع از قبل به طور گسترده توزیع شده بودند. اکثر کاربران اینترنت قبلاً از هوش مصنوعی به شکل سیستم‌های توصیه، انتخاب محتوای خبری و خدمات پیش‌بینی متن و تکمیل خودکار، برای نام بردن فقط چند کاربرد، استفاده می‌کردند. سناریوهای دیگری نیز وجود داشت که اغلب بحث‌برانگیز بودند. در سال ۲۰۱۶، ProPublicaگزارشی درباره یک الگوریتم منتشر کرد که در سراسر سیستم عدالت کیفری ایالات متحده استفاده می‌شد و "به ویژه احتمال بیشتری داشت که متهمان سیاهپوست را به عنوان مجرمان آینده به اشتباه پرچم‌گذاری کند و آن‌ها را با تقریباً دو برابر نرخ متهمان سفیدپوست به این شکل نادرست برچسب‌گذاری کند".

همان سال، ائتلافی از هفده سازمان حقوق مدنی بیانیه‌ای صادر کرد که در آن ابزارهای پلیس پیش‌بینی‌کننده را که از الگوریتم‌ها و داده‌های گزارش جرم تاریخی برای پیش‌بینی زمان و مکان وقوع جرایم استفاده می‌کنند، محکوم کرد. چند سال بعد، در سال ۲۰۱۹، سان فرانسیسکو اولین شهری شد که اداره پلیس و سایر سازمان‌های شهری خود را از استفاده از سیستم‌های تشخیص چهره منع کرد. تمام این موارد استفاده یک وجه مشترک داشتند: هیچ یک از افرادی که هوش مصنوعی در این سناریوها بر آن‌ها تأثیر می‌گذاشت، به صراحت استفاده از آن را انتخاب نکرده بودند. حتی با خدمات توصیه و انتخاب محتوای خبری، شما اغلب یک انتخاب صریح برای استفاده از هوش مصنوعی انجام نمی‌دهید. در عوض، شما صرفاً مسیری را دنبال می‌کنید که توسعه‌دهندگان برای شما تعیین کرده‌اند. آمازون می‌گوید که ممکن است من به این چیز علاقه داشته باشم چون اخیراً این چیز دیگر را خریده‌ام؟ شاید آن را بررسی کنم.

مدل‌های هوش مصنوعی جدیدی که OpenAI در سال ۲۰۲۰ شروع به انتشار آن‌ها کرد، با یک

LLMاولیه به نامGPT-2 ، عمیقاً متفاوت بودند. آن‌ها در سیستم‌های دیگر تعبیه نشده‌اند یا بدون اطلاع شما فعال نمی‌شوند. آن‌ها چیزی هستند که شما باید به طور مثبت انتخاب کنید که از آن‌ها استفاده کنید. هنگامی که این کار را انجام می‌دهید، می‌توانید از آن‌ها به روش‌های باز و خودتعیین‌کننده استفاده کنید.

با انتشار ChatGPT در نوامبر ۲۰۲۲، این رویکرد به سطح جدیدی از کاربرد متنوع دست یافت. البته، می‌توانستید از آن بخواهید که برای شما یک مقاله بنویسد. اما می‌توانستید از آن بخواهید که مقاله‌ای را که نوشته‌اید نقد کند. می‌توانستید از آن بخواهید که سوالاتی برای مصاحبه آتی شما در یک شرکت بسازد. می‌توانستید از آن بخواهید که یک روتین ورزشی چالش‌برانگیز اما بی‌صدا را پیشنهاد دهد که می‌توانید در آپارتمان خود بدون بیدار کردن هم‌اتاقی‌تان انجام دهید.

بالاخره، یک ابزار هوش مصنوعی بسیار در دسترس و آسان برای استفاده وجود داشت که به وضوح با شما و برای شما کار می‌کرد، نه بر روی شما. این یک تغییر حیاتی در توسعه هوش مصنوعی و توانمندسازی انسان بود. این حیاتی است زیرا کاربران فردی را در قلب تجربه قرار می‌دهد. و به همان اندازه مهم، فرصت‌هایی را به آن‌ها می‌دهد تا تجربیاتی را داشته باشند که خودشان جستجو کرده یا طراحی کرده‌اند. به جای توسعه این فناوری جدید پشت درهای بسته تا زمانی که یک کادر کوچک از کارشناسان تصمیم بگیرند که به اندازه کافی مؤثر و کاملاً ایمن عمل می‌کند، OpenAIاز عموم مردم دعوت کرد تا در فرآیند توسعه شرکت کنند. این رویکرد را "استقرار تکراری (iterative deployment) "توصیف کرد.

استقرار تکراری دارای بنیان‌های فنی، نظارتی و جامعه‌شناختی محکم است. به سبک کلاسیک سیلیکون ولی، بر تجربه کاربر و بازخورد صریح کاربر برای اطلاع‌رسانی تلاش‌های توسعه مداوم تکیه دارد؛ این بخش فنی است. به نیازهای نظارتی توجه دارد، زیرا تغییر را به تدریج معرفی می‌کند تا تأثیر اثرات نامطلوب احتمالی را محدود کند. از نظر جامعه‌شناختی نیز با این دیدگاه سازگار است که به افراد و جامعه زمان می‌دهد تا با این تغییرات سازگار شوند. در اصل، استقرار

تکراری تشخیص می‌دهد که رم یک روزه ساخته نشد – و شهروندان رومی نیز یک روزه ساخته نشدند. زمان می‌برد تا خطوط یک دنیای جدید و نحوه عملکرد در آن را درک کرد. اعتماد با قرار گرفتن در معرض شروع می‌شود و با استفاده تکامل می‌یابد. هنگامی که می‌آموزید چیزی چیست و چگونه کار می‌کند، شروع به اعتماد به آن می‌کنید. اعتماد برابر است با ثبات در طول زمان.

در زمینه هوش مصنوعی، ابتدا باید به خود فناوری‌ها اعتماد کنیم – کاری آسان نیست وقتی فناوری‌ها تا حدی غیرقابل پیش‌بینی و مستعد خطا هستند. اما اگر به ابزارها دسترسی داشته باشیم، می‌توانیم برای خودمان نظراتی را شکل دهیم: آیا آنقدر قابل اعتماد است که کارهایی را که می‌خواهم انجام دهم؟ آیا ارزش واقعی ارائه می‌دهد یا فقط تازگی؟ آیا من را توانمند می‌کند یا بیش از حد وابسته می‌سازد؟ اعتماد به فناوری‌ها تازه آغاز کار است. ما باید به توسعه‌دهندگان فناوری‌ها، تنظیم‌کننده‌های فناوری‌ها و شاید بیش از همه، سایر کاربران فناوری‌ها نیز اعتماد کنیم. بالاخره، چرا باید به دیگران اعتماد کنید که از هوش مصنوعی به روش‌های عمدتاً مثبت استفاده کنند؟ چرا باید به دولت خود – یا دولت‌های کشورهای دیگر – اعتماد کنید که همین کار را انجام دهند؟ چرا باید به توسعه‌دهندگان هوش مصنوعی اعتماد کنید که کار خود را به نحو شایسته و مسئولانه انجام دهند؟

نقشه‌برداری چشم‌انداز هوش مصنوعی: دیدگاه‌های متنوع و مسیر به سوی اجماع

پاسخ کوتاهی برای این سوالات وجود ندارد. پاسخ یک گفتگوی مداوم و اغلب بحث‌برانگیز است، درست همانطور که همیشه وقتی بشریت فناوری‌های جدید و تحول‌آفرین ایجاد می‌کند، بوده است. این گفتگوی خاص سال‌هاست که ادامه دارد، اما بیشتر در میان دانشمندان کامپیوتر، محققان آموزش عالی، توسعه‌دهندگان تجاری، سرمایه‌گذاران، اخلاق‌گرایان، روزنامه‌نگاران فناوری، سیاست‌گذاران، کارشناسان حقوقی و نویسندگان داستان‌های علمی تخیلی در جریان بوده است. تاکنون، حداقل چهار گروه اصلی این گفتمان را شکل داده‌اند: "نابودی‌خواهان(Doomers) "، "سیاه‌بینان(Gloomers) "، "سریع‌روندگان (Zoomers) "و

"شکوفا کنندگان .(Bloomers) "در حالی که اینها به وضوح برچسب‌های کلی هستند، هدف ما کاهش پیچیدگی دیدگاه‌های زیربنایی آن‌ها نیست. در عوض، صرفاً قصد داریم مواضع اساسی را که بحث‌ها پیرامون هوش مصنوعی را به حرکت درآورده‌اند، خلاصه کنیم. همه این مکاتب فکری بینش‌های منحصر به فرد و ارزشمندی را ارائه می‌دهند، مفروضات را به چالش می‌کشند و مرزهای درک جمعی ما را فراتر می‌برند.

نابودی‌خواهان بر این باورند که ما در مسیری به سوی آینده‌ای هستیم که در بدترین سناریوها، هوش مصنوعی‌های فوق‌هوشمند و کاملاً خودمختار که دیگر با ارزش‌های انسانی همسو نیستند، ممکن است تصمیم بگیرند که ما را به طور کلی نابود کنند، مگر شاید تعداد کمی از کارشناسان فناوری را که برای انجام کارهای پیش پا افتاده نگه دارند، به عنوان انتقام از ماشین‌های رباتیک خانگی.

سیاه‌بینان هم به شدت منتقد هوش مصنوعی هستند و هم به شدت منتقد نابودی‌خواهان. به تخمین آن‌ها، دیدگاه نابودی‌خواهان دو هدف را دنبال می‌کند. اول، این دیدگاه به عنوان تأیید ضمنی هوش مصنوعی عمل می‌کند – "آنقدر قدرتمند است که ممکن است ما را نابود کند!" دوم، ماهیت بلندمدت و انتزاعی آن توجه را به آینده معطوف می‌کند، در حالی که اولویت واقعی ما باید خطرات و آسیب‌های کوتاه‌مدت هوش مصنوعی مانند از دست دادن مشاغل بالقوه، اطلاعات غلط در مقیاس وسیع، تقویت تعصبات سیستمی موجود و تضعیف عاملیت فردی باشد. به طور کلی، سیاه‌بینان از یک رویکرد ممنوعیت‌گرایانه و از بالا به پایین حمایت می‌کنند، جایی که توسعه و استقرار باید توسط نهادهای نظارتی و مقررات رسمی به دقت کنترل و پایش شوند.

در مقابل، **سریع‌روندگان** استدلال می‌کنند که دستاوردهای بهره‌وری و نوآوری هوش مصنوعی به مراتب از هرگونه تأثیر منفی آن فراتر خواهد رفت. به طور کلی، آن‌ها نسبت به ایده مقررات پیشگیرانه که سعی در از بین بردن امکان خطر یا آسیب قبل از وقوع استقرار در دنیای واقعی دارد، شکاک هستند. در عوض، آن‌ها استدلال می‌کنند که دادن فضا به توسعه‌دهندگان برای

فعالیت به دلخواه خود، سریع‌ترین نتایج را به دست خواهد آورد. آن‌ها خواهان مقررات دولتی نیستند. آن‌ها خواهان حمایت دولتی نیستند. آن‌ها یک مسیر روشن و خودمختاری کامل برای نوآوری می‌خواهند.

در نهایت، **شکوفا کنندگان** وجود دارند. مانند سریع‌روندگان، دیدگاه آن‌ها اساساً خوش‌بینانه است. آن‌ها معتقدند هوش مصنوعی می‌تواند پیشرفت انسانی را در بی‌شمار حوزه‌ها تسریع کند. در عین حال، آن‌ها تشخیص می‌دهند که فناوری به این تحول‌آفرینی و چندوجهی بودن نمی‌تواند و نباید به صورت یکجانبه توسعه و استقرار یابد. هوش مصنوعی قرار است بر زندگی افراد زیادی به طرق مختلف تأثیر بگذارد. بنابراین شکوفا کنندگان مشارکت گسترده، در شرایط دنیای واقعی – که همان چیزی است که با استقرار تکراری به دست می‌آید – را دنبال می‌کنند. در حالی که آن‌ها بدون قید و شرط مخالف مقررات دولتی نیستند، معتقدند که سریع‌ترین و مؤثرترین راه برای توسعه ابزارهای هوش مصنوعی ایمن‌تر، عادلانه‌تر و مفیدتر، دسترسی به طیف متنوعی از کاربران با ارزش‌ها و نیت‌های مختلف است.

به جای دیدن هوش مصنوعی به عنوان یک صنعت استخراجی، همانطور که بسیاری از سیاه‌بینان انجام می‌دهند، شکوفا کنندگان آن را بیشتر شبیه به کشاورزی می‌دانند. شما یک دانه می‌کارید. شما رشد آن را تماشا می‌کنید و با شرایط داده شده سازگار می‌شوید. شما یاد می‌گیرید کدام محصولات در کجا بهتر عمل می‌کنند و شروع به مداخله به روش‌هایی می‌کنید که از هر آنچه در مورد مشکلات و چالش‌های پیش آمده و راه‌حل‌هایی که می‌توانند آن‌ها را کاهش دهند، می‌آموزید، آگاه شده است. این یک فرآیند بدون ریسک نیست. با گذشت زمان، دانش شما افزایش می‌یابد، تکنیک‌های شما بهبود می‌یابد، محصول شما رشد می‌کند. نویسنده این متن خود را در گروه شکوفا کنندگان قرار می‌دهد. از نقشش در راه‌اندازی لینکدین گرفته تا سرمایه‌گذاری در صدها استارتاپ اینترنتی در بیست و پنج سال گذشته، او حرفه خود را بر مبنای اصول حلقه‌های بازخورد شبکه و یادگیری و بهبود بر اساس استفاده در دنیای واقعی بنا نهاده است.

تأکید بر مشارکت گسترده به روش‌های خودگردان است که نویسنده را به سمت دیدگاه شکوفا کنندگان می‌کشاند، زیرا مشارکت هم نیازمند و هم پاداش‌دهنده عاملیت فردی است. در مقابل، هم نابودی‌خواهان و هم سیاه‌بینان فرض می‌کنند که از عاملیت انسانی محافظت می‌کنند – با ممنوعیت یا محدود کردن هوش مصنوعی، و در نتیجه سلب توانایی ما برای ارزیابی آن خودمان. با سریع‌روندگان، قضیه برعکس است: آن‌ها آزادی مطلق برای ساختن می‌خواهند، عموماً با این فرض که نتایج در نهایت عاملیت انسانی را افزایش خواهد داد. اما این دیدگاه اغلب در نظر نمی‌گیرد که سرعت سرسام‌آور مورد علاقه آن‌ها چگونه باعث می‌شود بقیه بشریت نسبت به سهم خود در این فرآیند احساس کنند.

با توانمندسازی افراد برای استفاده از ابزارهای هوش مصنوعی به روش‌های عملی، OpenAI فوراً گفتگوی پیرامون آینده بالقوه هوش مصنوعی را گسترش داد و آن را فراگیرتر کرد. این رویکرد همچنین به همه اطلاعات بیشتری برای ارزیابی فرصت‌ها و چالش‌هایی که این ابزارها ایجاد می‌کنند، می‌دهد. هرچقدر که پیشرفت‌های فناوری محرک هوش مصنوعی در چند سال گذشته هیجان‌انگیز بوده است، این دموکراتیزاسیون دسترسی به همان اندازه هیجان‌انگیز و مسلماً مهمتر بوده است. اگر ما واقعاً در میان گذری هم‌تراز با ظهور نیروی بخار هستیم، تصمیماتی که اکنون می‌گیریم فقط چند سال آینده را شکل نخواهند داد. آن‌ها چند دهه و حتی چند قرن آینده را شکل خواهند داد. در عصر شبکه‌ها و توانمندسازی دیجیتال، هیچ کس نباید به نقش یک شخصیت غیربازیکن در ایجاد آینده خود تنزل یابد. برای توسعه هوش مصنوعی به صورت دموکراتیک، مسئولانه، اخلاقی و کارآمد، باید مشارکت گسترده عمومی و بازخورد مستمر عمومی را در خود جای دهیم. بشریت وارد گفتگو شده است، و این یک تحول داستانی بسیار ضروری است.

ضرورت‌های جهانی: عاملیت، اعتماد و آینده دموکراسی‌ها

یک عنصر استراتژیک جهانی نیز در این جنبه از داستان وجود دارد، و آن به عاملیت فردی باز می‌گردد. زیرا در نهایت یک آینده مثبت هوش مصنوعی فقط به این بستگی ندارد که کدام

شرکت‌ها یا حتی کدام کشورها فناوری را سریع‌تر توسعه می‌دهند. نتایج همچنین به این بستگی خواهد داشت که کدام افراد، شرکت‌ها و کشورها از هوش مصنوعی مولدترین استفاده را می‌کنند. این راه دیگری برای گفتن این است که مسئله کلیدی در اینجا اعتماد است.

همانطور که توسعه‌دهندگان مدل‌های توانمندتری را معرفی می‌کنند، به ویژه مدل‌هایی که می‌توانند خودمختارتر و با حداقل نظارت یا تعامل انسانی عمل کنند، اختلافات فرهنگی که این هوش مصنوعی‌ها در حال حاضر الهام‌بخش آن هستند، تنها افزایش خواهد یافت. تمایل به تنظیم پیشگیرانه و حتی ممنوعیت آن‌ها قوی‌تر خواهد شد. اما همه اینها در مقیاس جهانی در حال وقوع است. هر تصمیم نظارتی که ایالات متحده یا هر کشور دیگری بگیرد، در خلاء آشکار نخواهد شد. آینده هوش مصنوعی آمریکا تا حدی توسط تصمیماتی که چین، اتحادیه اروپا، هند و ده‌ها کشور دیگر می‌گیرند، و بالعکس، تعیین خواهد شد. این بسیار شبیه یک بازی چندنفره است.

با توجه به این موضوع، بحث‌ها درباره پذیرش هوش مصنوعی اغلب پیرامون تنظیم مقررات می‌چرخند. سریع‌روندگان خواستار آزادی نوآوری هستند. سیاه‌بینان از توقف، ممنوعیت و رژیم‌های صدور مجوز از بالا به پایین دفاع می‌کنند. شکوفاکنندگان معتقدند آزادی نوآوری و تعهد به تنظیم مقررات هر دو مهم هستند و ما باید تعادل درستی بین این دو برقرار کنیم. تا حدی، این بدان معناست که باید اذعان کنیم که مردم دیدگاه‌های بسیار متفاوتی در مورد هوش مصنوعی دارند و هدف اصلی ما باید پیگیری اجماع و اعتماد گسترده باشد. و این چیزی است که به احتمال زیاد از دسترسی و استفاده پدید می‌آید. به عبارت دیگر، وضع قوانین درباره هوش مصنوعی کافی نیست. باورها، هنجارها و ارزش‌های مشترک نیز اهمیت دارند.

همانطور که در فصل‌های آینده بررسی خواهیم کرد، هوش مصنوعی فناوری‌ای است که می‌تواند به طور چشمگیری کاربران فردی را توانمند کند، در حالی که در سطوح بسیار گسترده اجتماعی نیز عمل می‌کند. از این نظر، شبیه خود اینترنت، یا جاده‌ها و بزرگراه‌ها، یا شبکه برق، یا سیستم آب است. به درجات مختلف، به دلایل مختلف، افرادی هستند که از همه این سیستم‌ها

انصراف می‌دهند، یا حداقل بدون آنها زندگی می‌کنند. اما تصور کنید که مثلاً ۳۰ درصد از مردم نسبت به خطوط برق نگرانی‌های عمیقی داشته باشند. یا ۲۰ درصد اصرار داشته باشند که چراغ راهنمایی را در زندگی خود جای ندهند. در نهایت، فرض کنید درصد بسیار کمی از مردم، نه بیش از چند هزار نفر، یا حتی چند صد نفر، از مخازن آب آنقدر متنفر باشند که بخش زیادی از وقت خود را صرف تلاش فعالانه برای تخریب یا تضعیف آنها کنند.

هیچ یک از این رفتارها احتمالاً تعهدات ما را نسبت به خدمات و زیرساخت‌هایی که ما را قادر می‌سازند زندگی‌های پرباری در قرن بیست و یکم داشته باشیم، تغییر نخواهد داد. اما احتمالاً بر نحوه عملکرد روان و مولد ما، به عنوان یک کشور، با پیامدهای داخلی و جهانی، تأثیر خواهد گذاشت. به همین دلیل بسیار مهم است که هوش مصنوعی را به روش‌هایی توسعه دهیم که نقش واقعی به مردم در این فرآیند بدهد، درست همانطور که هنگام توسعه خودروها و اینترنت انجام دادیم. بدیهی است که اجماع یکنواخت هم بعید و هم غیرضروری است. اما با این همه چیز در خطر، منطقی است که به روش‌هایی پیش برویم که بهترین شانس موفقیت را به ما بدهد.

این عمدتاً به این دلیل است که یک جنبه نامطلوب حکمرانی در عصر شبکه‌های دیجیتال این است که جوامعی که در آنها مشارکت دموکراتیک گسترده‌تر است، همچنین جوامعی هستند که در آنها ایجاد تغییر سخت‌تر می‌شود. آشفتگی بر جامعه و اجماع غلبه می‌کند. بنابراین این چالش ویژه است که اکنون با ایالات متحده و سایر دموکراسی‌ها روبرو است. اگر ما کوتاهی کنیم، این احتمال بسیار واقعی وجود دارد که کشورهای اقتدارگرا یا حداقل کشورهای با همبستگی اجتماعی بیشتر، فناوری‌های هوش مصنوعی را مؤثرتر از ما به کار گیرند و مستقر کنند. به تمام روش‌هایی فکر کنید که این می‌تواند عاملیت فردی را تضعیف کند. ما قبلاً می‌دانیم که ابزارهای هوش مصنوعی را می‌توان طوری طراحی کرد که بر روی شما کار کنند تا برای شما یا با شما، حتی در دموکراسی‌ها. در خودکامگی‌ها، این احتمال افزایش می‌یابد. همچنین رفاه ملی و امنیت ملی را در نظر بگیرید. حمایت‌گرایی به ما اجازه رقابت اقتصادی با کشورهایی را که مجهز به میلیون‌ها

کارگر رباتیک بسیار باهوش و تعداد بیشتری از سیستم‌های هوش مصنوعی هستند که به عنوان دانشمندان و مهندسان مصنوعی عمل می‌کنند، نخواهد داد. انزواگرایی ما را از سلاح‌های خودمختار جدید و عجیب و غریبی که آن دانشمندان و مهندسان مصنوعی ممکن است ابداع کنند، محافظت نخواهد کرد.

تمام این‌ها بدان معناست که در قرن بیست و یکم، عاملیت فردی بیش از هر زمان دیگری با عاملیت ملی همسو است. برای اطمینان از آینده هوش مصنوعی که به عنوان "گسترشی از اراده‌های انسانی فردی" عمل می‌کند، دموکراسی‌ها باید این تلاش را رهبری کنند. این بدان معناست که با یک ذهنیت نوآوری بلندپروازانه، تعهد به آزمایش و سازگاری، و مشارکت گسترده عمومی برای ایجاد اجماع در طول مسیر، توسعه یابد.

نتیجه‌گیری

ظهور هوش مصنوعی و به طور خاص مدل‌های زبانی بزرگ مانندChatGPT ، نشان‌دهنده یک نقطه عطف حیاتی در تاریخ بشریت است که با تحولات اقتصادی و اجتماعی اواخر سال ۲۰۲۲ همزمان شد. این فصل به تفصیل نشان داد که چگونه هوش مصنوعی، با وجود چالش‌ها و نگرانی‌های اولیه، به سرعت توانایی خود را برای تغییر چشم‌انداز فناوری و حتی ساختار اجتماعی به اثبات رسانده است. بررسی فنی این مدل‌ها، از معماری شبکه عصبی گرفته تا پدیده‌هایی مانند "توهمات" و "جعبه سیاه"، محدودیت‌های ذاتی آن‌ها را روشن ساخت، اما در عین حال پتانسیل بی‌نظیرشان را نیز برجسته کرد.

عاملیت انسانی به عنوان محور اصلی این گفتمان مطرح شد؛ هوش مصنوعی این پتانسیل را دارد که عاملیت ما را به شکلی بی‌سابقه افزایش دهد و ما را به سمت "ابراژانس" سوق دهد، همانطور که انقلاب صنعتی با نیروی بخار توانست بهره‌وری و ظرفیت‌های انسانی را دگرگون کند. با این حال، این پیشرفت عظیم با چالش‌های اخلاقی، اجتماعی و سیاسی همراه است که نیازمند توجه دقیق به مسائلی مانند جابجایی شغلی، اطلاعات غلط، حریم خصوصی و تعصبات

الگوریتمی است. رویکرد OpenAI در "استقرار تکراری" و تأکید بر مشارکت عمومی، راهبردی را برای توسعه مسئولانه و دموکراتیک هوش مصنوعی ارائه می‌دهد که در آن اعتماد متقابل میان توسعه‌دهندگان، کاربران و تنظیم‌کنندگان حیاتی است. در نهایت، آینده هوش مصنوعی نه تنها یک مسئله فناوری، بلکه یک چالش جهانی برای دموکراسی‌هاست تا اطمینان حاصل کنند که این فناوری جدید به عنوان گسترشی از اراده انسانی عمل می‌کند و به جای تضعیف، عاملیت فردی و جمعی را تقویت می‌کند. مسیری که در پیش داریم، نیازمند نوآوری جسورانه، آزمایش مداوم و مشارکت گسترده عمومی برای شکل‌دهی به آینده‌ای فراگیر و مفید برای همگان است.

نکات کلیدی

- اواخر سال ۲۰۲۲ شاهد بحران‌های اقتصادی و تعدیل نیروی گسترده در شرکت‌های بزرگ فناوری بود، در حالی که تقاضای مصرف‌کننده و رشد مشاغل نیز در جریان بود.
- معرفی ChatGPT در نوامبر ۲۰۲۲ به سرعت ۱۰۰ میلیون کاربر را در دو ماه جذب کرد و توجه صنعت فناوری را به هوش مصنوعی مولد جلب نمود.
- ChatGPTاز طریق معماری شبکه عصبی و پارامترهای متعدد، بر اساس پیش‌بینی توکن‌ها عمل می‌کند و درکی انسانی از مفاهیم ندارد.
- مدل‌های زبانی بزرگ (LLMs) مستعد تولید اطلاعات نادرست ("توهمات") هستند و ممکن است به دلیل داده‌های آموزشی، تعصباتی را بازتولید کنند.
- پدیده "جعبه سیاه" به دلیل پیچیدگیLLMs ، توضیح فرآیندهای تصمیم‌گیری آن‌ها را دشوار یا ناممکن می‌سازد.
- با وجود پیشرفت‌ها، برخی شکاکان معتقدند LLMs هرگز به هوش عمومی مصنوعی (AGI)دست نخواهند یافت و محدودیت‌های ذاتی خواهند داشت.
- هوش مصنوعی به عنوان یک "ابزار" پتانسیل افزایش عاملیت انسانی را دارد، همانطور که انرژی و هوش در طول تاریخ پیشرفت انسانی را شکل داده‌اند.

- مقایسه با انقلاب صنعتی نشان می‌دهد که هوش مصنوعی ترکیبی می‌تواند جهش‌های بزرگی در بهره‌وری و شکوفایی انسانی ایجاد کند، با وجود چالش‌های اولیه.
- فلسفه OpenAI بر این اساس است که هوش مصنوعی باید "گسترشی از اراده‌های انسانی فردی" باشد و به طور گسترده توزیع شود.
- "استقرار تکراری (Iterative Deployment) "رویکردی است که بر مشارکت عمومی و بازخورد کاربر برای توسعه تدریجی و مسئولانه هوش مصنوعی تأکید دارد.
- در گفتمان هوش مصنوعی، چهار دیدگاه اصلی وجود دارد: نابودی‌خواهان، سیاه‌بینان، سریع‌روندگان و شکوفاکنندگان، که هر یک دیدگاهی متفاوت در مورد خطرات و پتانسیل‌ها ارائه می‌دهند.
- عاملیت فردی و ملی در قرن ۲۱ به طور فزاینده‌ای به هم پیوسته‌اند و دموکراسی‌ها باید با نوآوری و مشارکت عمومی، آینده هوش مصنوعی را رهبری کنند.

سؤالات تفکربرانگیز

- چگونه می‌توان تعادل مناسبی بین سرعت نوآوری در هوش مصنوعی و نیاز به تنظیم مقررات مسئولانه برقرار کرد تا از آسیب‌های احتمالی جلوگیری شود؟
- با توجه به محدودیت‌های ذاتی مدل‌های زبانی بزرگ، مانند "توهمات" و "جعبه سیاه"، چگونه می‌توان اعتماد عمومی به این فناوری‌ها را تقویت کرد؟
- چگونه می‌توانیم اطمینان حاصل کنیم که هوش مصنوعی به عنوان ابزاری برای تقویت عاملیت انسانی عمل می‌کند، نه اینکه آن را تحلیل برده یا دستکاری کند؟
- درس‌های انقلاب صنعتی در مورد پیامدهای اجتماعی و اقتصادی فناوری‌های تحول‌آفرین چه بینش‌هایی را برای مدیریت انتقال به عصر هوش مصنوعی ارائه می‌دهد؟
- با توجه به نقش جهانی در توسعه و استقرار هوش مصنوعی، دموکراسی‌ها چگونه

می‌توانند از نفوذ مدل‌های هوش مصنوعی طراحی‌شده برای کار "بر روی شما" به جای "برای شما" جلوگیری کنند؟

۳۰ جمله کلیدی

۱. اواخر سال ۲۰۲۲، با موجی از اخراج‌ها در شرکت‌های فناوری و ورشکستگی‌های بزرگ، روایت قدرت مطلق "بیگ تک" به چالش کشیده شد.

۲. در این دوران عدم اطمینان، OpenAI با معرفی ChatGPT ، تغییرات چشمگیری را در صنعت فناوری رقم زد.

۳. ChatGPT، با دانش چشمگیر، کارایی خیره‌کننده و شباهت متقاعدکننده به انسان، به سرعت میلیون‌ها کاربر را جذب کرد.

۴. این سیستم حتی با اشتباهاتش، که "توهمات" نامیده می‌شدند، کنجکاوی و شگفتی ایجاد می‌کرد.

۵. در واکنش به ChatGPT ، غول‌های فناوری مانند گوگل و مایکروسافت، هوش مصنوعی را در اولویت اصلی خود قرار دادند.

۶. احساسات عمومی نسبت به هوش مصنوعی ۱۸۰ درجه تغییر کرد، از نگرانی‌های ضدانحصار به دغدغه‌های مربوط به سرعت بیش از حد نوآوری.

۷. مدل‌های زبانی بزرگ (LLMs) بر اساس معماری شبکه عصبی عمل می‌کنند و از طریق تنظیم پارامترها، الگوهای پیچیده زبان را می‌آموزند.

۸. LLMs هرگز حقایق را نمی‌دانند یا مفاهیم را به شیوه انسانی درک نمی‌کنند؛ آن‌ها صرفاً توکن‌های احتمالی بعدی را پیش‌بینی می‌کنند.

۹. "توهمات" زمانی رخ می‌دهند که مدل‌ها اطلاعات نادرست یا نامربوط تولید می‌کنند که مبنایی در واقعیت ندارند.

۱۰. مدل‌های هوش مصنوعی عینیت یا بی‌طرفی ندارند؛ آن‌ها نتیجه انتخاب‌های انسانی در جمع‌آوری و پردازش داده‌ها هستند.

۱۱. تعصبات موجود در داده‌های آموزشی می‌توانند منجر به خروجی‌های جنسیتی یا نژادپرستانه از سوی هوش مصنوعی شود.

۱۲. پدیده "جعبه سیاه" توضیح فرآیندهای تصمیم‌گیری در LLMs پیچیده را دشوار می‌سازد.

۱۳. محدودیت‌های اساسی LLMs شامل عدم توانایی در استدلال مشترک، تجربه زیسته و مدل زمینی‌شده از جهان است.

۱۴. با وجود پیشرفت‌ها، شکاکان نسبت به دستیابی LLMs به هوش عمومی مصنوعی (AGI)تردید دارند.

۱۵. هوش مصنوعی موجب یک "رویارویی وجودی" می‌شود، که شامل نگرانی‌هایی در مورد مشاغل، اطلاعات غلط، حریم خصوصی و خودمختاری است.

۱۶. "عاملیت انسانی" توانایی فرد برای انتخاب مستقل و تأثیرگذاری بر زندگی خود است.

۱۷. هوش مصنوعی می‌تواند به طور فزاینده‌ای بر حوزه‌هایی نفوذ کند که به طور سنتی توسط عاملیت انسانی اداره می‌شوند.

۱۸. بشریت به عنوان "Homo techne" با ابزارسازی و استفاده از فناوری برای گسترش عاملیت خود تعریف می‌شود.

۱۹. هوش مصنوعی به عنوان "ابزار" جدیدی تلقی می‌شود که می‌تواند عاملیت انسانی را به سطح بی‌سابقه‌ای برساند.

۲۰. هوش و انرژی، در کنار هم، عاملیت و پیشرفت انسانی را به حرکت درمی‌آورند.

۲۱. برای اولین بار، "هوش مصنوعی ترکیبی" مانند "انرژی ترکیبی" انقلاب صنعتی، به

ابزاری منعطف و مقیاس‌پذیر تبدیل شده است.

۲۲. "ابراژانس" وضعیتی است که افراد توانمند شده توسط هوش مصنوعی، در سطوحی فعالیت می‌کنند که تأثیرات آن‌ها در جامعه ترکیب می‌شود.

۲۳. انقلاب صنعتی، با وجود چالش‌های اولیه، در مجموع نیرویی انسانی‌ساز و عامل پیشرفت اجتماعی بود.

۲۴. هوش مصنوعی می‌تواند جهش تصاعدی مشابهی در "انقلاب شناختی" ایجاد کند که پتانسیل و عاملیت انسانی را گسترش دهد.

۲۵. فلسفه OpenAI بر این مبنا است که هوش مصنوعی باید "گسترشی از اراده‌های انسانی" باشد و به طور گسترده توزیع شود.

۲۶. "استقرار تکراری (Iterative Deployment) "رویکردی است که با مشارکت عمومی، اعتماد به فناوری و جامعه را در طول زمان می‌سازد.

۲۷. گفتمان هوش مصنوعی شامل چهار دیدگاه اصلی است: نابودی‌خواهان، سیاه‌بینان، سریع‌روندگان و شکوفا کنندگان.

۲۸. شکوفا کنندگان، که نویسنده نیز با آنها همسو است، معتقدند که مشارکت گسترده و دسترسی دموکراتیک به هوش مصنوعی برای توسعه مسئولانه آن ضروری است.

۲۹. دموکراتیزه کردن دسترسی به هوش مصنوعی به مردم امکان می‌دهد تا در شکل‌دهی آینده خود نقش فعال‌تری داشته باشند.

۳۰. عاملیت فردی و ملی در قرن ۲۱ بیش از پیش به هم پیوسته‌اند و دموکراسی‌ها باید در توسعه هوش مصنوعی با نوآوری و مشارکت عمومی پیشگام باشند.

فصل ۲

دانش بزرگ: از سایه اورول تا عصر هوش مصنوعی و ساختارهای اعتماد دیجیتال

" :در عصر حاضر، داده‌ها نه تنها سوخت موتور پیشرفت‌اند، بلکه سنگ‌بنای درک عمیق‌تر ما از جهان و خودمان را تشکیل می‌دهند. اما تبدیل این حجم عظیم از اطلاعات به دانشی قابل بهره‌برداری، مستلزم تعادلی ظریف میان حفظ حریم خصوصی و خلق هویت عمومی مبتنی بر اعتماد است — ".پژوهشگر برجسته حوزه اطلاعات

اهداف اصلی فصل:

- آشنایی با ریشه‌های نگرانی‌های اجتماعی درباره نظارت فناورانه و نقض حریم خصوصی در طول تاریخ.
- بررسی سیر تحول دیدگاه‌ها نسبت به جمع‌آوری و اشتراک‌گذاری داده‌ها از میانه قرن بیستم تاکنون.
- تحلیل نقش نوآوری‌های فناورانه در شکل‌دهی به مفاهیم هویت فردی و جمعی در دنیای دیجیتال.
- شناخت چگونگی تأثیر پلتفرم‌های مبتنی بر اعتماد بر گسترش دانش و فرصت‌های فردی در دنیای متصل امروز.
- درک اهمیت هوش مصنوعی در تبدیل داده‌های عظیم به دانش کاربردی و افزایش توانمندی‌های انسانی برای ورود به عصری نوین.

چکیده

این فصل به بررسی مفهوم "دانش بزرگ" و سیر تحول آن در بستر نوآوری‌های فناورانه می‌پردازد. آغازگر این بحث، نگاه پیشگویانه جورج اورول در رمان "۱۹۸۴" است که تصویری هراس‌انگیز از نظارت دولتی و خردکننده ارائه داد و بنیان‌گذار نگرانی‌های عمومی درباره فناوری شد. در ادامه، فصل به ظهور کامپیوترهای بزرگ در دهه ۱۹۶۰ و واکنش‌های دوگانه نسبت به آن‌ها، از جمله وحشت از "جامعه عریان" و در عین حال استفاده‌های کاربردی بی‌سابقه، می‌پردازد. طرح "مرکز داده فدرال" در آمریکا و مخالفت‌های شدید با آن، نمونه‌ای بارز از اوج‌گیری نگرانی‌ها در آن دوره است که با مفاهیم نوظهور حریم خصوصی پیوند خورد. با این حال، با گسترش کامپیوترهای شخصی و اینترنت، فناوری به جای آنکه عامل سرکوب باشد، به ابزاری برای گسترش عاملبودگی فردی، تنوع فرهنگی و شخصی‌سازی تبدیل شد. پلتفرم‌هایی مانند لینکدین نشان دادند که چگونه اعتماد مبتنی بر هویت عمومی و شبکه ارتباطات می‌تواند به خلق ارزش جمعی و "دانش بزرگ" بینجامد. در نهایت، فصل به چالش‌های عصر حاضر، از جمله نگرانی‌های ناشی از هوش مصنوعی و حجم بی‌سابقه داده‌ها می‌پردازد و استدلال می‌کند که هوش مصنوعی، با وجود خطرات احتمالی، ابزاری حیاتی برای تبدیل "داده‌های بزرگ" به "دانش بزرگ" و گشودن راه به سوی "عصر روشنگری" نوین است، جایی که اطلاعات نه تنها محدودکننده نیستند بلکه قدرت‌بخش و توانمندساز محسوب می‌شوند.

مقدمه

در میانه قرن بیستم، زمانی که جورج اورول رمان "۱۹۸۴" را منتشر کرد، جهان تنها تعداد انگشت‌شماری کامپیوتر داشت و شبکه‌های تلویزیونی در مراحل ابتدایی خود بودند. با این حال، دیدگاه اورول درباره چگونگی شکل‌گیری زندگی ما توسط شبکه‌های صوتی و تصویری الکترونیکی، به وضوح بسیار پیشگویانه بود و به همین دلیل است که "۱۹۸۴" همچنان پس از گذشت دهه‌ها، در فهرست پرفروش‌ترین کتاب‌ها ظاهر می‌شود. این رمان، تصویری تلخ از یک ابرقدرت توتالیتر به نام اقیانوسیه را ارائه می‌دهد که شهروندانش را تحت سلطه شدید نظارت بی‌وقفه، شکنجه‌های مکرر و پروپاگاندای پوچ قرار داده است. در جهانی که شعارهایی چون "نادانی قدرت است" و "آزادی بردگی است" پذیرفته شده‌اند، تنها اطاعت کامل باقی می‌ماند. این تصویر تاریک از نظارت خداگونه فناورانه و اثرات غیرانسانی آن، دیدگاه بشریت را برای همیشه دگرگون کرد و الهام‌بخش نسل‌های متعددی شد که آینده‌ای **آخرالزمانی** را پیش‌بینی می‌کردند.

با این حال، سیر تکاملی فناوری و جامعه نشان داده است که این داستان بسیار پیچیده‌تر از پیش‌بینی‌های اولیه است. در حالی که نگرانی‌ها در مورد حریم خصوصی و نظارت همچنان معتبرند، فناوری در بسیاری جهات مسیرهای غیرمنتظره‌ای را گشوده که به جای محدود کردن، به توانمندسازی فردی و جمعی منجر شده است. هدف این فصل، کاوش در این دوگانگی و بررسی چگونگی تبدیل "داده‌های بزرگ" – حجم عظیمی از اطلاعات که امروزه تولید می‌شود – به "دانش بزرگ" است. این دانش نه تنها به سازمان‌ها قدرت می‌دهد، بلکه به افراد نیز کمک می‌کند تا در یک جهان پیچیده و متصل، عامل‌بودگی بیشتری داشته باشند. از ترس‌های اولیه مرتبط با کامپیوترهای غول‌پیکر تا ظهور هوش مصنوعی پیشرفته، این فصل به تحولات اساسی در درک ما از اطلاعات، حریم خصوصی، هویت و اعتماد می‌پردازد و نشان می‌دهد که چگونه می‌توان از قدرت فناوری برای ساختن آینده‌ای روشن‌تر بهره برد.

سایه‌های ۱۹۸۴: ریشه‌های نگرانی‌های فناورانه

رمان "۱۹۸۴" جهانی را به تصویر می‌کشد که در آن انگلستان سابق، اکنون "بخش اول پرواز" نامیده می‌شود و سومین استان پرجمعیت اقیانوسیه است. این ابرقدرت توسط رژیمی به نام "حزب" و چهره همه‌جا حاضر آن، "برادر بزرگ"، اداره می‌شود که حتی ممکن است یک دیپ‌فیک باشد. حزب با ترکیبی خردکننده از نظارت دائمی، شکنجه‌های مکرر و تبلیغاتی آنقدر پوچ که هرگونه میل به منطق، کرامت یا امید را از بین می‌برد، شهروندان خود را به وحشیانه‌ترین شکل ممکن سرکوب می‌کند. در اقیانوسیه، همه چیز خاکستری و لکه‌دار است؛ خیابان‌ها و آسمان، گذشته و آینده، حتی خود حقیقت. تنها پوسترهای رهبر سبیلو کشور، اشاره‌ای به پویایی و رنگ دارند، و حتی آن‌ها نیز تاریک و تهدیدآمیز هستند. شعار "برادر بزرگ تو را می‌بیند" نه تنها اطمینان‌بخش نیست، بلکه نویدبخش نظارت همه‌جانبه است.

همه‌جا حاضر بودن دولت از طریق شبکه‌ای جامع از "تلسکرین‌ها"ی عمومی و خصوصی انجام می‌شود که به طور مداوم تبلیغات دولتی را پخش می‌کنند و در عین حال شهروندان اقیانوسیه را زیر نظر دارند. این شبکه توسط "پلیس فکر" نظارت می‌شود که به دقت نشانه‌های بی‌وفایی به برادر بزرگ یا حزب را، و همچنین هرگونه ابراز احساسات انسانی، زندگی درونی، یا انحراف از هنجار را جستجو می‌کنند. در اقیانوسیه، حتی کوچک‌ترین نشانه از علاقه‌ای فراتر از مرزهای حزب، به عنوان عدم اطاعت تلقی می‌شود. اورول با ارائه چنین دیدگاه وحشتناکی از نظارت فناورانه در سطح خدایی و اثرات غیرانسانی آن، تأثیری عمیق بر ذهنیت عمومی گذاشت. برای بیش از هفتاد و پنج سال، نسل‌های متعددی از افراد بدبین و نگران، پیش‌بینی آخرالزمانی قریب‌الوقوع را مطرح کرده‌اند، پیش‌بینی‌هایی که ریشه در ترس از کنترل مطلق و از دست دادن فردیت دارند.

ظهور غول‌های محاسباتی و آغاز دوگانگی نگاه به داده

در اوایل دهه ۱۹۶۰، با افزایش اهمیت "کامپیوترهای بزرگ" یا همان "مین‌فریم‌ها"، موجی از این پیش‌بینی‌های بدبینانه به اوج خود رسید. تا سال ۱۹۶۳، صنعت کامپیوتر ایالات متحده حدود

۱۵,۵۰۰ کامپیوتر تولید کرده بود. حدود هزار دستگاه از آن‌ها توسط دولت ایالات متحده اجاره یا خریداری شده بود و مابقی در شرکت‌های بزرگ و دانشگاه‌ها جای گرفته بودند، جایی که اعداد را با سرعت بی‌سابقه‌ای پردازش می‌کردند و به این مؤسسات قدرت ادراک و بینشی می‌بخشیدند که قبلاً غیرممکن بود. به عنوان مثال، هنگامی که سازمان امور مالیاتی در سال ۱۹۶۳ اعلام کرد که یک دستگاه IBM 7074 به ارزش ۱٫۲ میلیون دلار را برای پردازش و تأیید اظهارنامه‌های مالیاتی به کار گرفته است، بسیاری از مودیان نگران به دفاتر سازمان هجوم آوردند تا بدهی‌های مالیاتی پرداخت نشده خود را تسویه کنند. در مجموع، یک مقام IRS اعلام کرد که این سازمان ۷۰۰,۰۰۰ دلار مالیات پرداخت نشده جمع‌آوری کرده است.

حتی در این مرحله اولیه از توسعه، گستره کارهایی که کامپیوترها انجام می‌دادند، چشمگیر بود. سازمان‌های خبری از آن‌ها برای پیش‌بینی نتایج انتخابات استفاده می‌کردند. تولیدکنندگان بستنی برای کنترل دقیق فرآیندهای ترکیب خود از آن‌ها بهره می‌بردند. شهر نیویورک از کامپیوترها برای مدیریت الگوهای زمان‌بندی چراغ‌های راهنمایی استفاده می‌کرد. حتی آسترید لیندگرن، نویسنده سوئدی کتاب‌های "پی‌پی جوراب‌بلند"، ۹,۰۰۰ کرون اضافی (حدود ۱۱,۷۰۰ دلار آمریکا در سال ۲۰۲۴) بابت حق امتیاز از سیستم کتابخانه‌های عمومی سوئد دریافت کرد، جایی که یک سیستم جبران خسارت نویسنده برای امانت گرفتن عمومی کتاب‌ها از سال ۱۹۵۴ راه‌اندازی شده بود. این پرداخت بر اساس ۸۵۰,۰۰۰ امانت کل کتاب‌های او از هزاران مدرسه و کتابخانه بود و چنین حسابداری تنها با یک کامپیوتر الکترونیکی ممکن بود.

البته، آنچه می‌توانست به عنوان یک "عصر روشنگری" جدید توصیف شود – جایی که مؤسسات از هر نوع از مین‌فریم‌ها برای جمع‌آوری داده‌های بیشتر، دستیابی به وضوح جدید در عملیات خود و انجام اقدامات بر اساس این دانش جامع‌تر استفاده می‌کردند – جنبه تاریکی نیز داشت. ونس پاکارد، یکی از پرخواننده‌ترین نویسندگان دوران خود، در کتاب پرفروش "جامعه عریان" در سال ۱۹۶۴ هشدار داد: "بانک‌هایی از ماشین‌های حافظه غول‌پیکر وجود دارند که به طور بالقوه می‌توانند در چند ثانیه هر اقدام مرتبط – از جمله شکست‌ها، شرمندگی‌ها، یا شاید

اقدامات مجرمانه – را از طول عمر هر شهروند به یاد آورند." او همچنین به پیشرفت تحقیقات مغزی اشاره کرد و اظهار داشت که "بسیار باورپذیر است که یک برادر بزرگ بتواند در زمان تولد در مغز هر نوزاد یک الکترود بکارد و پس از آن با کنترل از راه دور، درجه‌ای از محدودیت را بر حالت‌ها و رفتار فرد اعمال کند، حداقل تا زمانی که شخصیت او به طور مناسب شکل گرفته باشد." این دیدگاه، ترس از کنترل همه‌جانبه و از دست رفتن کامل عامل‌بودگی فردی را در کانون توجه قرار داد. با این حال، همان ماشین‌های حافظه غول‌پیکر، با احتمالی کمتر نگران‌کننده، می‌توانستند به اقتصاددانان کمک کنند تا داده‌های سلامت عمومی را به گونه‌ای تحلیل کنند که روابط پیچیده بین عوامل اقتصادی، استفاده از مراقبت‌های بهداشتی و نتایج سلامت برای جمعیت‌های مختلف را آشکار سازد.

رویای مرکز داده فدرال و طوفان حریم خصوصی

در آوریل ۱۹۶۵، کمیته حفظ و استفاده از داده‌های اقتصادی، گروهی از دانشمندان اجتماعی که به درخواست دولت فدرال برای چندین سال کار می‌کردند، گزارشی منتشر کرد که پیشنهاد می‌کرد دولت یک "مرکز داده فدرال" را تأسیس کند. این مرکز می‌توانست دسترسی متمرکزتر و آسان‌تری به حجم فزاینده داده‌هایی که آژانس‌های مختلف دولتی جمع‌آوری می‌کردند، فراهم کند. ایده اصلی این بود که طیف وسیعی از داده‌ها را که بیست آژانس فدرال مختلف جمع‌آوری کرده بودند، گردآوری کرده و آن‌ها را برای محققان قابل دسترس‌تر و قابل استفاده‌تر سازد. این به نوبه خود می‌توانست به سیاست‌گذاری فدرال در زمانی کمک کند که دولت در حال گسترش برنامه‌های اجتماعی به عنوان بخشی از "چشم‌انداز جامعه بزرگ" رئیس‌جمهور لیندون بی. جانسون بود. اما در آن زمان، داده‌هایی که دولت شروع به جمع‌آوری آن‌ها کرده بود، در بیش از ۶۰۰ مجموعه داده، به صورت سیلویی و به روشی بسیار غیرمتمرکز و اغلب ناسازگار، روی بیش از ۱۰۰ میلیون کارت پانچ و ۳۰،۰۰۰ نوار مغناطیسی ذخیره شده بودند. دانشمندان اجتماعی استدلال کردند که یکپارچه‌سازی این داده‌ها در یک مکان فیزیکی واحد – یک مرکز داده ملی – داده‌ها را مفیدتر کرده و بنابراین ارزش بیشتری برای مالیات‌دهندگانی که هزینه جمع‌آوری آن را پرداخته

بودند، ایجاد می‌کند.

زمان‌بندی کمیته نمی‌توانست بدتر باشد. همراه با شیوع رو به رشد کامپیوترهای مین‌فریم، پیشرفت‌های فناورانه در طراحی ترانزیستور و مدارهای چاپی، تولیدکنندگان را قادر ساخته بود تا دستگاه‌های الکترونیکی را هم‌زمان کوچک‌تر و قدرتمندتر سازند. کاتالوگ‌های پستی در مجلات ماجراجویی مردانه، میکروفون‌هایی به اندازه حبه قند و دوربین‌های مداربسته که می‌توانستند در جیب جلیقه پنهان شوند، تبلیغ می‌کردند. محققان خصوصی، جاسوسان بالقوه شرکتی و افراد کنجکاو می‌توانستند میکروفون‌های پارابولیک که قادر به شنیدن مکالمات در فاصله یک بلوک شهری بودند، ضبط‌کننده‌های صوتی فعال و دوربین‌های مادون قرمز که می‌توانستند در تاریکی کامل عکس بگیرند، خریداری کنند. همه اینها، به گفته بدبینان میانه قرن، به سوی جهانی فراابزاری مانند "بخش اول پرواز" و "اقیانوسیه" اورول اشاره داشت. به عنوان مثال، مایران برنتون در کتاب "مهاجمان حریم خصوصی" نوشت: "ترسناک‌ترین پیامد برای آینده، مفهوم نظارت داخلی است – یعنی سیستم‌های نظارتی که مستقیماً در هتل‌ها، مدارس، زندان‌ها، ساختمان‌های اداری یا سایر سازه‌های جدید ساخته شده‌اند. در این کابوس اورولی، سیم‌کشی‌ها می‌توانند به یک ایستگاه مرکزی شنود و ضبط منتهی شوند".

همراه با نگرانی‌های عمومی فزاینده درباره انواع جدید جاسوسی الکترونیکی، کمیته حفظ و استفاده از داده‌های اقتصادی با پرونده "گریزولد علیه کنتیکت" نیز مواجه شد. تنها دو ماه پس از انتشار گزارش این دانشمندان اجتماعی، دیوان عالی قانونی در کنتیکت را که استفاده از وسایل پیشگیری از بارداری را ممنوع می‌کرد، لغو کرد. همانطور که سارا ای. ایگو، تاریخ‌دان، در کتاب "شهروند شناخته شده: تاریخچه حریم خصوصی در آمریکای مدرن" پیشنهاد می‌کند، حکم دادگاه در "گریزولد علیه کنتیکت" برای اولین بار در تاریخ آمریکا "یک 'حق حریم خصوصی' قانون اساسی را شناسایی کرد." علاوه بر این، "امکان ایجاد ادعاهای جدید حریم خصوصی و همچنین دور زدن دیگران" را فراهم آورد. بنابراین، زمان بسیار نامناسبی بود که اعلام شود دولت ایالات متحده "جامع‌ترین برنامه آماری در جهان" را گردآوری کرده و برنامه‌های بلندپروازانه‌ای برای

دسترس‌پذیری بیشتر آن اطلاعات به تکنوکرات‌های روشنفکر دارد. هنگامی که دولت جانسون اعلام کرد که با این نمونه از خوش‌بینی "جامعه بزرگ" و حکمرانی خوب پیش می‌رود، واکنش کنگره "سریع و شدید" بود. تعدادی از سناتورها و اعضای کنگره قبلاً جلساتی را برای رسیدگی به مسائل مربوط به حریم خصوصی که در کتاب‌هایی مانند پاکارد و برنتون مطرح شده بود، برگزار کرده بودند. آن‌ها به سرعت یک جلسه دیگر به تقویم خود اضافه کردند.

در جلسه کمیته فرعی ژوئیه ۱۹۶۶ با عنوان "کامپیوتر و تهاجم به حریم خصوصی"، قانونگذاران اذعان کردند که لزوماً مخالف هدف مرکز داده ملی – یعنی استفاده از داده‌های واقعی برای اطلاع‌رسانی به سیاست‌ها و قوانین دولتی – نیستند، اما نگران بودند که چنین مرکزی ممکن است به چه چیزی تبدیل شود. کورنلیوس ای. گالاگر، نماینده دموکرات کنگره از منطقه سیزدهم نیوجرسی، هشدار داد: "ادعای ما این است که اگر تمهیدات حفاظتی در چنین تأسیساتی ایجاد نشود، می‌تواند به ایجاد چیزی که من آن را 'انسان کامپیوتری‌شده' می‌نامم، منجر شود. 'انسان کامپیوتری‌شده'، آن‌گونه که من او را می‌بینم، از فردیت و حریم خصوصی خود محروم خواهد شد. از طریق استانداردسازی که توسط پیشرفت فناوری به ارمغان آورده می‌شود، وضعیت او در جامعه توسط کامپیوتر اندازه‌گیری خواهد شد و او هویت شخصی خود را از دست خواهد داد. زندگی، استعداد و ظرفیت کسب درآمد او به یک نوار با گزینه‌های بسیار محدود کاهش می‌یابد." فرانک هورتون، نماینده جمهوری‌خواه از نیویورک، نیز تکرار کرد: "این خطر وجود دارد که کامپیوترها، چون ماشین هستند، با ما به عنوان ماشین رفتار کنند. آن‌ها می‌توانند حقایق را ارائه دهند و در واقع، ما را از تولد تا مرگ هدایت کنند. آن‌ها می‌توانند ما را همانطور که نوارهایشان حکم می‌کنند، 'قفسه‌بندی' کنند، و در یک محدوده باریک، تحصیلاتی که می‌گیریم، مشاغلی که انجام می‌دهیم، پولی که می‌توانیم کسب کنیم و حتی دختری که با او ازدواج می‌کنیم را انتخاب کنند".

ونس پاکارد خود به ساختمان تازه تکمیل شده ریبرن هاوس در کاپیتول هیل رفت تا کمیته فرعی را در مورد اینکه چگونه "نگهدارندگان پرونده" دولتی و بخش خصوصی میلیاردها سابقه

درباره میلیون‌ها شهروند ایالات متحده، پر از "اطلاعات تحقیرآمیز از یک نوع یا دیگری" درباره هر جنبه از زندگی آن‌ها، جمع‌آوری کرده‌اند، آموزش دهد. پاکارد پیشنهاد کرد که از طریق اظهارنامه‌های مالیاتی، ارزیابی‌های شخصیتی مربوط به شغل، تست‌های دروغ‌سنجی، سوابق پزشکی، گزارش‌های اعتباری، تحقیقات شرکت بیمه، سوابق پلیس و حتی موجودی شرکت‌های حمل و نقل، مقامات می‌توانستند پرونده‌هایی را جمع‌آوری کنند که شامل اطلاعاتی درباره وضعیت سلامت و رفاه عاطفی افراد، زندگی جنسی آن‌ها، وضعیت ازدواجشان، مواجهاتشان با قانون و ارتباطات و وابستگی‌های شناخته‌شده‌شان باشد. به تخمین پاکارد، ایجاد "یک سیستم غول‌پیکر که بتواند برای درج یا بازیابی اطلاعات از مکان‌های متعدد فعال شود"، در نهایت به "غیرشخصی شدن شیوه زندگی آمریکایی" و "حس خفقان‌آور نظارت با آگاه شدن مردم از اینکه دولتشان در حال توسعه یک چشم همه‌بین است" منجر خواهد شد. پاکارد در پایان شهادت خود خاطرنشان کرد: "بیایید به یاد داشته باشیم، ۱۹۸۴ تنها هجده سال دیگر است. گمان خودم این است که برادر بزرگ، اگر هرگز به ایالات متحده بیاید، ممکن است نه یک قدرت‌طلب حریص، بلکه یک بوروکرات بی‌رحم و وسواسی نسبت به کارایی باشد. و او، بیش از یک قدرت‌طلب ساده، می‌تواند ما را به آن نهایت وحشت‌ها، یعنی بشریتی در زنجیره‌های نوار پلاستیکی، سوق دهد".

پوشش خبری رسانه‌ها از جلسه و پیشنهاد مرکز داده ملی نیز این احساسات را بازتاب داد. روزنامه "چارلستون پست" نگران بود که سوابق دولتی به شدت تقویت شده منجر به پرونده‌هایی مانند پرونده‌های جمع‌آوری شده توسط "گشتاپوی هیتلر" شود، اما بدتر. "بنتون هاربر نیوز-پالادیوم" پیشنهاد کرد که مقامات منتخب ممکن است از آن برای اجبار به کمک‌های مالی کمپین از شهروندان یا باج‌گیری از مخالفان سیاسی استفاده کنند. یک سرمقاله "نیویورک تایمز" با آه و افسوس پرسید: "آیا حریم خصوصی شخصی می‌تواند در برابر پیشرفت‌های بی‌وقفه غول فناوری زنده بماند؟" و سپس به "کابوس اورولی" اشاره کرد که در صورت اجازه کنگره به آن بیست آژانس فدرال برای سازماندهی کارآمدتر داده‌هایشان، بر آمریکا نازل خواهد شد. برای چندین سال دیگر، اعضای کنگره همچنان از ایده مرکز داده ملی به عنوان یک کیسه بوکس ایدئولوژیک

در جلسات و گزارش‌ها استفاده کردند. سپس، در سال ۱۹۶۹، این ابتکار به دلیل تغییرات بوروکراتیک که با پایان دوره لیندون جانسون و انتقال نظارت بر آینده داده‌های آمریکا به ریچارد نیکسون آغاز شد، سرانجام به پایان رسید.

از کامپیوتر شخصی تا عصر دیجیتال: دگرگونی هویت و عامل‌بودگی فردی

امروزه، ونس پاکارد تا حد زیادی فراموش شده است. مرکز داده ملی و فریادهای اعتراضی که از سوی کنگره، نویسندگان مقالات و عموم مردم آمریکا الهام گرفت، تقریباً ناشناخته‌اند. بخشی از آن به این دلیل است که حتی در کوتاه مدت، شکست مرکز داده ملی تأثیر چندانی بر کند کردن سرعت نوآوری فناورانه نداشت. ابتدا، دیسک‌های سخت جایگزین نوارهای مغناطیسی کند و با ظرفیت محدود مین‌فریم‌های میانه قرن شدند. سپس سیستم‌های اشتراک‌گذاری زمان، مینی‌کامپیوترها، شبکه‌های محلی، انقلاب کامپیوتر شخصی، دوربین‌های ویدئویی، شبکه جهانی وب، رسانه‌های اجتماعی، فضای ابری، تلفن‌های همراه و ردیاب‌های تناسب اندام پدید آمدند. و همچنین حسگرهای حرکت، سیستم‌های تشخیص چهره، اسکنرهای اثر انگشت، اسکنرهای عنبیه، ترموستات‌های هوشمند، ردیاب‌های خواب، ایمپلنت‌هایRFID ، تتوهای دیجیتالی و دلالان داده که پروفایل‌های دیجیتالی شما را با هزاران نقطه داده درباره جمعیت‌شناسی پایه، وضعیت‌های پزشکی، وضعیت مالی، ترجیحات جنسی، گرایش‌های سیاسی و حتی پیش‌بینی‌هایی درباره رفتار آینده شما – همه بدون اطلاع یا رضایت مستقیم شما – جمع‌آوری می‌کنند.

از بسیاری جهات، نگرانی‌هایی که بدبینان میانه قرن درباره عامل‌بودگی فردی، حریم خصوصی و چشم همه‌بین فناوری ابراز می‌کردند، همچنان معتبرند. در واقع، آسان است تصور کنیم پاکارد مقالات بسیار نگران‌کننده‌ای را در ساب‌استک خود درباره ظرفیت مدل‌های زبانی بزرگ برای دستکاری رفتار عمومی از طریق تولید انبوه محتوای متقاعدکننده که برای بهره‌برداری از پروفایل‌های روانشناختی فردی طراحی شده است، منتشر می‌کند. یا کورنلیوس گالاگر که از سال

۱۹۶۶ از طریـق کرم‌چالـه‌ای در روتونـدای کنگـره بـه زمـان حـال سـفر کـرده و سـوندار پیچـای، مدیرعامل گوگل، را درباره تصمیم‌گیری الگوریتمی مورد بازجویی قرار می‌دهد.

امـا همچنـین لازم بـه ذکـر اسـت کـه پاکـارد و یـارانش تـا چـه حـد دربـاره آینـده در اشـتباه بودنـد. هنگـامی کـه پاکـارد بـه کنگـره دربـاره جهـانی هشـدار داد کـه پیشـرفت‌های فناورانـه "بشـریت را در زنجیـر" رهـا خواهـد کـرد، بسـیاری از ایالت‌هـا هنـوز قوانینـی بـرای محـدود کـردن ازدواج بـین نـژادی داشـتند. ازدواج همجنس‌گرایـان آنقـدر غیرقابـل تصـور بـود کـه هـیچ ایـالتی حتـی بـه فکـر تصـویب قـوانینی بـرای ممنوعیـت صـریح آن نیفتـاده بـود. پرونـده "گریزولـد علیـه کنتیکـت" حـق حـریم خصوصـی را بـرای همـه کسـانی کـه ممکـن بـود بخواهنـد از وسـایل پیشـگیری از بـارداری اسـتفاده کننـد، ایجـاد نکـرد. نظـر اکثریـت قاضـی ویلیـام او. داگلاس بـه طـور مکـرر اشـاره می‌کنـد کـه حکـم دادگاه بـه طـور خـاص "حـق حریـم خصوصـی زناشـویی" را تنهـا محافظـت می‌کنـد. در یـک نظـر موافـق، قاضـی آرتـور گلدبـرگ حتـی روشـن‌تر بیـان کـرد کـه تصمیـم دادگاه قصـد نـدارد "مقـررات مناسـب دولـت در مـورد بی‌بنـدوباری یـا سـوءرفتار جنسـی" را نقـض کنـد یـا قـدرت آن را بـرای "ممنـوع کـردن روابـط جنسـی خـارج از ازدواج... یـا گفتـن اینکـه چـه کسـی می‌توانـد ازدواج کنـد" سـلب کنـد. بنـابراین در حالـی کـه مخالفـان مرکـز داده ملـی معتقـد بودنـد کـه موضعـی اصولـی بـرای آزادی و رهایـی آمریکایـی می‌گیرنـد، بسـیاری از هموطنـان آن‌هـا بـه شـدت درگیـر زنجیرهـای قانونـی از نوعـی بودنـد.

امـا ایـن نیـز درسـت اسـت کـه هنجارهـای فرهنگـی در حـال تغییـر بودنـد. و بـا گسـترش سیسـتم‌های کامپیوتـری، مـا بـه دنیـایی بسـیار آزادتـر از آنچـه قبـل از آن بـود، دسـت یافتیـم. ایـن دنیـای جدیـد، کاوش خلاقانـه، شـورش علیـه همرنگـی، سـبک‌های زنـدگی و ارزش‌هـای جایگزیـن را تلویحـاً تشـویق می‌کـرد. ایـن امـر عامـل‌بودگی و خودمختـاری فـردی را بـه گروه‌هـای سـنتی بـه حاشـیه رانـده شـده، از جملـه زنـان، افـراد رنگیـن‌پوسـت و جامعـه LGBTQ گسـترش داد. عوامـل بسـیاری ایـن تغییـرات را بـه پیـش راندنـد – بـه ویـژه سـازماندهی و فعالیت‌هـای مردمـی کـه منجـر بـه قوانیـن و مقـررات جدیـد شـد. امـا نوآوری فناورانـه نیـز نقـش کلیـدی ایفـا کـرد. بـا وجـود تمـام تبلیغـات مجـلات بـرای میکروفون‌هـای مخفـی و دوربیـن‌هـای جاسوسـی مـادون قرمـز، اواخـر قـرن بیسـتم بیشـتر توسـط دسـتگاه‌های کمـی

متفاوت شکل گرفت: کنترل از راه دور، تلویزیون کابلی، دستگاه‌های ویدئویی، واکمن‌های سونی و دوربین‌های فیلمبرداری دستی. هر کدام از این‌ها، همرنگی دوران رسانه‌های جمعی را از بین بردند و آینده‌ای شخصی‌تر، قابل تنظیم‌تر و خودشکوفاترتری را به ارمغان آوردند.

قابل توجه‌ترین نکته این است که کامپیوتر مین‌فریم، آن دستگاه ترسناک توتالیتاریسم خزنده، تا حد زیادی توسط ماشین جدیدی جایگزین شد که ممکن بود از نظر ونس پاکارد غیرقابل قبول و متناقض به نظر برسد: کامپیوتر شخصی. یک آگهی اولیه اپل می‌پرسید: "چه نوع مردی کامپیوتر خودش را دارد؟" و با تصویری از نماد طبقه خلاق و نمونه‌ای از خودسازی و خودمختاری فردی، بنجامین فرانکلین، پاسخ می‌داد. آگهی ادامه می‌داد: "داشتن کامپیوتر شخصی، انقلابی به نظر می‌رسد؟ نه، اگر شما یک دیپلمات، چاپگر، دانشمند، مخترع... یا یک طراح بادبادک باشید. امروز کامپیوتر اپل وجود دارد. این کامپیوتر برای شخصی بودن طراحی شده است. برای ساده‌تر کردن زندگی شما. و کارآمدتر کردن شما".

اما فقط این نبود که ابزارهای جدیدی مانند کامپیوتر به افراد قابلیت‌های جدید و سطوح جدیدی از عامل‌بودگی در زندگی شخصی و کاری‌شان می‌داد. در حالی که پاکارد در اشاره به اینکه چگونه شرکت‌ها، همراه با دولت، شروع به جمع‌آوری اطلاعات بیشتری درباره افراد کرده بودند، پیشگام بود، این عمل گزینه‌ها و جایگزین‌های موجود برای مردم را به طور چشمگیری محدود نکرد، آنگونه که پاکارد و دیگران می‌ترسیدند. در عوض، ماشین‌های حافظه غول‌پیکری که پاکارد هشدار می‌داد نسل‌های آینده آمریکایی‌ها را سرکوب خواهند کرد، به آزادی آن‌ها کمک کردند. ظرفیت جدید برای تجمیع و تحلیل مقادیر عظیمی از داده‌های مصرف‌کننده، به تولیدکنندگان، بازاریابان و سازندگان محتوا این قدرت را داد که بازارها را تقسیم‌بندی کنند، جمعیت‌شناسی‌های دقیق‌تر را هدف قرار دهند و به خرده‌فرهنگ‌ها، سبک‌های زندگی و علایق خاص به گونه‌ای که در دوران پیش از کامپیوتر هرگز نمی‌توانست اتفاق بیفتد، پاسخ دهند. شخصی‌سازی در چندین حوزه، جایگزین رویکرد استاندارد "یک اندازه برای همه" صنعت‌گرایی قرن بیستم شد. این آغاز یک جهان متنوع‌تر و فراگیرتر مبتنی بر دانش جامع بود. برادر بزرگ، یا شاید مناسب‌تر، کسب‌وکار

بزرگ، نه آنقدر شما را تماشا می‌کرد که باعث می‌شد "احساس دیده شدن" کنید. با آگاهی بیشتر از آنچه مردم می‌خریدند، اطلاعاتی که مصرف می‌کردند و به طور کلی چگونه زندگی می‌کردند، شرکت‌های تجاری می‌توانستند به طور مؤثرتری به آن‌ها خدمت کنند.

شبکه‌های اعتماد: معماری جدید تعاملات انسانی

در دنیای بسیار دیجیتالی و شبکه‌ای امروز، آنچه زمانی مبهم بود، به راحتی قابل تحلیل می‌شود — در این مورد شکی نیست. اما چه چیزی نقض حریم خصوصی است و چه چیزی تأیید هویت؟ چه زمانی تولید و انتشار دانش جدید، عامل‌بودگی فردی را تقویت می‌کند و چه زمانی آن را کاهش می‌دهد؟ تا چه حد وفاداری به حریم خصوصی، به پنهان نگه داشتن اطلاعات، فرصت‌های ما را برای به اشتراک گذاشتن دانش، یادگیری چیزهای جدید و رشد کاهش می‌دهد؟ لوئیز براندیس، قاضی آینده دیوان عالی، در مقاله‌ای در سال ۱۸۹۰ نوشت: "هر کس 'حق تنها ماندن' را دارد" که به شکل‌گیری مفاهیم حریم خصوصی در قرن بیستم کمک کرد. اما به عنوان عضوی از جامعه، و به ویژه به عنوان عضوی از یک جامعه شبکه‌ای، این همیشه سازنده‌ترین حق برای سازماندهی زندگی نیست. در دهه ۱۹۶۰، زمانی که "۱۹۸۴" هنوز سایه بلندی افکنده بود و سرعت سرسام‌آور تغییرات گیج‌کننده بود، به نظر غیرممکن می‌آمد که جهانی را تصور کنیم که در آن دستگاه‌های قادر به ضبط، تحلیل و به اشتراک گذاشتن اطلاعات بتوانند عامل‌بودگی انسان را افزایش دهند، نه اینکه آن را کاهش دهند. برای برخی، هنوز هم همینطور است. امروز در پلتفرم‌های رسانه‌های اجتماعی، لازم نیست زیاد جستجو کنید تا منتقدان مترقی حریم خصوصی یا نظریه‌پردازان راست‌گرا را بیابید که از افراط‌گری اورولی فناوری بزرگ آنلاین و تأثیر خفقان‌آور آن بر حریم خصوصی، خودمختاری شخصی و فرهنگ قرن بیست و یکم به طور کلی گلایه می‌کنند. اما واقعاً، امروز چه کسی گزینه‌های کمتری نسبت به دهه ۱۹۶۰، ۱۹۸۰ یا ۲۰۰۰ برای بیان آنچه هستند، اولویت‌بندی آنچه برایشان ارزشمند است و دنبال کردن زندگی مورد نظر خود دارد؟

در بسیاری جهات، برادر بزرگ برای چندین دهه مرده است. اگر این اغراق‌آمیز به نظر می‌رسد، فقط تصور کنید وینستون اسمیت، قهرمان رمان ۱۹۸۴، چه شوکی را تجربه می‌کرد اگر او را در

خیابان‌های منهتن امروزی رها می‌کردید. تنها ویترین‌های رنگارنگ فروشگاه‌ها ممکن بود او را به حالت گیجی غیرقابل درکی بیندازد. سپس نمایش‌های خیره‌کننده خود، نمایش‌های جسورانه از صمیمیت‌ها و شور و شوق‌های شخصی که هیچ ارتباطی با حزب حاکم ندارند را اضافه کنید. بیلبوردهای غول‌پیکر پر از تصاویری از افرادی که برادر بزرگ نیستند. گوشی‌های هوشمندی که مردم مانند طلسم‌های عزیز به قلب خود می‌چسبانند. در ایالات متحده امروز، محیط کنونی ما آنقدر به سمت ابراز وجود، تنوع و عدم همرنگی گرایش دارد که بخش بزرگی از مردم ظاهراً مشتاق احیای یک فرد قدرتمند شبیه برادر بزرگ هستند که بتواند وفاداری مطلق را الهام بخشد و انسجام اجتماعی را به جمهوری چندصدایی و بسیار تجزیه شده ما بازگرداند.

با این حال، به دلیل هوش مصنوعی و اشتهای آن برای داده‌ها، بسیاری از منتقدان امروزی بیش از همیشه مطمئن هستند که "زنجیرهای نوار پلاستیکی" که پاکارد شصت سال پیش پیشگویی کرد، ممکن است سرانجام ما را به دام اندازند. البته، این احتمالی نیست که باید به سادگی از آن چشم‌پوشی کرد. عملکرد گذشته هیچ تضمینی برای عملکرد آینده نیست، به ویژه در صنعت تحول‌آفرینی. اما همانطور که نباید لزوماً پیش‌بینی‌های شکست‌خورده گذشته را نادیده بگیریم، نباید درس‌های واقعی آن را نیز نادیده بگیریم. و آنچه تاریخ در شصت سال گذشته، بارها و بارها نشان داده است، این است که اطلاعات، به طور گسترده به اشتراک گذاشته شده، جمع‌آوری شده، تحلیل شده و به روش‌های مختلف به کار گرفته شده، می‌تواند فوق‌العاده توانمندساز و الهام‌بخش باشد — نه آنقدر زنجیرهای نوار پلاستیکی، بلکه طناب‌های کوهنوردی که برای صعود به ارتفاعات بیشتر و دستیابی به سطوح جدیدی از معنا و رضایت استفاده می‌کنید.

این به معنای بی‌ارزش بودن حریم خصوصی نیست. به ویژه در سیصد سال اخیر، پس از آنکه شهرنشینی واقعاً آغاز شد و تعداد زیادی از مردم در نزدیکی غریبه‌ها شروع به زندگی کردند، ما حریم خصوصی را به طور فزاینده‌ای مقدس می‌دانیم. خارج از فشارهای، پرهیزگاری‌ها و اغلب هنجارهای محدودکننده زندگی عمومی، حریم خصوصی فضایی برای تأمل صادقانه و خودارزیابی

فراهم می‌کند. ما را قادر می‌سازد تا خود واقعی‌مان را کشف و تعریف کنیم. این به نوبه خود به ما کمک می‌کند تا با خودمختاری بیشتر و ظرفیت بیشتری برای عامل‌بودگی عمل کنیم. اما حریم خصوصی تنها راه دستیابی به این اهداف نیست. به ویژه در یک دنیای شبکه‌ای، یک هویت عمومی قوی نیز خودمختاری و عامل‌بودگی ایجاد می‌کند. حتی اگر گاهی از افشای دائمی زندگی دیجیتال از طریق نام‌های مستعار، VPNها و رمزگذاری به پناه می‌بریم، ما همچنین به تأیید، پذیرش و حاکمیت شخصی که هویت عمومی تسهیل می‌کند، نیاز داریم. به همین دلیل، یکی از آرمان‌های برجسته عصر دیجیتال – شاید برجسته‌ترین آرمان – "دیده شدن" است.

هویت عمومی به معنای قابلیت کشف، قابل اعتماد بودن، نفوذ، قدرت و عامل‌بودگی است. این نوعی سرمایه اجتماعی – و گاهی سرمایه مالی – است که می‌تواند به شما کمک کند تا با بهره‌وری بیشتری در جهان حرکت کنید. این چیزی است که مخالفان پروژه مرکز داده ملی شصت سال پیش نتوانستند ببینند. به همین دلیل است که بدبینان معاصر سنت‌های مقدس قبیله خود را حفظ می‌کنند و اصرار دارند که، طبق معمول، ما در آستانه یک دیستوپیای اورولی هستیم، حتی در حالی که نوآوری فناورانه همچنان به ما ابرقدرت‌های جدیدی می‌بخشد که ما را قادر می‌سازند زندگی معنادارتر و رضایت‌بخش‌تری داشته باشیم.

چشم‌انداز بنیان‌گذاری هوش مصنوعی به عنوان "امتداد اراده‌های فردی انسان" برای بسیاری الهام‌بخش بوده است. همین تمایل به ایجاد راه‌های جدید برای دستیابی افراد به عامل‌بودگی بیشتر در زندگی‌شان – قدرت انتخاب‌های آگاهانه و عمل مؤثر بر آن‌ها – همان چیزی است که در ابتدا الهام‌بخش پلتفرم‌هایی مانند لینکدین شد. همانطور که در سال ۲۰۰۳، سال راه‌اندازی، در وب‌سایت توضیح داده شد: "شما از طریق ارتباطات کاری که از قبل دارید، با افرادی که برای دستیابی به اهداف کاری خود نیاز دارید، در ارتباط هستید. با هزاران متخصص، از طریق ارتباطات شبکه‌ای مورد اعتماد، تماس بگیرید و به خود و همکارانتان کمک کنید تا کارها را انجام دهید".

با این حال، در سال‌های اولیه، لینکدین اغلب به سادگی به عنوان "سایت رزومه" یا "سایت

آگهی‌های شغلی" توصیف می‌شد. در واقع، پس از بیش از بیست سال، هنوز برخی افراد آن را اینگونه می‌پندارند. اما از ابتدا، لینکدین همیشه به عنوان چیزی اساسی‌تر از آن تصور می‌شد. لینکدین همیشه در مورد استفاده از شبکه‌ها برای به اشتراک گذاشتن و کشف اطلاعات به روش‌های جدید، با استفاده از هویت برای افزایش اعتماد بود. ایجاد یک پروفایل عمومی قابل دسترسی که شامل اطلاعات اساسی درباره شغل و آرزوهای شغلی شما باشد، اکنون ممکن است بدیهی به نظر برسد، اما در اوایل دهه ۲۰۰۰ هنوز رفتار پیش‌فرض نبود. در واقع، زمانی که اینترنت با معرفی شبکه جهانی وب یک دهه قبل شروع به عمومی شدن کرده بود، گمنامی به عنوان یکی از بزرگترین فضیلت‌های آن تلقی می‌شد. همانطور که یک کارتون کلاسیک نیویورکر در سال ۱۹۹۳ می‌گفت: "در اینترنت، هیچ‌کس نمی‌داند شما یک سگ هستید".

اما اگر می‌خواستید مردم بدانند شما یک سگ هستید چه؟ یا به عبارت دقیق‌تر، اگر می‌خواستید مردم بدانند شما یک مهندس پایگاه داده با تسلط عمیق بر MySQL و SQL Server هستید؟ و به راحتی به این عنوان قابل کشف باشید؟ و حتی بهتر از آن، به این عنوان به راحتی قابل اعتماد باشید؟ قبل از اینترنت، تقریباً فقط افراد مشهور و مدیرانی که به اندازه کافی برجسته بودند تا در مجلاتی مانند "فورچون" یا بخش کسب و کار "نیویورک تایمز" پوشش داده شوند، یک هویت حرفه‌ای داشتند که فراتر از همکاران و دایره محدودی از افرادی که مستقیماً با آن‌ها تجارت می‌کردند، گسترش می‌یافت. با ظهور وب، هر کسی می‌توانست برای خود یک هویت حرفه‌ای ایجاد و تبلیغ کند. این یک تغییر عمده با مزایای بالقوه بسیاری بود، اما مسائل مربوط به اعتماد را نیز ایجاد کرد. بدون "فورچون" یا "نیویورک تایمز" برای تأیید ادعاهای شما درباره خودتان، چگونه می‌توانستید به مردم اطمینان دهید که واقعاً یک مهندس پایگاه داده با تسلط عمیق بر MySQL هستید، و نه فقط یک سگ که وانمود می‌کند مهندس است؟ قرار دادن هویت فردی در یک شبکه بزرگتر از ارتباطات حرفه‌ای، یکی از راه‌های انجام این کار است. شما ممکن است مارک را نشناسید، اما اگر لیندا را بشناسید و بدانید لیندا مارک را می‌شناسد، مارک فوراً کمی قابل اعتمادتر به نظر می‌رسد.

در اصل، آنچه لینکدین و بسیاری دیگر از پلتفرم‌های موفق اینترنتی انجام می‌دهند، مقیاس‌بندی اعتماد است. به این فکر کنید که چگونه eBay، PayPal، Airbnb، Uber و Lyft، تنها به چند مورد اشاره کنیم، از مکانیزم‌های نوآورانه اعتماد مختلفی برای امکان‌پذیر ساختن طیف وسیعی از تعاملات، تراکنش‌ها و رفتارها در مقیاس جهانی استفاده کردند. در حالی که اینترنت قطعاً فرصت‌های جدیدی برای کلاهبرداری و اطلاعات نادرست ایجاد کرده است، داستان بزرگتر آن این است که چگونه به عنوان یک ماشین اعتماد بی‌سابقه عمل کرده است. به عنوان مثال، در سال ۱۹۹۵، ما با استفاده از نام‌های مستعار در اینترنت گشت و گذار می‌کردیم. تا سال ۱۹۹۹، ما ماشین‌های دست دوم را در eBay Motors بدون دیدن خریداری می‌کردیم. تا سال ۲۰۱۲، پس از یک شب تفریح در شهر، وارد یک تویوتا کرولای تصادفی با سبیل صورتی روی جلوپنجره آن می‌شدیم تا به سلامت به خانه برسیم. وقتی به آن فکر می‌کنید، واقعاً شگفت‌انگیز است. در مورد لینکدین، با افزایش پایگاه کاربری آن و بزرگتر و متصل‌تر شدن شبکه‌های فردی افراد، پلتفرم به طور فزاینده‌ای قابل اعتمادتر شد. این به این دلیل است که هر کاربر به طور مؤثر بسیاری از کاربران دیگر را تأیید می‌کرد، در نوع جدیدی از پلتفرم اعتماد توزیع شده که ریشه در هویت فردی واقعی داشت.

برای موفقیت این نوع جدید پلتفرم اعتماد، ما باید مردم را متقاعد می‌کردیم که اطلاعات بیشتری درباره خودشان به اشتراک بگذارند تا آنچه قبلاً به صورت آنلاین به اشتراک می‌گذاشتند. و فقط این نبود که از کاربرانمان خواستیم پروفایل‌هایی برای خود با نام‌های واقعی خود ایجاد کنند، که این خود در آن زمان هنوز یک عمل نسبتاً غیرمعمول بود. ما همچنین آن‌ها را تشویق کردیم تا عناوین شغلی، شرکت‌هایی که در آن‌ها کار می‌کردند، مهارت‌ها و تجربیات خود و شبکه‌های خود را به اشتراک بگذارند. البته، ما به کاربران کنترل‌های حریم خصوصی صریحی دادیم که آن‌ها را قادر می‌ساخت تا مدیریت کنند چه کسی می‌تواند اطلاعات آن‌ها را ببیند و چگونه به اشتراک گذاشته می‌شود. اما به طور کلی، ایده این بود که حداقل تا حدی خود را برای سایر اعضای لینکدین فراتر از ارتباطات مستقیم خود قابل مشاهده کنید، تا ارتباطات جدید را

تسهیل کنید. در سال ۲۰۰۳، این قطعاً کاری معمول نبود. در واقع، بسیاری از افراد و شرکت‌ها آن را کاملاً بحث‌برانگیز یافتند.

قبل از اینترنت، یکی از دلایلی که مردم عموماً اطلاعات دقیق درباره شغل خود را به طور گسترده به اشتراک نمی‌گذاشتند، صرفاً این بود که راه‌های کمی برای انجام این کار وجود داشت. شما زمانی درباره تجربه و تخصص خود به مردم می‌گفتید که فعالانه به دنبال شغل جدید بودید. یا شاید زمانی که در یک بار فرودگاه در حال گپ زدن با یک غریبه بودید در حالی که منتظر شروع پروازتان بودید. رزومه‌ها مستقیماً به بخش‌های منابع انسانی شرکت‌ها، استخدام‌کنندگان حرفه‌ای و آژانس‌های کاریابی ارسال می‌شد. در دوران کامپیوترهای رومیزی دهه‌های ۱۹۸۰ و اوایل ۱۹۹۰، برخی افراد از برنامه‌هایی مانند ACT! Contact Manager و Sidekick برای پیگیری شبکه‌های حرفه‌ای خود استفاده می‌کردند – اما این‌ها شبکه‌هایی به معنایی که امروز این اصطلاح را می‌فهمیم، نبودند. در عوض، شما یک پایگاه داده بر روی کامپیوتر خود از نام افراد، شماره تلفن، آدرس ایمیل، کارفرما و هر اطلاعات دیگری که درباره آن‌ها به دست آورده بودید، نگهداری می‌کردید. هنگامی که اطلاعات آن‌ها تغییر می‌کرد، باید آن را به صورت دستی در فایل خود به‌روزرسانی می‌کردید – مشروط بر اینکه می‌دانستید تغییر رخ داده است. در اوایل دهه ۲۰۰۰، برخی از توسعه‌دهندگان برنامه‌های مدیریت تماس مبتنی بر وب را ایجاد کرده بودند. اما این سیستم‌ها به سادگی رویکردی را که برنامه‌های رومیزی استفاده کرده بودند، تکرار می‌کردند. بنابراین، در حالی که کاربران اکنون می‌توانستند به لیست تماس‌های خود از هر کامپیوتری با اتصال اینترنت دسترسی پیدا کنند، همه این لیست‌ها عملاً هنوز جزایر کوچکی از داده‌های خود بودند، که از یکدیگر جدا شده بودند. این رویکرد، بدون قصد واقعی، حریم خصوصی را به حداکثر می‌رساند – اما همچنین نتوانست از تمام قابلیت‌های ارزشمند جدیدی که یک شبکه واقعی می‌توانست ارائه دهد، بهره ببرد.

این همان کاری بود که لینکدین انجام داد. در پلتفرم ما، کاربران با یکدیگر ارتباط برقرار می‌کردند و پروفایل‌های خود را ایجاد و نگهداری می‌کردند. شما دیگر مجبور نبودید به طور

دوره‌ای با رئیس قدیمی خود از شغل قبلی‌تان تماس بگیرید تا ببینید اطلاعات تماس شما برای او هنوز به‌روز است یا خیر، و اگر تغییری کرده بود، سابقه او را در پایگاه داده خود به‌روز کنید. در عوض، او پروفایل لینکدین خود را هر زمان که تغییری رخ می‌داد، به‌روز می‌کرد و هر کسی که با او در ارتباط بود، فوراً به جدیدترین اطلاعات او دسترسی پیدا می‌کرد. حل مشکل به‌روزرسانی، تنها آغاز مزایای اتصال اطلاعات به این شیوه مشترک و عمومی‌تر بود. اول از همه، این واقعیت بود که شما اکنون مسئول هویت شخصی خود بودید. به جای اینکه کاربران دیگر یک ورودی برای شما در لیست تماس خود ایجاد کنند و آنچه را که درباره شما می‌دانستند در آن قرار دهند، شما اطلاعاتی را ارائه می‌دادید که می‌خواستید آن‌ها بدانند. شما مسئول تعریف خودتان بودید. هنگامی که با شخص دیگری ارتباط برقرار می‌کردید، اکنون می‌توانستید مخاطبین او را نیز ببینید، چیزی که هرگز در دوران برنامه‌های تماس رومیزی امکان‌پذیر نبود. اکنون ناگهان یافتن مخاطبین جدید و قابل کشف شدن برای دیگران بسیار آسان‌تر شد.

چالشی که با آن روبرو بودیم این بود که این‌ها همه رفتارهای جدیدی بودند. هنگامی که لینکدین راه‌اندازی شد، زندگی واقعی و "فضای مجازی" به روش‌های بیشتری در حال همگرایی بودند، اما گمنامی یا نام مستعار هنوز حالت‌های غالب برای تعامل عمومی در اینترنت بودند. و مردم به ویژه درباره به اشتراک گذاشتن جزئیات زندگی حرفه‌ای خود محتاط بودند. در واقع، در روزهای اولیه، حتی با شروع محبوبیت شبکه‌های اجتماعی، اکثر مردم احساس می‌کردند که انگیزه‌های روشنی برای عدم ایجاد پروفایل در لینکدین دارند، دقیقاً به این دلیل که ما از اعضای خود خواستیم از نام‌های واقعی خود استفاده کنند. برخی افراد نمی‌خواستند رؤسایشان فکر کنند که آن‌ها ناراضی هستند و قصد ترک شغل را دارند. دیگران فایده‌ای در به اشتراک گذاشتن عمومی اطلاعات واقعی درباره خودشان نمی‌دیدند اگر در حال حاضر به دنبال شغل نبودند یا به دنبال پاداش کوتاه‌مدت دیگری نبودند. در نهایت، اگر شما از آن دسته افرادی بودید که فعالانه یک شبکه حرفه‌ای از مخاطبین را جمع‌آوری می‌کردید – که در آن زمان به این معنی بود که احتمال زیادی وجود داشت که در فروش کار کنید – شبکه شما یکی از مزایای رقابتی بزرگ شما

بود. پس چرا آن را با رقبای احتمالی به اشتراک بگذارید؟ چرا خود را در معرض جریان جدیدی از تماس‌گیرندگان دیجیتالی قرار دهید که برای معرفی آسان به افرادی که می‌شناختید، مزاحمت ایجاد می‌کردند؟ کارفرمایان نیز لزوماً از این پلتفرم راضی نبودند. زیرا حتی اگر کارمندانشان صرفاً می‌خواستند به گونه‌ای شبکه‌سازی کنند که به آن‌ها در انجام بهتر کارهای فعلی‌شان کمک کند، چرا به آن‌ها اجازه دهند که اینقدر برای استخدام‌کنندگان و رقبا قابل مشاهده باشند؟ و نه تنها قابل مشاهده، بلکه آزاد باشند تا شغل خود را آنگونه که خودشان می‌دیدند، توصیف کنند، که شاید دقیقاً با دیدگاه کارفرمایانشان مطابقت نداشت. خوشبختانه، به اندازه کافی از کاربران اولیه احساس کردند که مزایای بالقوه فراوانی که پلتفرم ما می‌توانست در نهایت ارائه دهد، مشارکت را به یک ریسک هوشمندانه و ارزشمند تبدیل می‌کند. با به اشتراک گذاشتن اطلاعات درباره خودشان، آن‌ها قادر خواهند بود سودمندی و دامنه شبکه‌های حرفه‌ای خود را افزایش دهند. این به نوبه خود می‌تواند طیف وسیعی از مزایا را در طول زمان فراهم کند: فرصت‌های جدید برای توسعه هویت حرفه‌ای، منابع جدید اطلاعات صنعتی و شرکای بالقوه جدید برای همکاری به روش‌های مختلف.

پایان برادر بزرگ و آغاز چالش‌های جدید

اگر چیزی شبیه لینکدین در دهه ۱۹۶۰ وجود داشت، ونس پاکارد احتمالاً آن را نگران‌کننده می‌یافت. این پلتفرم، پس از همه، نه تنها یک مرکز داده ملی، بلکه یک مرکز داده جهانی است. حجم عظیمی از اطلاعات را در یک مخزن واحد جمع‌آوری می‌کند و جستجو در آن را آسان می‌سازد. بسیاری از این اطلاعات بی‌شک از نظر پاکارد اساساً "خصوصی" و بنابراین مشکل‌ساز تلقی می‌شد، حتی اگر به صورت داوطلبانه منتشر شده بود. (بسیاری از داده‌هایی که دولت و بازیگران بخش خصوصی در دهه ۱۹۶۰ جمع‌آوری می‌کردند نیز به صورت داوطلبانه افشا شده بودند.)

با این حال، این تمایل به زندگی عمومی‌تر، مقدار زیادی ارزش جمعی ایجاد می‌کند. با توجه به اینکه بسیاری از افراد دیگر جزئیات سوابق کاری و آرزوهای خود را منتشر می‌کنند، شما می‌توانید

بیاموزید که چه نوع مهارت‌هایی برای افرادی که شغلی را که دوست دارید، در شرکتی که دوست دارید کار کنید، مهم هستند. می‌توانید ببینید کدام شرکت‌ها در بین سایر متقاضیان محبوب‌تر هستند. هنگامی که اطلاعاتی که قبلاً سیلویی شده بود آسان‌تر در دسترس قرار می‌گیرد، همه شرکت‌کنندگان در پلتفرم، از جمله افراد و شرکت‌ها به طور یکسان، می‌توانند با آگاهی بیشتری عمل کنند. برای مثال، هر دقیقه هفت نفر از طریق لینکدین استخدام می‌شوند، نرخی که معادل ۳٫۶۷ میلیون نفر در سال است.

این نیز درست است که زندگی‌های عمومی‌تری که ما اکنون داریم، همیشه به روش‌های مفید پیش نمی‌رود. یک تبادل غیررسمی و ظاهراً زودگذر در فیس‌بوک یا X.com ممکن است اکنون برای همیشه عمومی باقی بماند. یک اظهار نظر بی‌مقدمه که توسط تماشاگران دیجیتال ضبط یا بازنشر شده است، می‌تواند بسیار عمومی‌تر از آنچه نویسنده آن قصد داشته است، شود. در این واقعیت جدید، منطقی است که پناهگاهی که حریم خصوصی فراهم می‌کند، برای ما ارزشمندتر می‌شود. با این حال، جوامع – و به ویژه دموکراسی‌ها – همیشه به جریان آزاد اطلاعات به همان اندازه که به حریم خصوصی وابسته بوده‌اند، وابسته بوده‌اند. آندرئا بلیگر و دیوید جی. کرایگر، نظریه‌پردازان اینترنت، در کتاب خود "حکمرانی عمومی شبکه‌ای" در سال ۲۰۱۸ می‌نویسند: "به نظر می‌رسد ما در حال ورود به قرن ۲۱ با جامعه‌ای هستیم که به دو دسته تقسیم شده است: کسانی که معتقدند هر چه بیشتر اطلاعات باید در همه حوزه‌ها و در همه سطوح یکپارچه شود و کسانی که معتقدند آزادی، خودمختاری و حتی کرامت انسانی به رازداری و پنهان نگه داشتن هر چه بیشتر اطلاعات بستگی دارد".

چگونه می‌توانیم تعادل صحیح بین این دو دیدگاه را برقرار کنیم؟ در مواجهه با نوآوری‌های محاسباتی بزرگ در دهه ۱۹۶۰، قانونگذاران و مفسران آمریکایی تقریباً منحصراً بر آنچه می‌توانست اشتباه پیش برود، حداقل در مورد مرکز داده ملی، تمرکز کردند. در تلاش‌های موفق خود برای ممنوع کردن آن پروژه، آن‌ها همچنین هر گونه منافعی را که ممکن بود از دسترس‌پذیرتر کردن ذخایر رو به رشد داده‌های دولت برای محققان و سیاست‌گذاران، زودتر از موعد، حاصل شود،

ممنوع کردند.

از اطلاعات انباشته تا دانش استخراجی: نقش هوش مصنوعی

چرا عجله کنیم؟ هرچه بیشتر هوش مصنوعی را بپذیریم، به طور مؤثرتری می‌توانیم از داده‌ها و اطلاعاتی که در حال حاضر ایجاد می‌کنیم به گونه‌ای بهره ببریم که عامل‌بودگی ما را در تقریباً هر کاری که به عنوان انسان انجام می‌دهیم، افزایش دهد. این به این دلیل است که بشریت از قبل به نقطه‌ای رسیده است که اطلاعات بیشتری تولید می‌کند که می‌تواند به تنهایی از آن استفاده کند. در مقاله‌ای در سال ۱۹۹۸، هال واریان، اقتصاددان و استاد دانشکده اطلاعات دانشگاه کالیفرنیا، برکلی که بعدها اقتصاددان ارشد گوگل شد، نوشت که میزان اطلاعاتی که جهان تولید می‌کند و در واقع مصرف می‌شود، "به سمت صفر میل می‌کند".

برای حمایت از این ادعا، واریان به کتاب "جریان‌های ارتباطات: سرشماری در ایالات متحده و ژاپن" که در سال ۱۹۸۴ توسط ایتل دسولا پول، استاد قدیمی مؤسسه فناوری ماساچوست، و همکارانش نوشته شده بود، اشاره کرد. در آن، پول و همکارانش تلاش کردند تا داده‌های واقعی پشت "انفجار اطلاعات" را کمی‌سازی کنند. برای این کار، آن‌ها تعداد کل "کلمات عرضه شده" و "کلمات مصرف شده" را هر روز در ایالات متحده و ژاپن برای یک سال معین، از هجده منبع رسانه‌ای مختلف، از جمله رادیو، تلویزیون، کتاب‌ها، روزنامه‌ها، مجلات و آنچه در آن زمان یک دسته نسبتاً جدید اما به سرعت در حال رشد بود، یعنی "ارتباطات داده"، جدول‌بندی کردند. در ایالات متحده، آن‌ها تخمین زدند که "کلمات عرضه شده" از ۸٫۷ میلیون در روز در سال ۱۹۷۵ به ۱۱ میلیون در روز در سال ۱۹۸۰ رشد کرد. در مقابل، "کلمات مصرف شده" هر روز برای آن سال‌ها از ۴۵,۰۰۰ به ۴۸,۰۰۰ افزایش یافت.

می‌توانید حس کنید که این روند چگونه در طول زمان پیش رفت. از یک سو، طیف وسیعی از ایستگاه‌های رادیویی، کانال‌های کابلی، خرده‌فروشان کاتالوگی، ناشران سی‌دی‌رام، گروه‌های خبری اینترنتی، ارسال‌کنندگان ایمیل‌های اسپم، پلتفرم‌های پخش ویدئو، پادکسترها و

اینفلوئنسرهای تیک‌تاک، میلیون‌ها و میلیون‌ها کلمه بیشتر در روز عرضه می‌کردند. و از سوی دیگر، مغزهای کند، به راحتی پریشان و با محدودیت زمانی ما وجود داشتند که احتمالاً در اوایل دهه ۱۹۹۰ به حداکثر "کلمات مصرف شده" رسیدند. بر اساس تخمین‌های پول، هر ساکن ایالات متحده در سال ۱۹۸۰ تنها ۰٫۰۰۴ از اطلاعات موجود در هر روز را مصرف می‌کرد. و امروز؟ در مدت زمانی که طول می‌کشد تا این جمله را بخوانید، جهان اکنون به اندازه ۲۳ میلیارد کتاب الکترونیکی داده و اطلاعات تولید می‌کند. بخشی از آن از انسان‌ها می‌آید، در قالب توییت‌ها، ورودی‌های ویکی‌پدیا، مخازن گیت‌هاب، مقالات سفید منتشر شده درarXiv.org ، راهنماهای IRSو رقص‌های تیک‌تاک. بخشی از آن از گوشی‌های هوشمند، ترموستات‌های هوشمند، دوربین‌های امنیتی و سایر زیرساخت‌های اینترنت اشیا (IoT) می‌آید، در قالب داده‌هایGPS ، خوانش‌های دما، فیلم‌های ویدئویی و موارد دیگر. این بدان معناست که حتی اگر هر قسمت از پادکست کارا سویشر را ببینید و بیشتر توییت‌های مت ایگلزیاس را بخوانید، همچنان اساساً در "عصر تاریکی" شخصی خود زندگی می‌کنید و تقریباً از تمام دانش جهانی بی‌خبر هستید.

و دقیقاً به همین دلیل است که هوش مصنوعی برای آینده ما، به عنوان افراد و به صورت جمعی، بسیار حیاتی است. همانطور که ما از قدرت بخار و اکنون از چندین منبع نیروی کار مصنوعی برای تقویت عامل‌بودگی انسان استفاده کردیم، اکنون می‌توانیم از هوش مصنوعی نیز استفاده کنیم. با توزیع گسترده هوش، توانمندسازی مردم با ابزارهای هوش مصنوعی که به عنوان امتداد اراده‌های فردی انسان عمل می‌کنند، می‌توانیم داده‌های بزرگ را به دانش بزرگ تبدیل کنیم تا به یک "عصر روشنگری" جدید از وضوح و رشد مبتنی بر داده دست یابیم.

نتیجه‌گیری

سفر از کابوس اورولی ۱۹۸۴ به واقعیت پیچیده عصر دیجیتال، پر از تناقضات و تحولات غیرمنتظره است. در حالی که ترس از نظارت همه‌جانبه و از دست دادن حریم خصوصی در میانه قرن بیستم، با ظهور کامپیوترهای مین‌فریم و طرح‌های دولتی مانند مرکز داده فدرال، اوج گرفت، تاریخ نشان داد که فناوری مسیری متفاوت را طی کرده است. کامپیوترهای شخصی و سپس اینترنت، به جای محدود کردن آزادی، ابزاری قدرتمند برای گسترش عامل‌بودگی فردی، ابراز وجود و ایجاد جوامع متنوع‌تر فراهم آوردند. مفهوم "شبکه اعتماد"، که توسط پلتفرم‌هایی مانند لینکدین پیشگامی شد، نشان داد که چگونه به اشتراک‌گذاری هویت واقعی در یک بستر متصل می‌تواند ارزش جمعی بی‌سابقه‌ای ایجاد کند و منجر به فرصت‌های جدید برای افراد و سازمان‌ها شود.

امروزه، با حجم بی‌سابقه داده‌های تولید شده توسط انسان و ماشین، چالش اصلی نه در کمبود اطلاعات، بلکه در ظرفیت ما برای تبدیل این "داده‌های بزرگ" به "دانش بزرگ" نهفته است. در این میان، هوش مصنوعی نقشی محوری ایفا می‌کند. هوش مصنوعی نه تنها می‌تواند به ما در پردازش و تحلیل این حجم عظیم از اطلاعات کمک کند، بلکه می‌تواند به عنوان امتدادی از اراده‌های فردی انسان، عامل‌بودگی ما را در تمامی ابعاد زندگی افزایش دهد. پذیرش هوش مصنوعی با احتیاط و با در نظر گرفتن درس‌های گذشته، می‌تواند ما را از یک "عصر تاریکی" شخصی به یک "عصر روشنگری" جدید مبتنی بر داده رهنمون سازد، جایی که اطلاعات نه تنها محدودکننده نیستند بلکه قدرت‌بخش، الهام‌بخش و راهی برای دستیابی به سطوح جدیدی از معنا و رضایت در زندگی محسوب می‌شوند. تعادل بین حفظ حریم خصوصی و بهره‌برداری از قدرت دانش جمعی، چالش همیشگی ماست، اما با رویکردی هوشمندانه به فناوری، می‌توانیم آینده‌ای را بسازیم که در آن "دانش بزرگ" به ابزاری برای رشد و توانمندسازی همگانی تبدیل شود.

نکات کلیدی

- "1984"جورج اورول، تصویری پیشگویانه از نظارت دولتی و محدود کردن فردیت را ارائه داد که الهام‌بخش نگرانی‌های اولیه درباره فناوری شد.
- در دهه ۱۹۶۰، ظهور کامپیوترهای بزرگ هم توانایی‌های بی‌سابقه‌ای را فراهم آورد و هم ترس از "جامعه عریان" و کنترل اطلاعاتی را دامن زد.
- طرح "مرکز داده فدرال" در آمریکا با هدف یکپارچه‌سازی داده‌ها برای سیاست‌گذاری بهتر، به دلیل نگرانی‌های شدید درباره حریم خصوصی و سوءاستفاده از اطلاعات، با مخالفت گسترده مواجه شد و شکست خورد.
- پیشرفت‌های فناورانه، از جمله کامپیوترهای شخصی و اینترنت، به جای محدود کردن، به افزایش عامل‌بودگی فردی، شخصی‌سازی و تنوع فرهنگی کمک کردند.
- پلتفرم‌هایی مانند لینکدین، با محوریت هویت واقعی و ایجاد "شبکه‌های اعتماد"، نشان دادند که به اشتراک‌گذاری اطلاعات می‌تواند ارزش جمعی عظیمی ایجاد کند.
- در عصر حاضر، حجم داده‌های تولید شده بسیار فراتر از ظرفیت مصرف انسان است، که منجر به "عصر تاریکی شخصی" می‌شود.
- هوش مصنوعی نقشی حیاتی در تبدیل "داده‌های بزرگ" به "دانش بزرگ" دارد و می‌تواند به عنوان ابزاری برای تقویت عامل‌بودگی انسانی عمل کند.
- تعادل میان حفظ حریم خصوصی و بهره‌برداری از قدرت دانش اشتراکی، یک چالش مداوم است، اما اطلاعات به اشتراک گذاشته شده می‌تواند توانمندساز باشد.
- هویت عمومی در دنیای شبکه‌ای امروز به معنای قابل کشف بودن، قابل اعتماد بودن و داشتن نفوذ است که نوعی سرمایه اجتماعی محسوب می‌شود.
- با وجود نگرانی‌های جدید مرتبط با هوش مصنوعی، تاریخ نشان داده است که فناوری پتانسیل زیادی برای آزاد کردن و توانمندسازی انسان‌ها دارد.

سؤالات تفکربرانگیز

۱. با توجه به تفاوت‌های پیش‌بینی‌های اورولی و واقعیت‌های امروزین، چگونه می‌توانیم بهترین درس‌ها را از تاریخ فناوری و جامعه‌شناسی اطلاعات بگیریم تا آینده‌ای متعادل‌تر بسازیم؟

۲. در دنیایی که هویت عمومی به طور فزاینده‌ای به سرمایه اجتماعی و اقتصادی تبدیل می‌شود، چگونه می‌توان مرز بین ابراز وجود و حفاظت از حریم خصوصی را به گونه‌ای ترسیم کرد که برای افراد و جامعه مفید باشد؟

۳. با ظهور هوش مصنوعی و توانایی آن در تحلیل مقادیر بی‌سابقه داده، چه مسئولیت‌های اخلاقی و قانونی جدیدی برای توسعه‌دهندگان و کاربران این فناوری ایجاد می‌شود تا از سوءاستفاده جلوگیری شود؟

۴. چگونه می‌توان به شهروندان کمک کرد تا سواد اطلاعاتی و دیجیتالی خود را افزایش دهند تا در برابر خطرات احتمالی داده‌های بزرگ و هوش مصنوعی محافظت شوند، در حالی که از مزایای آن بهره‌مند می‌شوند؟

۵. آیا جامعه امروز، با وجود دسترسی بی‌سابقه به اطلاعات و ابزارهای ارتباطی، به دلیل حجم بیش از حد اطلاعات، در معرض خطر از دست دادن توانایی تفکر انتقادی و قضاوت مستقل قرار دارد؟

۳۰ جمله کلیدی

۱. نگاه اورول در "۱۹۸۴" تصویری هراس‌انگیز از نظارت فناورانه ارائه داد که نگرانی‌های عمومی را شکل داد.

۲. اقیانوسیه تحت سلطه "حزب" و "برادر بزرگ" با نظارت دائمی و تبلیغات خردکننده اداره می‌شد.

۳. "تلسکرین‌ها" در ۱۹۸۴ هم پروپاگاندا پخش می‌کردند و هم شهروندان را زیر نظر

داشتند.

۴. "پلیس فکر" برای ردیابی هرگونه نشانه بی‌وفایی یا احساسات انسانی عمل می‌کرد.

۵. در اوایل دهه ۱۹۶۰، ظهور کامپیوترهای مین‌فریم موجی از پیش‌بینی‌های بدبینانه را برانگیخت.

۶. ونس پاکارد در "جامعه عریان" از ماشین‌های حافظه غول‌پیکر که می‌توانند زندگی افراد را ثبت کنند، هشدار داد.

۷. طرح "مرکز داده فدرال" در ۱۹۶۵ برای یکپارچه‌سازی داده‌های دولتی جهت سیاست‌گذاری عمومی بود.

۸. این طرح با مخالفت شدید کنگره و عموم مردم به دلیل نگرانی از نقض حریم خصوصی مواجه شد.

۹. قانونگذاران نگران تبدیل شهروندان به "انسان‌های کامپیوتری‌شده" و از دست دادن هویت بودند.

۱۰. پرونده "گریزولد علیه کنتیکت" در ۱۹۶۵، "حق حریم خصوصی" را برای اولین بار در آمریکا شناسایی کرد.

۱۱. شکست مرکز داده ملی، مانع از سرعت نوآوری‌های فناورانه نشد و ابزارهای نظارتی پیچیده‌تر پدید آمدند.

۱۲. با این حال، پیش‌بینی‌های پاکارد درباره "زنجیرهای فناوری" تا حد زیادی اشتباه از آب درآمد.

۱۳. کامپیوترها و فناوری‌های شخصی‌سازی، عامل بودگی فردی و تنوع فرهنگی را گسترش دادند.

۱۴. "برادر بزرگ" یا "کسب‌وکار بزرگ" به جای سرکوب، باعث "احساس دیده شدن" افراد شد.

۱۵. هویت عمومی در دنیای شبکه‌ای، به کشف پذیری، قابلیت اعتماد و نفوذ منجر می‌شود.

۱۶. پلتفرم‌هایی مانند لینکدین، مفهوم "شبکه اعتماد" را از طریق هویت واقعی شکل دادند.

۱۷. لینکدین با تشویق کاربران به اشتراک‌گذاری اطلاعات حرفه‌ای، مشکلات اعتماد را حل کرد.

۱۸. قبل از اینترنت، شبکه‌های حرفه‌ای شخصی و به‌روزرسانی اطلاعات دشوار بود.

۱۹. پلتفرم‌های اجتماعی به جای "جزایر داده" شخصی، یک سیستم متصل و خود به‌روز شونده ایجاد کردند.

۲۰. چالش‌های اولیه لینکدین شامل مقاومت در برابر به اشتراک‌گذاری نام واقعی و اطلاعات شغلی بود.

۲۱. با این حال، مزایای جمعی مانند فرصت‌های شغلی و توسعه هویت حرفه‌ای بر این مقاومت غلبه کرد.

۲۲. داده‌های جمعی در لینکدین به افراد کمک می‌کند تا مهارت‌های مورد نیاز و شرکت‌های محبوب را شناسایی کنند.

۲۳. زندگی عمومی در عصر دیجیتال می‌تواند منجر به ماندگاری و گسترش ناخواسته اطلاعات شود.

۲۴. جوامع به جریان آزاد اطلاعات و همچنین حریم خصوصی برای حفظ دموکراسی وابسته هستند.

۲۵. بشریت در حال حاضر اطلاعات بسیار بیشتری از آنچه که می‌تواند به تنهایی مصرف کند، تولید می‌کند.

۲۶. نسبت "کلمات عرضه شده" به "کلمات مصرف شده" به شدت افزایش یافته و منجر

به اضافه‌بار اطلاعاتی شده است.

۲۷. هوش مصنوعی برای تبدیل این "داده‌های بزرگ" به "دانش بزرگ" ضروری است.

۲۸. هوش مصنوعی می‌تواند به عنوان امتداد اراده‌های فردی انسان، عامل‌بودگی را تقویت کند.

۲۹. درس‌های تاریخ نشان می‌دهد که اطلاعات به اشتراک گذاشته شده می‌تواند توانمندساز و الهام‌بخش باشد.

۳۰. تعادل میان حریم خصوصی و دانش بزرگ، کلید دستیابی به "عصر روشنگری" نوین است.

فصل ۳

چه اتفاق خوبی ممکن است رخ دهد؟

"معیار واقعی هوش ما نه در چالش‌هایی که می‌آفرینیم، بلکه در راه‌حل‌هایی است که برای اعتلای بشریت ارائه می‌دهیم — ".یک پژوهشگر برجسته در اخلاق هوش مصنوعی

اهداف فصل:

- ارائه تعریفی دقیق از «راه‌حل‌گرایی» و «مشکل‌گرایی» و بررسی تأثیر آن‌ها بر نوآوری.
- بررسی پتانسیل هوش مصنوعی در ایجاد تغییرات مثبت در مقیاس وسیع و مقابله با چالش‌های جهانی.
- ارائه تحلیلی عمیق از بحران سلامت روان جهانی و محدودیت‌های رویکردهای سنتی و فناوری‌های موجود.
- تشریح نقش تحول‌آفرین هوش مصنوعی در توسعه، ارزیابی و ارائه خدمات سلامت روان در آینده.
- ترغیب به اتخاذ رویکرد «چه اتفاق خوبی ممکن است رخ دهد؟» به جای تمرکز صرف بر خطرات احتمالی.

چکیده

این فصل به بررسی دو رویکرد متضاد در مواجهه با پیشرفت‌های تکنولوژیک، به ویژه هوش مصنوعی، می‌پردازد: «راه‌حل‌گرایی» که به راه‌حل‌های فناورانه برای چالش‌های پیچیده اجتماعی باور دارد، و «مشکل‌گرایی» که فناوری را نیرویی ذاتاً مشکوک و ضدبشری می‌داند. در حالی که انتقادات سازنده از فناوری می‌تواند ارزشمند باشد، تأکید صرف بر خطرات احتمالی، فرصت‌های بی‌شماری را که ممکن است از نوآوری حاصل شود، نادیده می‌گیرد. این فصل استدلال می‌کند که هوش مصنوعی پتانسیل فوق‌العاده‌ای برای حل مبرم‌ترین چالش‌های جهانی از جمله انرژی پایدار، مراقبت‌های بهداشتی، آموزش و امنیت سایبری دارد. تمرکز اصلی بر روی پتانسیل تحول‌آفرین هوش مصنوعی در حوزه سلامت روان است، جایی که بحران فزاینده‌ای ناشی از کمبود متخصصان، دسترسی محدود و ناکارآمدی‌های سیستماتیک وجود دارد. با بررسی مورد بحث‌برانگیز پلتفرم «کوکو» و تحقیقات پیشرفته در مورد تحلیل داده‌های درمانی، نشان داده می‌شود که چگونه هوش مصنوعی می‌تواند به طور چشمگیری دسترسی به مراقبت‌ها را گسترش داده، درمان‌ها را شخصی‌سازی کرده و کارایی آن‌ها را افزایش دهد. در نهایت، این فصل با ترویج دیدگاهی موسوم به «چه اتفاق خوبی ممکن است رخ دهد؟» و مفهوم «فوق‌بشری» استدلال می‌کند که تعامل با هوش مصنوعی می‌تواند نه تنها مشکلات را حل کند، بلکه ظرفیت انسانی برای همدلی و مهربانی را نیز افزایش دهد و به یک جامعه بهتر و انسانی‌تر منجر شود.

مقدمه

در عصر حاضر، که با پیشرفت‌های خیره‌کننده تکنولوژیک، به ویژه در حوزه هوش مصنوعی، مشخص شده است، جامعه انسانی در برابر پرسش‌های بنیادین در مورد آینده و نقش این فناوری‌ها قرار گرفته است. این دوران، که برخی آن را «عصر روشنایی جدید» می‌نامند، شاهد ظهور ابزارهایی است که از مسواک‌های مجهز به هوش مصنوعی برای بهینه‌سازی شیوه مسواک زدن گرفته تا مدل‌های زبان بزرگ (LLM) که پتانسیل تغییر پارادایم در مراقبت‌های بهداشتی را دارند، را شامل می‌شود. با این حال، واکنش به این پیشرفت‌ها اغلب دو قطبی است، که می‌توان آن را در چارچوب دو دیدگاه متضاد «راه‌حل‌گرایی» و «مشکل‌گرایی» درک کرد.

راه‌حل‌گرایی، باوری است که مدعی است حتی پیچیده‌ترین چالش‌های جامعه، از جمله نابرابری‌های عمیق سیاسی، اقتصادی و فرهنگی، دارای راه‌حل‌های فناورانه ساده‌ای هستند. این دیدگاه، که اغلب به بنیان‌گذاران و سرمایه‌گذاران سیلیکون ولی نسبت داده می‌شود، گاهی اوقات به دلیل ساده‌انگاری در مواجهه با مشکلات چندوجهی مورد انتقاد قرار می‌گیرد. در مقابل، مشکل‌گرایی، رویکرد غالب «افراد بدبین» است که فناوری را به طور کلی یک نیروی مشکوک، ضدبشری و ضدانسانی می‌بیند که آلوده به بوی تازه سرمایه‌داری است. این افراد بدبین، خود را به عنوان محافظان منافع عمومی و نیروهای اخلاقی در برابر «فقر فزاینده ناشی از راه‌حل‌گرایی» معرفی می‌کنند و می‌توانند در پاسخگو کردن شرکت‌های بزرگ فناوری برای اشتباهات و کوتاهی‌ها که منجر به پیامدهای مضر اجتماعی می‌شود، ارزشمند باشند. با این حال، تمرکز صرف بر نقد به جای عمل، و ترجیح احتیاط و حتی ممنوعیت بر نوآوری، می‌تواند به جامعه آسیب جدی وارد کند. وقتی تنها بر آنچه ممکن است اشتباه پیش برود تمرکز می‌شود، ناگزیر آنچه ممکن است درست پیش برود، نادیده گرفته می‌شود.

این فصل با این پیش‌فرض آغاز می‌شود که در حالی که هوش مصنوعی ممکن است خطراتی داشته باشد که باید به آن‌ها رسیدگی شود، پتانسیل آن برای کمک به حل مبرم‌ترین چالش‌های جهانی غیرقابل انکار است. فناوری به عنوان یکی از مؤثرترین اهرم‌های بشریت برای ایجاد

تغییرات مثبت در مقیاس وسیع شناخته شده است. خواه در تولید انرژی پایدار، مراقبت‌های بهداشتی، آموزش یا امنیت سایبری، فناوری بین ۳۰ تا ۸۰ درصد از هر راه‌حل مؤثر را تشکیل می‌دهد. حتی برای مقابله با مسائل ناشی از خود هوش مصنوعی، مانند «دیپ‌فیک» و اطلاعات نادرست، به سیستم‌های تشخیص هوش مصنوعی نیاز داریم که بتوانند پلتفرم‌ها را در مقیاس وسیع نظارت کرده و به سرعت با چالش‌های جدید سازگار شوند.

بنابراین، هر تلاشی برای پیشگیری از پیامدهای منفی بالقوه یک نوآوری، همچنین پیامدهای مثبت بالقوه‌ای را که ممکن است تولید کند، از بین می‌برد. این فصل به تفصیل پتانسیل هوش مصنوعی را در حوزه‌های مختلف بررسی می‌کند، اما تأکید ویژه بر تحولاتی است که می‌تواند در بخش سلامت روان ایجاد کند. با وجود بحران جهانی سلامت روان، کمبود شدید متخصصان و موانع دسترسی، هوش مصنوعی می‌تواند دسترسی به مراقبت‌ها را به طور چشمگیری گسترش داده، درمان‌ها را شخصی‌سازی کرده و کارایی آن‌ها را افزایش دهد. هدف این فصل، ارائه دیدگاهی جامع و متعادل است که فراتر از تقابل ساده‌انگارانه راه‌حل‌گرایی و مشکل‌گرایی می‌رود و خواننده را به پذیرش یک ذهنیت «چه اتفاق خوبی ممکن است رخ دهد؟» تشویق می‌کند؛ ذهنیتی که نه تنها به خطرات احتمالی بی‌اعتنا نیست، بلکه با تصویرسازی آینده مطلوب و حرکت به سوی آن، به دنبال جلوگیری از این خطرات است.

۱. تقابل راه‌حل‌گرایی و مشکل‌گرایی در عصر هوش مصنوعی

همانطور که در مقدمه اشاره شد، گفتمان پیرامون هوش مصنوعی اغلب بین دو قطب افراطی «راه‌حل‌گرایی» و «مشکل‌گرایی» در نوسان است. راه‌حل‌گرایی، با ایمان راسخ به توانایی تکنولوژی برای رفع تمامی معضلات اجتماعی، گاهی اوقات به ساده‌سازی چالش‌هایی متهم می‌شود که ریشه‌های عمیق سیاسی، اقتصادی و فرهنگی دارند. این رویکرد، در برخی مواقع، به نادیده گرفتن پیچیدگی‌های انسانی و اجتماعی منجر می‌شود. به عنوان مثال، توسعه مسواک‌های هوش مصنوعی که شیوه مسواک زدن را بهینه می‌کنند، هرچند نمونه‌ای از نوآوری

است، اما نمی‌تواند راه حلی برای نابرابری‌های ساختاری در دسترسی به مراقبت‌های بهداشتی باشد. با این حال، نادیده گرفتن نقش فناوری به عنوان یکی از قدرتمندترین اهرم‌های بشریت برای ایجاد تغییرات مثبت در مقیاس وسیع، خود یک رویکرد ساده‌انگارانه است.

در مقابل، مشکل‌گرایی، که مشخصه افراد بدبین است، فناوری را ذاتاً نیرویی مشکوک، ضدبشری و ضدانسانی می‌داند که با منطق سرمایه‌داری در هم آمیخته است. این افراد، که اغلب خود را به عنوان دیده‌بانان منافع عمومی و نیروهای اخلاقی در برابر پیامدهای منفی فناوری می‌پندارند، می‌توانند نقش حیاتی در پاسخگو کردن شرکت‌های بزرگ فناوری در قبال اشتباهات، کوتاهی‌ها و پیامدهای مضر داشته باشند. این نظارت، برای اطمینان از توسعه مسئولانه و اخلاقی هوش مصنوعی، کاملاً ضروری است. با این حال، زمانی که این دیدگاه تنها بر نقد تمرکز می‌کند و احتیاط و ممنوعیت را بر نوآوری ارجح می‌داند، می‌تواند آسیب‌های واقعی به جامعه وارد کند. تمرکز انحصاری بر آنچه ممکن است اشتباه پیش رود، ناگزیر، پتانسیل عظیم آنچه ممکن است درست پیش رود را نادیده می‌گیرد. این رویکرد می‌تواند به رکود و حفظ وضعیت موجود منجر شود، زیرا هر تلاشی برای پیشگیری از پیامدهای منفی بالقوه یک نوآوری، همچنین از پیامدهای مثبت بالقوه‌ای که ممکن است ایجاد کند، جلوگیری می‌کند.

برای نمونه، یک سیستم هوش مصنوعی را تصور کنید که بتواند نحوه تفسیر و ترجمه صداهای حیوانات را یاد بگیرد. این قابلیت می‌تواند به انسان‌ها امکان دهد تا نیازهای گونه‌های در معرض خطر را به روش‌هایی که قبلاً ممکن نبوده است، درک کنند و در نتیجه منجر به مداخلات مؤثرتر برای حفاظت از تنوع زیستی شود. یا سیستم هوش مصنوعی دیگری را در نظر بگیرید که راهی به طرز چشمگیری کارآمدتر برای بهینه‌سازی زنجیره‌های تأمین توزیع مواد غذایی ابداع می‌کند، به طور قابل توجهی ضایعات را کاهش داده و دسترسی به تغذیه را در مناطق ناامن غذایی بهبود می‌بخشد و بدین ترتیب، نتایج سلامت میلیون‌ها نفر را در سال ارتقا می‌دهد. به تعویق انداختن چنین ابزارهایی که می‌توانند به دانشمندان کمک کرده و تأثیر آن‌ها را تقویت کنند، یا تأخیر در توسعه مشاوران سلامت هوش مصنوعی و دستیاران حقوقی هوش مصنوعی که می‌توانند به افراد

و جمعیت‌های حاشیه‌نشین قدرت بیشتری ببخشند، از نظر اخلاقی قابل دفاع نیست. این به معنای قربانی کردن نقش پیشرو در ایجاد آینده فناوری جهان برای ایالات متحده و دموکراسی نیز خواهد بود.

در حالی که احتیاط، مشورت و حتی شک‌گرایی در زمینه توسعه هوش مصنوعی بسیار ضروری است، هدف نهایی پیشرفت است. این بدان معناست که ما باید سطح معینی از ریسک و عدم قطعیت را بپذیریم تا بتوانیم اقدام کنیم و به جلو حرکت کنیم. نوآوری نیازمند کاوش است. هرکس که خلاف این را بگوید، نه تنها در حال بیان حقیقت به قدرت نیست، بلکه در حال رأی دادن به منافع و نابرابری‌های ریشه‌دار است. در بسیاری موارد، افراد بدبین که در حالت مشکل‌گرایی عمل می‌کنند، چیزی بیش از قهرمانان رضایت و بی‌تفاوتی نیستند. به همین دلیل، پذیرش ذهنیت «چه اتفاق خوبی ممکن است رخ دهد؟» برای ایجاد تغییرات مثبت ضروری است. این نوع تفکر به معنای نادیده گرفتن پیامدهای منفی احتمالی نیست؛ بلکه به معنای جلوگیری از آن پیامدها با تصور آینده مطلوب و حرکت به سوی آن است.

۲ .تهدید وجودی وضعیت موجود: بحران سلامت روان

از منظر مشکل‌گرایی، کاربرد هوش مصنوعی در مراقبت‌های سلامت روان، یک «راه‌حل‌گرایی دسته‌بندی ۵» تلقی می‌شود؛ یک طوفان الگوریتمی از ناآگاهی متخصصان فناوری که با سرعت بالا به سمت فاجعه حرکت می‌کند. "درمان افراد بالقوه توهم‌زا در بحران با سیستمی که توهم تولید می‌کند؟ چه چیزی ممکن است اشتباه پیش برود؟" این دیدگاه، اگرچه به ظاهر محتاطانه است، اما اغلب تهدید وجودی وضعیت موجود را نادیده می‌گیرد. تهدید وجودی وضعیت موجود به هرگونه کاستی، آسیب، ناکارآمدی یا نابرابری اشاره دارد که برای مدت طولانی وجود داشته و تقریباً نامرئی شده است. در حوزه سلامت روان، این تهدید خاص، نیازهای مراقبت سلامت روان است که به طور مزمن برآورده نمی‌شوند.

آمارها نشان می‌دهد که نرخ افسردگی، اضطراب و خودکشی در ایالات متحده طی بیست سال

اخیر به طرز چشمگیری افزایش یافته است، در حالی که نتایج برای بیماری‌هایی مانند سرطان و بیماری‌های قلبی در حال بهبود است. در سال ۲۰۲۲، نزدیک به ۵۰,۰۰۰ آمریکایی جان خود را گرفتند که بالاترین نرخ سرانه از سال ۱۹۴۱ بود. علاوه بر این، بیش از ۱۰۰,۰۰۰ نفر دیگر به دلیل مصرف بیش از حد مواد مخدر جان باختند، که ارتباط قوی با مسائل سلامت روان دارد. این آمارها نه تنها نشانگرهای مهم بحران جاری سلامت روان هستند، بلکه مطالعات نشان می‌دهند که مسائل سلامت روان خطر مرگ و میر را از هر نوع افزایش می‌دهند و می‌توانند به کاهش ۱۰ تا ۲۰ ساله امید به زندگی منجر شوند. اختلالات روانی با معدل‌های پایین‌تر و نرخ ترک تحصیل بالاتر در میان دانشجویان نیز مرتبط بوده‌اند.

علاوه بر این، کارکنانی که سلامت روان خود را «متوسط یا ضعیف» ارزیابی می‌کنند، تقریباً چهار برابر بیشتر از همتایان خود که سلامت روان «خوب، بسیار خوب یا عالی» گزارش می‌کنند، غیبت‌های برنامه‌ریزی نشده به دلیل سلامت روان ضعیف دارند. افرادی که به دلیل اختلالات روانی مداوم، از دست دادن مکرر شغل را تجربه می‌کنند، می‌توانند در چرخه بیکاری طولانی‌مدت قرار گیرند، زیرا از دست دادن مهارت‌ها، شبکه‌های حرفه‌ای و اعتماد به نفس، همراه با شکاف‌های شغلی و تعصبات احتمالی کارفرمایان، یافتن و حفظ مشاغل جدید را به طور فزاینده‌ای دشوار می‌کند. یک گزارش در سال ۲۰۱۱ از مجمع جهانی اقتصاد پیش‌بینی کرد که هزینه‌های مستقیم و غیرمستقیم جهانی ناشی از سلامت روان ضعیف در سال ۲۰۳۰ به ۶ تریلیون دلار خواهد رسید.

گزارش‌ها نشان می‌دهد که ۱۲۹ میلیون آمریکایی در مناطقی زندگی می‌کنند که با کمبود متخصصان سلامت روان مواجه هستند. کمتر از یک‌سوم جمعیت ایالات متحده در مکان‌هایی زندگی می‌کنند که دسترسی کافی به متخصصان سلامت روان برای رفع نیازهای محلی دارند. در نتیجه، اغلب ممکن است سه تا شش ماه طول بکشد تا از یک متخصص واجد شرایط مراقبت دریافت شود. این کمبود منابع موجود زمانی بارزتر می‌شود که در نظر بگیریم بسیاری از افراد تنها زمانی به دنبال خدمات سلامت روان می‌روند که با چالش‌ها یا پریشانی‌های قابل توجهی روبرو

هستند. اگر به سمت یک رویکرد پیشگیرانه‌تر حرکت کنیم و مراقبت‌های سلامت روان را جزء مداوم رفاه کلی – منبعی برای زمان‌های نسبی رفاه و بحران – بدانیم، شکاف بین مراقبت‌های موجود و تقاضای بالقوه احتمالاً حتی بزرگتر از آن چیزی خواهد بود که معمولاً تصور می‌کنیم.

۳ .محدودیت‌های راه حل‌های موجود و پتانسیل هوش مصنوعی

استفاده از فناوری و اتوماسیون برای گسترش دسترسی به مراقبت‌های سلامت روان فراتر از جلسات سنتی حضوری ۵۰ دقیقه‌ای، یک عمل طولانی‌مدت است. خطوط کمک بحران از دهه ۱۹۵۰ وجود داشته‌اند. تله‌تراپی، گروه‌های پشتیبانی آنلاین، برنامه‌های سلامت روان خودراهنما و پلتفرم‌های مشاوره آنلاین مانند BetterHelp و Talkspace همگی در سال‌های اخیر، به ویژه در دوران همه‌گیری کووید-۱۹، به طور فزاینده‌ای رایج شده‌اند. اکنون بیش از ۱۰،۰۰۰ برنامه سلامت روان وجود دارد که شامل جلسات درمانی خودراهنما، تمرینات ذهن‌آگاهی و مدیتیشن، ردیابی خلق و خو، تکنیک‌های CBT و موارد دیگر است. با این حال، دسترسی به مراقبت برای بخش‌های بزرگی از جمعیت همچنان دست‌نیافتنی باقی مانده است. قوانین مهم فدرال که شرکت‌های بیمه سلامت را ملزم به پوشش گسترده‌تر خدمات سلامت روان می‌کنند، در سال ۲۰۰۸ و ۲۰۱۰ به تصویب رسیدند، اما شکاف‌های درمانی همچنان پابرجاست.

یکی از چالش‌های اصلی در اینجا، موضوع «جذب و تعامل» است. حتی با وجود هزاران برنامه هوشمند سلامت روان، این به معنای تأثیر واقعی در دنیای واقعی نیست. همانند هر نوع مداخله پزشکی، دیجیتال یا غیره، اولین گام کلیدی در دستیابی به بهبودهای قابل اندازه‌گیری در نتایج بیمار، اعتبار سنجی بالینی در آزمایش‌های کنترل شده تصادفی است. اما حتی دسترسی به علاوه اثربخشی اثبات شده نیز اگر مردم از محصول استفاده نکنند، یا پس از شروع استفاده به آن پایبند نباشند، اهمیت کمی دارد. بسیاری از مطالعات سلامت روان در دو دهه گذشته نشان داده‌اند که ماهیت عمدتاً خودراهنمای برنامه‌ها منجر به سطوح پایین تعامل و نرخ بالای ترک می‌شود. یک مطالعه در سال ۲۰۱۹ روی چنین برنامه‌هایی نشان داد که نرخ نگهداری پس از تنها پانزده روز

استفاده به میانگین ۳.۹ درصد کاهش یافته است.

یکی از راه‌های بهبود نرخ تعامل و نگهداری، گنجاندن عناصر مکالمه‌ای یا اجتماعی بیشتر در برنامه‌ها و سایر مداخلات است. بالاخره، یکی از دلایلی که گفتاردرمانی سنتی ستون فقرات روان‌درمانی است، این است که مردم صحبت با دیگران را جذاب می‌یابند؛ ما حیوانات اجتماعی هستیم. این همان کاری است که راب موریس با پلتفرم کوکو از طریق پشتیبانی همتایان انجام می‌دهد. همچنین کاری است که توسعه‌دهندگان برنامه‌هایی مانند Woebot و Wysa از طریق چت‌بات‌هایی انجام می‌دهند که با درگیر شدن در گفتگوی مداوم با کاربران، گفتاردرمانی سنتی را شبیه‌سازی می‌کنند. با این حال، در حالی که چت‌بات‌های تخصصی سلامت روان مانند Woebot و Wysa از پردازش زبان طبیعی پیشرفته استفاده می‌کنند، اما به اندازه مدل‌های بنیادی عمومی مانند GPT-4 مولد نیستند. در عوض، آن‌ها به ساختارهای از پیش تعریف‌شده‌ای به نام «فریم» متکی هستند که به چت‌بات‌های سنتی کیفیت تا حدودی سفت و سخت و قابل پیش‌بینی یک بازاریاب تلفنی ماهر را می‌دهد که در حال کار بر روی یک اسکریپت است، هرچند با مسیرهای شاخه‌دار متعدد. بنابراین، در حالی که این چت‌بات‌های تخصصی سلامت روان ظرفیت ارائه پاسخ‌های همدلانه و شبیه‌سازی یک گفتگوی درمانی را دارند، اما در نحوه پاسخگویی به ورودی کاربر بسیار محدود هستند. این بدان معناست که آن‌ها مستعد «توهم» مانند ChatGPT نیستند – محتوای منبع آن‌ها بر اساس تکنیک‌های درمانی مبتنی بر شواهد است که توسط متخصصان سلامت روان تأیید شده‌اند. با این حال، نقطه ضعف این است که ظرفیت آن‌ها برای پاسخگویی به روش‌هایی که برای یک زمینه یا اظهار نظر کاربر مناسب‌تر است، به اندازه درمانگران انسانی یا مدل‌های هوش مصنوعی پیشرفته‌تر، ظریف یا سازگار نیست.

۴ .درس‌هایی از ماجرای «کوکو»: بینش‌های حاصل از کاربرد هوش مصنوعی در سلامت روان

پلتفرم کوکو(Koko) ، که یک سرویس پیام‌رسان سلامت روان مبتنی بر همتا است، با افزودن قابلیت‌های تولید متن مبتنی بر3-GPT ، به طور موقت به محور بحث‌های داغ در مورد پتانسیل و خطرات هوش مصنوعی در این حوزه تبدیل شد. راب موریس، توسعه‌دهنده این پلتفرم، در ژانویه ۲۰۲۳ اعلام کرد که کاربران پیام‌هایی که توسط GPT-3 تولید یا هم‌تولید شده بودند را به طور قابل توجهی بالاتر از پیام‌هایی که توسط کاربران انسانی کوکو به تنهایی نوشته شده بودند، ارزیابی کرده بودند. با این حال، تیم تصمیم گرفت این گزینه جدید را به سرعت از سرویس خود حذف کند، زیرا "وقتی مردم متوجه شدند پیام‌ها توسط یک ماشین هم‌تولید شده‌اند، دیگر کار نکرد." موریس اظهار داشت: "همدلی شبیه‌سازی شده عجیب و خالی به نظر می‌رسد".

این توییت به سرعت وایرال شد و واکنش‌های شدیدی را برانگیخت، به طوری که بسیاری از کاربران X.com (توییتر سابق) و رسانه‌ها به شدت کوکو را به گمراه‌کردن و خیانت به کاربران آسیب‌پذیرش متهم کردند. این واکنش‌ها بدون درک دقیق از نحوه عملکرد کوکو و با فرض آسیب‌های جدی به کاربران صورت گرفت. اما موریس بعداً توضیح داد که عبارت "وقتی مردم متوجه شدند" به خود او و تیمش اشاره دارد، نه کاربران کوکو. آن‌ها پس از چند روز، زمانی که پیام‌های تولید شده صرفاً توسط GPT-3 و بدون اصلاحات کاربر را شناسایی کردند، آن‌ها را به طور فزاینده‌ای "بی‌روح" یافتند. همچنین موریس تأکید کرد که در طول مدتی که پیام‌رسانی با کمک GPT-3 در دسترس کاربران کوکو بود، هر پیامی که توسط «کوکوبوت» (یک چت‌بات مبتنی بر هوش مصنوعی که به عنوان میزبان یا راهنمای کلی در سرویس پیام‌رسان کوکو عمل می‌کند) نوشته یا هم‌تولید شده بود، دارای یک سلب مسئولیت بود که نشان می‌داد "با همکاری کوکوبوت نوشته شده است." در واقع، این یک رویکرد «کمک‌راننده (copilot) «کاملاً شفاف به

هوش مصنوعی بود و مشارکت کوکوبوت در هر مرحله از فرآیند به وضوح اطلاع‌رسانی می‌شد.

این ماجرا، نمونه‌ای کلاسیک از «مشکل‌گرایی» در عمل بود. صدها نفر در شبکه‌های اجتماعی خشم خود را به نمایندگی از کاربران گمراه و بالقوه آسیب‌دیده کوکو ابراز کردند، در حالی که هیچ کاربری از کوکو شکایتی را در مورد این ویژگی یا نحوه عملکرد آن مطرح نکرده بود. تمام این درخواست‌ها برای پیگرد قانونی و تحریم علیه کوکو در پاسخ به آسیب‌های کاملاً فرضی و کاربران خیالی بود.

در مقابل، راب موریس بر «تهدید وجودی وضعیت موجود» تمرکز داشت: نیازهای مراقبت سلامت روان که به طور مداوم برآورده نمی‌شوند. کوکو با ارائه یک رابط کاربری ساده برای مشارکت در یک تمرین مبتنی بر همتا و مبتنی بر شواهد به نام «ارزیابی مجدد شناختی»، تلاش می‌کند تا وضعیت موجود را دگرگون کند. این سرویس کاربران را از طریق یک چت‌بات به نام کوکوبوت راهنمایی می‌کند و به آن‌ها کمک می‌کند تا پیام‌هایی را ایجاد کنند که اصول ارزیابی مجدد شناختی را در بر می‌گیرند. کوکوبوت همچنین به عنوان یک واسطه برای هر تعامل همتا به همتا عمل می‌کند، محتوای پیام را در زمان واقعی تجزیه و تحلیل و تعدیل می‌کند و در صورت لزوم، کاربران را به خطوط کمک بحران هدایت می‌کند. این سیستم، قبل از افزودن قابلیت‌های GPT-3، به پردازش زبان طبیعی سنتی و تکنیک‌های یادگیری ماشین متکی بود. این مورد نشان می‌دهد که نوآوری در هوش مصنوعی، حتی با وجود پتانسیل‌های بی‌شمار، باید با شفافیت و دقت فراوان در جمعیت‌های آسیب‌پذیر انجام شود. با این حال، انتقاداتی که بدون درک کامل از زمینه و با تمرکز صرف بر فرضیات منفی مطرح می‌شوند، می‌توانند مانع پیشرفت‌های واقعی شوند.

۵ .هوش مصنوعی به عنوان کاتالیزور تحول در سلامت روان

با توجه به کمبود شدید متخصصان سلامت روان و دسترسی محدود به مراقبت‌ها، هوش مصنوعی می‌تواند نقش تحول‌آفرینی در غلبه بر «تهدید وجودی وضعیت موجود» ایفا کند. در

حال حاضر، تشخیص‌ها و روش‌های درمانی در سلامت روان اغلب بر اساس کلمات و ارزیابی‌های ذهنی بنا شده‌اند که مستعد تفسیر نادرست و تحریف‌های شناختی هستند. هوش مصنوعی می‌تواند با ارائه روش‌های جمع‌آوری داده عینی‌تر و مستمر، این عوامل را خنثی کند. از طریق گوشی‌های هوشمند و دستگاه‌های پوشیدنی، می‌توان رفتار را به صورت غیرفعال نظارت کرد. برنامه‌های سلامت روان می‌توانند کاربران را به صورت فعال در مورد خلق‌وخو و رفتارهایشان سؤال کنند تا به خود-نظارتی مداوم‌تر کمک کنند. حتی سیستم‌های هوش مصنوعی که از داده‌های موقعیت مکانی، فرکانس ارسال پیامک و مدت مکالمات استفاده می‌کنند، می‌توانند شروع افسردگی یا دوره‌های دوقطبی را پیش‌بینی کنند.

علاوه بر این، سیستم‌های هوش مصنوعی می‌توانند مجموعه‌های عظیمی از متن جلسات درمانی را تجزیه و تحلیل کنند تا بهتر بفهمند کدام مداخلات در زمینه‌های مختلف بهترین عملکرد را دارند و کدام رفتارهای درمانگر به نتایج درمانی موفق منجر می‌شوند. به عنوان مثال، در ژانویه ۲۰۲۴، محققان وابسته به Lyssn.io و Talkspace مطالعه‌ای را منتشر کردند که در آن بیش از ۱۶۰,۰۰۰ جلسه مشاوره متنی ناشناس که بین سال‌های ۲۰۱۴ تا ۲۰۱۹ در Talkspace انجام شده بود، با استفاده از هوش مصنوعی تجزیه و تحلیل شد. در این جلسات مشاوره، بیماران بیش از ۲۰ میلیون پیام با درمانگران دارای مجوز رد و بدل کرده بودند. محققان با استفاده از طبقه‌بندی‌کننده‌های متنی توسعه‌یافته توسط Lyssn.io ، توانستند تمامی این پیام‌ها را در سطحی بسیار دقیق، هم از نظر موضوعات مورد بحث و هم از نظر انواع خاص رفتارهای درمانگر مورد استفاده، دسته‌بندی کنند. هر پیام می‌توانست با کدهایی مانند «کار»، «والدین»، «پرسیدن سؤالات باز» و «اظهارات گوش دادن تأملی» برچسب‌گذاری شود.

سپس محققان توانستند این اظهارات کدگذاری شده را با داده‌های مربوط به علائم بیمار، سطوح رضایت مشتری، مدت زمان تعاملات مشتری و تغییر علائم که از طریق یک پرسشنامه خودارزیابی مشتری برای علائم افسردگی (PHQ-8) بیان شده بود، مرتبط کنند. به این ترتیب،

آن‌ها توانستند بینش‌های جدیدی در مورد اینکه کدام انواع رفتارهای درمانگر و مداخلات در کدام زمینه‌های خاص بهترین عملکرد را دارند، به دست آورند. برای مثال، این داده‌ها نشان داد که در حالی که «ارائه اطلاعات» رایج‌ترین مداخله درمانگر بود و در بیش از نیمی از پیام‌های درمانگر رخ می‌داد، اما ارتباط کوچک اما معنی‌داری با نتایج بدتر داشت. این یافته قدرت تجزیه و تحلیل مجموعه‌داده‌های عظیم را نشان می‌دهد، زیرا مطالعات کوچک‌تر ممکن است این بینش غیرمنتظره را که ارائه بیش از حد آموزش روان‌شناختی می‌تواند برای پیشرفت درمان مضر باشد، آشکار نمی‌کردند.

در مقابل، مداخلاتی مانند «بازتاب‌های پیچیده» یا بیان مجددی که معنا یا تأکید قابل توجهی به آنچه مشتری گفته است اضافه می‌کند، و «تأییدات» یا اظهاراتی که نقاط قوت و تلاش‌های مشتری را به رسمیت می‌شناسد و تأیید می‌کند، با نتایج بهتر مرتبط بودند. با آشکار کردن الگوها و روابطی که در موارد فردی، یا حتی ده‌ها یا صدها مورد بعید است، چنین تجزیه و تحلیل‌های در مقیاس بزرگ می‌تواند به متخصصان کمک کند تا ظرافت‌های درمان مؤثر را بهتر درک کرده و عملکردهای درمانی خود را بهبود بخشند.

پژوهشگران در حوزه روان‌شناسی بالینی محاسباتی نیز در حال بررسی چگونگی ادغام مدل‌های زبان بزرگ (LLMs) به عنوان دستیاران در عمل خود هستند. آن‌ها سه مرحله بالقوه از ادغام هوش مصنوعی در فضای سلامت روان را پیش‌بینی می‌کنند. **مرحله ۱** شامل استفاده‌های کمکی نسبتاً ساده از هوش مصنوعی است، مانند استفاده از ضبط جلسات و یادداشت‌های واقعی روان‌درمانگر برای پیش‌نویس یادداشت‌های رسمی‌تر جلسه، برنامه‌های درمانی و سایر اسناد اداری که روان‌درمانگران باید برای هر بیمار خود نگهداری کنند. **مرحله ۲** شامل تعاملات مشارکتی‌تر است، مانند بررسی کامل متن جلسات برای ارزیابی میزان پایبندی روان‌درمانگران کارآموز به روش‌های مبتنی بر شواهد، یا ارائه کمک و راهنمایی به مشتریان در انجام تکالیف بین جلسات، مانند تکمیل انواع مختلف برگه‌های کار CBT. **مرحله ۳** شامل مراقبت کاملاً خودمختار

است: مدل‌های زبان بزرگ بالینی که بر روی روش‌های مبتنی بر شواهد آموزش دیده‌اند و به شدت از نظر سودمندی، کارایی و ایمنی ارزیابی شده‌اند، می‌توانند تمامی وظایف و مداخلات درمانی را که متخصصان بالینی انسانی انجام می‌دهند، مدیریت کنند.

۶. چشم‌انداز سلامت روان فراوان و دسترس‌پذیر

در حالی که برخی ممکن است با حرکت به سمت مرحله ۳، پیش‌بینی از دست دادن شغل را داشته باشند، یک روایت بسیار جاه‌طلبانه‌تر و به همان اندازه معقول نیز ممکن است: روایتی که در آن مراقبت‌های سلامت روان به همان شیوه که موسیقی در اسپاتیفای و ویدئو در نتفلیکس ارائه می‌شود، فراهم گردد. یعنی، به طور دسترس‌پذیر، اقتصادی، مقیاس‌پذیر، مبتنی بر تحلیل داده‌ها، و بسیار قابل تنظیم بر اساس ترجیحات و نیازهای منحصربه‌فرد کاربران فردی. این یک فرصت بی‌نظیر برای استفاده از علم داده است تا کارهایی فراتر از بهینه‌سازی کشف ویدئوهای گربه یا پیش‌بینی دقیق اوج تقاضا برای شناورهای استخری فلامینگوی صورتی انجام شود.

امروزه، انتخاب متخصصان سلامت روان اغلب به راحتی و عمل‌گرایی بستگی دارد: چه کسی در شبکه بیمه شماست؟ چه کسی توسط یک دوست یا خانواده مورد اعتماد توصیه شده است؟ چه کسی در منطقه شماست؟ آیا اصلا بیمار جدید می‌پذیرند؟ تصور کنید که این وضعیت چگونه می‌تواند با فراگیر شدن هوش فراوان تغییر کند. وقتی مراقبت‌های سلامت روان واقعاً مقرون به صرفه و دسترس‌پذیر باشد، رویکردهای درمانی و تخصص می‌توانند اولویت پیدا کنند. پلتفرم‌های موجود مانند BetterHelp و Talkspace تا حدی این رویکرد را دنبال می‌کنند، اما اتکای انحصاری آن‌ها به متخصصان انسانی محدودیت‌هایی را ایجاد می‌کند. خدمات به اندازه اسپاتیفای یا نتفلیکس در دسترس نیستند؛ همچنان نیاز به برنامه‌ریزی وقت ملاقات وجود دارد. حتی اگر پیام‌رسانی متنی نامتقارن را به عنوان فرم ترجیحی تعامل خود انتخاب کنید، درمانگر شما احتمالاً ساعات مشخصی برای پاسخگویی خواهد داشت و ۷/۲۴ کار نمی‌کند. تعداد مشتریانی که یک درمانگر انسانی می‌تواند مدیریت کند نیز محدود است. و در نهایت، اگر بیمه

نداشته باشید، مطمئناً بسیار بیشتر از هزینه اشتراک نتفلیکس پریمیوم پرداخت خواهید کرد.

اما یک پلتفرم که در آن هم درمانگران کاملاً مجازی و هم درمانگران انسانی حضور دارند، به روش‌های بسیار کمتری محدود خواهد بود. چگونه مراقبت‌ها تغییر می‌کنند وقتی می‌توانید یک درمانگر از نظر بالینی آزمایش شده و اثبات شده را هر زمان که بخواهید، برای دو دقیقه یا دو ساعت، با هزینه ثابت ماهانه مثلاً ۱۹.۹۹ دلار در دسترس داشته باشید؟ ویکی‌پدیا نزدیک به ۲۰۰ نوع مختلف روان‌درمانی را فهرست می‌کند؛ منابع دیگر بیش از ۵۰۰ مورد را پیشنهاد می‌دهند. شاید روان‌شناسی تحلیلی بهترین رویکرد برای شما باشد، شاید نظریه رفتار دیالکتیکی. چگونه می‌توانید بدانید؟

از آنجا که هزینه دیگر یک مسئله نیست، می‌توانید به سادگی شروع به کاوش کنید. به سرعت می‌توانید چندین درمانگر مجازی را امتحان کنید و بدین ترتیب درک بهتری از آنچه فکر می‌کنید برای شما بهترین کارایی را دارد، به دست آورید. اگر به معنای وقت ملاقات‌های ساختاریافته است، آن‌ها را انتخاب می‌کنید. اگر می‌خواهید بتوانید به صورت خودجوش و آزادانه به درمانگر خود دسترسی داشته باشید، درست مانند جستجو در ویکی‌پدیا، این کار را انجام می‌دهید. شاید به طور منظم با بیش از یک درمانگر ملاقات کنید. شاید تیمی از درمانگران را گرد هم آورید که با آن‌ها به صورت مشترک ملاقات می‌کنید تا بتوانید در زمان واقعی از نظرات دوم و سوم بهره‌مند شوید. یافتن تناسب درمانی ایده‌آل شما آسان‌تر از همیشه خواهد بود و شاید اصلاً نیازی به جستجوی زیاد نداشته باشید. اشتراک‌گذاری داده‌ها در زمینه‌های پزشکی یک مسئله حساس و بحث‌برانگیز است، اما می‌تواند به طور قابل توجهی شخصی‌سازی درمان را بهبود بخشد. یک پلتفرم سلامت روان مبتنی بر هوش مصنوعی می‌تواند داده‌های هزاران بیمار را که در محدوده سنی، جنسیت و علائم اضطراب خاص شما هستند، تجزیه و تحلیل کند.

چنین پلتفرمی می‌داند که کدام تمرینات رفتار درمانی شناختی یا رژیم‌های دارویی منجر به قابل توجه‌ترین بهبودها برای افراد مشابه شما شده است. این پلتفرم می‌تواند رویکردهایی را که از آستانه‌های خاصی از کارایی در سناریوهای واقعی فراتر رفته‌اند، اولویت‌بندی کند و برنامه‌های

درمانی شخصی‌سازی شده را توصیه کند. حتی می‌تواند گزینه‌های مختلفی را به شما ارائه دهد، نرخ موفقیت مربوطه آن‌ها را با شما به اشتراک بگذارد و به شما اجازه دهد تا موردی را که بیشتر برای شما جذاب است، انتخاب کنید. شاید هر متخصصی در این پلتفرم، اعم از انسانی و هوش مصنوعی، دارای نظرات کاربران در پروفایل خود باشد، درست مانند آمازون. در نهایت، این پلتفرم ممکن است یک «ترکیب درمانی» الهام گرفته از اسپاتیفای ارائه دهد که ترکیبی منحصربه‌فرد از رویکردهای درمانی مختلف را بر اساس سابقه تعامل شما، انتخاب می‌کند.

آیا همه اینها مراقبت‌های روان‌درمانی را بی‌اهمیت می‌کند؟ تنها اگر فکر کنیم که اگر مراقبت‌های سلامت روان دشوارتر، کمتر مبتنی بر تحلیل داده‌ها و کمتر شخصی‌سازی شده از محتوای نتفلیکس باشد، برای جامعه بهتر است. آیا می‌تواند چالش‌های اخلاقی جدیدی را معرفی کند؟ ممکن است، و به همین دلیل شفافیت، استقرار تکراری و ارزیابی دقیق هر تلاش برای استقرار در این حوزه بسیار ضروری است. حتی اگر همه اینها را مدیریت کنیم، همچنان می‌توان استدلال کرد که «اتحاد درمانی» که بین بیماران و درمانگرانشان وجود دارد – حس همدلی و اعتمادی که با گذشت زمان، از طریق تعامل مداوم توسعه می‌یابد – چیزی است که نباید آن را به طور کامل خودکار کرد، مهم نیست که مدل‌های زبان بزرگ درمانی چقدر قابل اعتماد، محرمانه و از نظر بالینی تأیید شده شوند. برخی ممکن است استدلال کنند که با انتخاب این مسیر، بخشی بنیادی از انسانیت خود را از دست خواهیم داد. اما این تفکر مشکل‌گرایی است، که تنها بر یک طرف معادله تمرکز دارد. بیایید طرف دیگر را در نظر بگیریم: سیستم‌های هوش مصنوعی که می‌توانند متخصصان انسانی بیشتری را آموزش دهند تا آنچه اکنون قادر به آموزش آن هستیم. سیستم‌های هوش مصنوعی که می‌توانند از متخصصان انسانی به گونه‌ای حمایت کنند که به آن‌ها امکان دهد با بیماران بیشتری نسبت به آنچه در حال حاضر می‌توانند مدیریت کنند، تعامل داشته باشند. سیستم‌های هوش مصنوعی که میلیون‌ها نفر را که توسط رویکردهای کنونی کمتر خدمت‌رسانی می‌شوند، قادر می‌سازند تا مراقبت‌های فراوان و مقرون به صرفه دریافت کنند. یا آیا فکر می‌کنیم که ادامه محرومیت افراد کمتر خدمت‌رسانی شده در

تعقیب حفظ انسانیت ما مناسب‌تر است؟

۷ .فراتر از وضعیت موجود: سلامت روان پیشگیرانه و «فوق‌بشری» شدن

یکی دیگر از کلیشه‌های مشکل‌گرایی این ایده است که مراقبتی که بیش از حد دسترس‌پذیر و بیش از حد مقرون‌به‌صرفه می‌شود، می‌تواند منجر به وابستگی بیش از حد به مدل‌های زبان بزرگ درمانی و کاهش عاملیت فردی شود. اما این نگرانی عمدتاً بر اساس نحوه تعریف ما از دامنه و سودمندی مراقبت‌های سلامت روان است—یا به طور خاص‌تر، بر اساس نحوه تعریف سنتی ما از آن. تا حدی به دلیل پرهزینه بودن ارائه مراقبت توسط انسان، ما آن را عمدتاً به عنوان یک مکانیسم واکنشی برای پاسخ به مشکلات حاد می‌بینیم. اگر این خط پایه باشد، پس تفسیر استفاده مکرر به عنوان وابستگی بیش از حد آسان است. اما چند بار شنیده‌اید که مردم در مورد اعتیاد به عینک، یا وابستگی مشکل‌ساز به ضربان‌ساز یا کمربند ایمنی ابراز نگرانی می‌کنند؟

چه می‌شود اگر به چشم‌انداز مراقبت سلامت روان پیشگیرانه و مدل‌های هوش مصنوعی که حمایت عاطفی را به صورت دائمی ارائه می‌دهند، نگاه کنیم؟ وقتی به ذهنیت «چه اتفاق خوبی ممکن است رخ دهد؟» روی می‌آوریم، مزایا شروع به انباشت می‌کنند. مردم زمانی که بیشتر نیاز به دیده شدن دارند، دیده می‌شوند. آن‌ها در هر زمان، شب یا روز، به پاسخ‌های همدلانه و استراتژی‌های مفید دسترسی پیدا می‌کنند. آن‌ها مکانیسم‌های مقابله‌ای سالم‌تر و انعطاف‌پذیری عاطفی را از طریق حمایت مداوم و شخصی‌سازی شده توسعه می‌دهند. این چشم‌انداز جاه‌طلبانه صرفاً با هدف تکرار و مقیاس‌پذیری عملکردهای درمانی فعلی نیست. در عوض، هدف آن تغییر و ارتقای آن‌ها است، که به طور بالقوه می‌تواند عصر جدیدی از مراقبت سلامت روان را رقم بزند که جامع‌تر، مستمرتر و عمیقاً در زندگی روزمره ادغام شده است.

هنگامی که نوآوری‌های جدید به قدری قدرتمند می‌شوند که شروع به تغییر هنجارهای اجتماعی و رفتارهای فردی می‌کنند، بدبین‌ها اغلب چنین تغییراتی را ذاتاً مخرب می‌دانند. قبل از اینکه بیشتر مردم حتی بدانند «WWW» به چه معناست، کتاب‌هایی در مورد «اعتیاد به

اینترنت» وجود داشت. چرا؟ زیرا ایده میلیون‌ها نفر که ناگهان ساعت‌ها را صرف تایپ کردن با غریبه‌های ناشناس می‌کنند، با قراردادهای موجود در مورد تعامل اجتماعی سالم و مدیریت زمان بهره‌ور سازگار نبود. با افزایش تعامل مردم با مدل‌های زبان بزرگ از هر نوع، از جمله مدل‌های درمانی، ارزش آن را دارد که به یاد داشته باشیم یکی از پایدارترین رفتارهای انسانی ما شامل ایجاد پیوندهای فوق‌العاده نزدیک و مهم با هوش‌های غیرانسانی است. میلیاردها نفر می‌گویند با خدا یا دیگر خدایان مذهبی رابطه شخصی دارند، که اکثر آن‌ها به عنوان هوش‌های برتر تصور می‌شوند که قدرت درک و عادات ذهنی آن‌ها به طور کامل برای ما فانیان قابل تشخیص نیست. میلیاردها نفر برخی از معنادارترین روابط خود را با سگ‌ها، گربه‌ها و سایر حیوانات ایجاد می‌کنند که دامنه قدرت‌های ارتباطی نسبتاً محدودی دارند. کودکان این کار را با عروسک‌ها، حیوانات عروسکی و دوستان خیالی انجام می‌دهند. اینکه ما ممکن است به سرعت پیوندهای عمیق و ماندگاری با هوش‌هایی ایجاد کنیم که به اندازه ما گویا و پاسخگو هستند، اجتناب‌ناپذیر به نظر می‌رسد، نشانه‌ای از طبیعت انسان بیش از تجاوز تکنولوژیکی.

در بیشتر موارد، ما توانایی خود را برای پرورش روابط با موجودات غیرانسانی یکی از ارزشمندترین ویژگی‌های خود می‌دانیم، زیرا این روابط می‌توانند هوش هیجانی ما را افزایش داده و روابط ما با افراد دیگر را تکمیل و تحت تأثیر قرار دهند. بسیاری از این روابط به عنوان منبع حمایت عاطفی بدون قضاوت عمل می‌کنند. آن‌ها به ایجاد محیط‌هایی کمک می‌کنند که در آن افراد به اندازه کافی احساس راحتی می‌کنند تا خود را صریحاً ابراز کنند. آن‌ها اغلب حس هدفمندی را فراهم می‌کنند و به رفاه کلی کمک می‌کنند.

در طول سال ۲۰۲۳، همانطور که میلیون‌ها نفر برای اولین بار با مدل‌های زبان بزرگ تعامل داشتند، یک عبارت تکراری در گفتمان، از منتقدان، حامیان، و خود مدل‌ها (اگر از آن‌ها می‌پرسیدید)، این بود که مدل‌ها درک واقعی از جهان ندارند و بنابراین هوش هیجانی واقعی و بینش یا درک واقعی از احساسات یا حالات ذهنی کسانی که با آن‌ها تعامل دارند، ندارند. به طور خلاصه، همدلی ندارند. و این یکی از دلایلی بود که «نگه داشتن انسان‌ها در چرخه» بسیار مهم

تلقی می‌شد.

با این حال، به سرعت بسیاری از مردم کشف کردند که مدل‌های هوش مصنوعی مکالمه‌گرا چقدر می‌توانند پاسخگو، حاضر، صبور و به ظاهر از نظر عاطفی هماهنگ باشند. یک مطالعه از آوریل ۲۰۲۳ که در JAMA Internal Medicine منتشر شد، مثال گویایی است. نویسندگان این مطالعه به طور تصادفی ۱۹۵ مکالمه را از یک ساب‌ردیت Reddit به نام AskDocs انتخاب کردند، جایی که کاربران Reddit سؤال می‌پرسند و پزشکان تأیید شده به آن‌ها پاسخ می‌دهند. این مطالعه از ChatGPT نیز خواست تا به آن‌ها پاسخ دهد، سپس از هیئت‌هایی متشکل از سه پزشک دارای مجوز (نه روان‌درمانگر) خواست تا دو پاسخ مختلف را برای هر یک از ۱۹۵ مکالمه – یکی از ChatGPT و یکی از ارائه‌دهنده مراقبت انسانی – بدون شناسایی نویسنده، رتبه‌بندی کنند. در ۷۸.۶ درصد از موارد، پزشکان پاسخ‌های ChatGPT را بالاتر از پاسخ‌های انسانی رتبه‌بندی کردند.

در واقع، در حالی که مدل‌های هوش مصنوعی صراحتاً آگاه یا خودآگاه نیستند، آن‌ها به شیوه احتمالاتی آماری خود، به طور نمایشی مهربان و همدلانه هستند، به روش‌هایی که اغلب از هنجارهای انسانی فراتر می‌روند. پیامد بالقوه این موضوع برای بسیاری افراد در مکالمات با محققان روشن شده است. محققان برجسته در زمینه هوش مصنوعی و رباتیک اجتماعی تأکید کرده‌اند که چگونه انواع مختلف هوش مصنوعی که همدلی را شبیه‌سازی می‌کنند، می‌توانند تأثیر عمیقی بر بشریت در کل داشته باشند. همانطور که برخی پیشنهاد داده‌اند، همه به طور قابل اعتماد به مهربانی و حمایت انسانی دسترسی ندارند. اما وقتی این موضوع به چیزی تبدیل می‌شود که شما «همیشه در دسترس» دارید، ظرفیت خود شما برای «توانایی مهربان بودن با دیگران» را افزایش می‌دهد. مطالعات متعدد نشان داده‌اند که رفتار همدلانه سلامت جسمانی را بهبود می‌بخشد.

پس تصور کنید اثرات ترکیبی که احتمالاً با گذشت زمان، با ادغام کامل‌تر مدل‌های هوش

مصنوعی در زندگی مردم، مشاهده خواهیم کرد. علاوه بر کمک به ما برای کشف اسرار تاخوردگی پروتئین، یا افزایش کارایی سیستم‌های انرژی تجدیدپذیر، چه می‌شود اگر تعاملات مداوم ما با مدل‌های هوش مصنوعی که از آرامش پیش‌بینی‌ناپذیر یک دالایی لاما برخوردارند، به سادگی به ما کمک کند تا نسخه‌های مهربان‌تر، صبورتر و از نظر عاطفی سخاوتمندتر از خودمان شویم؟ از طریق شبکه‌های عصبی پیچیده و خوشه‌های سرور جهانی عظیم، جهان می‌تواند «فوق‌بشری» شود.

نتیجه‌گیری

این فصل به ما نشان داد که مسیر توسعه هوش مصنوعی با دو دیدگاه متضاد «راه‌حل‌گرایی» و «مشکل‌گرایی» همراه است. در حالی که راه‌حل‌گرایی ممکن است در برخی موارد به ساده‌سازی مسائل بپردازد، مشکل‌گرایی با تمرکز صرف بر خطرات و احتمالات منفی، می‌تواند مانع نوآوری و پیشرفت‌های عظیم اجتماعی شود. هوش مصنوعی به عنوان یک اهرم قدرتمند برای ایجاد تغییرات مثبت در مقیاس وسیع، پتانسیل چشمگیری برای حل چالش‌های جهانی از جمله در حوزه سلامت روان دارد. بحران جاری سلامت روان، با کمبود متخصصان، دسترسی محدود و هزینه‌های گزاف، نمونه‌ای بارز از «تهدید وجودی وضعیت موجود» است که فناوری می‌تواند به آن پاسخ دهد.

مطالعه موردی پلتفرم کوکو نشان داد که چگونه ذهنیت مشکل‌گرا می‌تواند بدون درک کامل از شفافیت و نوآوری یک پروژه، به سرعت به قضاوت‌های نادرست منجر شود. در مقابل، رویکرد کوکو با استفاده از هوش مصنوعی برای گسترش دسترسی به پشتیبانی مبتنی بر شواهد در سلامت روان، مثالی از تلاش برای غلبه بر این تهدید وجودی بود. تحقیقات پیشرفته‌ای که در این فصل بررسی شد، به ویژه مطالعه تجزیه و تحلیل جلسات درمانی Talkspace توسط Lyssn.io، بر پتانسیل هوش مصنوعی در درک ظرافت‌های درمان مؤثر و ارائه بینش‌های بی‌سابقه برای بهبود عملکردهای بالینی تأکید کرد.

این فصل همچنین چشم‌انداز آینده‌ای را ترسیم کرد که در آن مراقبت‌های سلامت روان، شبیه به پلتفرم‌های رسانه‌ای دیجیتال، به طور فراوان، مقرون‌به‌صرفه، دسترس‌پذیر و فوق‌العاده شخصی‌سازی شده ارائه می‌شوند. این مدل می‌تواند محدودیت‌های فعلی را که ناشی از اتکای صرف به متخصصان انسانی است، از بین ببرد و به کاربران امکان دهد تا رویکردهای درمانی مختلف را کاوش کرده و تیم‌های درمانی شخصی‌سازی شده خود را تشکیل دهند. نگرانی‌ها در مورد وابستگی بیش از حد یا بی‌اهمیت شدن مراقبت‌ها، در صورت اتخاذ یک دیدگاه پیشگیرانه و متمرکز بر «چه اتفاق خوبی ممکن است رخ دهد؟»، می‌تواند برطرف شود.

در نهایت، این فصل استدلال کرد که تعامل فزاینده با هوش مصنوعی می‌تواند به فراتر رفتن از مفهوم «انسانیت» به سمت «فوق‌بشری» شدن منجر شود. با توجه به تمایل ذاتی انسان برای ایجاد پیوندهای عمیق با هوش‌های غیرانسانی و توانایی مدل‌های هوش مصنوعی در نمایش همدلی و مهربانی به شیوه‌هایی که اغلب از هنجارهای انسانی فراتر می‌رود، این فناوری‌ها نه تنها می‌توانند مشکلات را حل کنند، بلکه ظرفیت‌های انسانی برای شفقت، صبر و سخاوت عاطفی را نیز تقویت کرده و به ایجاد جامعه‌ای مهربان‌تر و انسانی‌تر کمک کنند. پیشرفت مستلزم پذیرش ریسک‌های حساب شده و تمرکز بر پتانسیل‌های بی‌نظیر هوش مصنوعی برای آینده‌ای بهتر است.

نکات کلیدی

- **راه‌حل‌گرایی** اعتقاد دارد که مشکلات پیچیده اجتماعی دارای راه‌حل‌های فناورانه ساده‌ای هستند.
- **مشکل‌گرایی** فناوری را به عنوان یک نیروی ذاتاً مشکوک و ضدبشری می‌بیند که می‌تواند مانع نوآوری شود.
- افراد بدبین می‌توانند نقش مهمی در **پاسخگو کردن شرکت‌های بزرگ فناوری** داشته باشند، اما تمرکز صرف بر نقد می‌تواند آسیب‌زا باشد.

- **هوش مصنوعی یک اهرم اثبات شده برای تغییرات مثبت در مقیاس وسیع است** که می‌تواند ۳۰ تا ۸۰ درصد از راه‌حل‌های مؤثر برای چالش‌های جهانی را تشکیل دهد.
- **تهدید وجودی وضعیت موجود** به ناکارآمدی‌ها و نابرابری‌های طولانی‌مدتی اشاره دارد که کمتر مورد توجه قرار می‌گیرند، مانند بحران سلامت روان.
- **بحران سلامت روان** با افزایش نرخ افسردگی، اضطراب و خودکشی، کمبود شدید متخصصان و موانع دسترسی مشخص می‌شود.
- **فناوری‌های موجود سلامت روان** مانند تله‌تراپی و برنامه‌های خودراهنما با چالش‌های جذب و تعامل پایین روبرو هستند.
- **چت‌بات‌های تخصصی سلامت روان** مانند Woebot وWysa ، اگرچه مفیدند، اما به دلیل اتکا به «فریم‌ها» و ساختارهای از پیش تعریف‌شده، در پاسخگویی به ورودی‌های ظریف کاربر محدود هستند.
- **ماجرای کوکو** نشان داد که چگونه برداشت‌های غلط و عدم شفافیت اولیه (از جانب نویسنده توییت) می‌تواند به واکنش‌های مشکل‌گرایانه شدید منجر شود، حتی در صورت عدم وجود آسیب واقعی به کاربران.
- **عملیات کوکو شفاف بود** و مشارکت هوش مصنوعی به کاربران اطلاع‌رسانی می‌شد، و این پلتفرم در حال مبارزه با تهدید وجودی نیازهای برآورده‌نشده سلامت روان بود.
- **هوش مصنوعی می‌تواند داده‌های درمانی را تجزیه و تحلیل کند** تا اثربخشی روش‌های مبتنی بر شواهد و رفتارهای درمانگر را شناسایی کند) مثال Lyssn.io و Talkspace).
- یافته‌ها نشان داد که «ارائه اطلاعات» بیش از حد می‌تواند به نتایج درمانی بدتر منجر شود، در حالی که «بازتاب‌های پیچیده» و «تأییدات» با نتایج بهتر مرتبط بودند.

- **مراحل سه‌گانه ادغام هوش مصنوعی در سلامت روان**)بر اساس Stade و همکاران :(۱ .استفاده‌های کمکی ساده، ۲. تعاملات مشارکتی، ۳. مراقبت کاملاً خودمختار.
- **چشم‌انداز «سلامت روان به سبک اسپاتیفای و نتفلیکس** «به معنای دسترسی فراوان، مقرون‌به‌صرفه، مقیاس‌پذیر و شخصی‌سازی شده به مراقبت‌ها است.
- هوش مصنوعی می‌تواند به **شخصی‌سازی درمان‌ها** بر اساس داده‌های هزاران بیمار مشابه کمک کند و «ترکیب درمانی» منحصر به فردی را ارائه دهد.
- نگرانی‌ها در مورد **وابستگی بیش از حد به هوش مصنوعی** می‌تواند با تغییر دیدگاه به مراقبت سلامت روان پیشگیرانه و مستمر برطرف شود.
- انسان‌ها تمایل ذاتی به ایجاد پیوندهای عمیق با هوش‌های غیرانسانی دارند، مانند خدایان، حیوانات خانگی و دوستان خیالی.
- مدل‌های هوش مصنوعی می‌توانند همدلی و مهربانی را به صورت نمایشی ارائه دهند که گاهی اوقات از هنجارهای انسانی فراتر می‌رود) مطالعه JAMA Internal Medicine).
- تعامل با هوش مصنوعی می‌تواند ظرفیت انسانی برای مهربانی، صبر و همدلی را افزایش دهد و به جهانی «فوق‌بشری» منجر شود.
- برای پیشرفت در هوش مصنوعی، پذیرش سطح معینی از ریسک و عدم قطعیت ضروری است، در حالی که شفافیت، ارزیابی دقیق و استقرار تکراری حفظ شود.

سؤالات تفکربرانگیز

۱. چگونه می‌توانیم یک تعادل سالم بین احتیاط لازم در برابر خطرات هوش مصنوعی و تشویق نوآوری برای بهره‌برداری از پتانسیل‌های مثبت آن ایجاد کنیم؟

۲. با توجه به «تهدید وجودی وضعیت موجود» در حوزه‌هایی مانند سلامت روان، آیا تأخیر در توسعه و به‌کارگیری راه‌حل‌های مبتنی بر هوش مصنوعی از نظر اخلاقی قابل

دفاع است؟

۳. چه تدابیر و سیاست‌هایی باید اتخاذ شود تا اطمینان حاصل شود که هوش مصنوعی در سلامت روان به شکلی اخلاقی، شفاف و عادلانه برای همه جوامع، به ویژه جمعیت‌های آسیب‌پذیر، ارائه شود؟

۴. چگونه می‌توانیم اعتماد عمومی را به سیستم‌های هوش مصنوعی در حوزه‌های حساس مانند سلامت روان جلب کنیم، با توجه به واکنش‌های منفی احتمالی مانند آنچه در مورد پلتفرم کوکو رخ داد؟

۵. چه ابعادی از «اتحاد درمانی» بین بیمار و درمانگر انسانی وجود دارد که هرگز نباید به طور کامل خودکار شود و چگونه می‌توانیم این ابعاد را در کنار پیشرفت‌های هوش مصنوعی حفظ کنیم؟

۶. اگر هوش مصنوعی بتواند همدلی و مهربانی را به شیوه‌ای «فوق‌بشری» ارائه دهد، این چه تأثیری بر درک ما از روابط انسانی و توسعه هوش هیجانی در جامعه خواهد داشت؟

۷. چگونه می‌توانیم مدل «سلامت روان به سبک اسپاتیفای و نتفلیکس» را بدون بی‌اهمیت جلوه دادن عمق و پیچیدگی مراقبت‌های روان‌درمانی پیاده‌سازی کنیم؟

۸. آیا پذیرش رویکرد «چه اتفاق خوبی ممکن است رخ دهد؟» به معنای نادیده گرفتن مسئولیت‌های اخلاقی توسعه‌دهندگان و سیاست‌گذاران در قبال پیامدهای منفی بالقوه است؟ چرا یا چرا نه؟

۳۰ جمله کلیدی

۱. راه‌حل‌گرایی باوری است که به راه‌حل‌های فناورانه برای چالش‌های عمیق اجتماعی معتقد است.

۲. مشکل‌گرایی فناوری را نیرویی ذاتاً مشکوک و ضدبشری می‌داند که آلوده به منطق

سرمایه‌داری است.

۳. در حالی که افراد بدبین می‌توانند ارزشمند باشند، تمرکز صرف بر نقد می‌تواند مانع نوآوری و ایجاد تغییر مثبت شود.

۴. فناوری یکی از مؤثرترین اهرم‌های بشریت برای ایجاد تغییرات مثبت در مقیاس وسیع است.

۵. هوش مصنوعی می‌تواند ۳۰ تا ۸۰ درصد از هر راه‌حل مؤثر برای چالش‌های جهانی را تشکیل دهد.

۶. تهدید وجودی وضعیت موجود به ناکارآمدی‌ها و نابرابری‌های طولانی‌مدت اشاره دارد که اغلب نادیده گرفته می‌شوند.

۷. بحران سلامت روان جهانی با افزایش نرخ افسردگی، اضطراب، خودکشی و کمبود شدید متخصصان مشخص می‌شود.

۸. دسترسی به مراقبت‌های سلامت روان برای میلیون‌ها نفر در سراسر جهان محدود یا غیرممکن است.

۹. فناوری‌های موجود سلامت روان، مانند برنامه‌های خودراهنما، اغلب با نرخ پایین تعامل و ترک استفاده مواجه هستند.

۱۰. چت‌بات‌های تخصصی سلامت روان به دلیل ساختارهای از پیش تعریف‌شده، در ارائه پاسخ‌های ظریف و شخصی‌سازی شده محدودیت دارند.

۱۱. پلتفرم کوکو (Koko) از هوش مصنوعی برای گسترش دسترسی به پشتیبانی سلامت روان مبتنی بر همتا استفاده می‌کند.

۱۲. ماجرای کوکو نشان داد که چگونه واکنش‌های مشکل‌گرایانه بدون اطلاعات کافی می‌تواند به قضاوت‌های نادرست منجر شود.

۱۳. کوکو مشارکت هوش مصنوعی در پیام‌ها را به طور شفاف به کاربران اطلاع‌رسانی

می‌کرد.

۱۴. هوش مصنوعی می‌تواند مجموعه‌های عظیمی از داده‌های درمانی را برای شناسایی مداخلات مؤثر و رفتارهای درمانگر تجزیه و تحلیل کند.

۱۵. مطالعه Lyssn.io و Talkspace نشان داد که ارائه بیش از حد اطلاعات می‌تواند به نتایج درمانی بدتر منجر شود.

۱۶. بازتاب‌های پیچیده و تأییدات توسط درمانگر با نتایج مثبت‌تر درمانی مرتبط هستند.

۱۷. مدل‌های زبان بزرگ (LLMs) می‌توانند به عنوان دستیاران در کارهای اداری و آموزشی برای متخصصان سلامت روان عمل کنند.

۱۸. چشم‌انداز آینده شامل مراقبت‌های سلامت روان کاملاً خودمختار و مبتنی بر هوش مصنوعی است.

۱۹. هوش مصنوعی می‌تواند به مراقبت‌های سلامت روان «فراوان، اقتصادی، مقیاس‌پذیر و شخصی‌سازی شده» منجر شود، مانند اسپاتیفای و نتفلیکس.

۲۰. دسترسی به درمانگر مجازی در هر زمان و مکان می‌تواند موانع سنتی انتخاب و دسترسی به مراقبت را از بین ببرد.

۲۱. پلتفرم‌های مبتنی بر هوش مصنوعی می‌توانند برنامه‌های درمانی شخصی‌سازی شده را بر اساس داده‌های بیماران مشابه توصیه کنند.

۲۲. نگرانی‌ها در مورد وابستگی بیش از حد به هوش مصنوعی باید در چارچوب تغییر به سوی مراقبت سلامت روان پیشگیرانه بررسی شود.

۲۳. رویکرد «چه اتفاق خوبی ممکن است رخ دهد؟» به معنای تصور آینده مطلوب و حرکت فعال به سوی آن است.

۲۴. انسان‌ها به طور طبیعی تمایل دارند با هوش‌های غیرانسانی مانند خدایان، حیوانات

خانگی و دوستان خیالی پیوندهای عمیق برقرار کنند.

۲۵. مدل‌های هوش مصنوعی مکالمه‌گرا می‌توانند پاسخ‌های همدلانه و مهربانانه را به شیوه‌ای نمایشی ارائه دهند.

۲۶. مطالعه AskDocs نشان داد که پاسخ‌های ChatGPT در ۷۸.۶ درصد موارد از پاسخ‌های پزشکان انسانی بهتر ارزیابی شد.

۲۷. هوش مصنوعی می‌تواند مهربانی و حمایت را «همیشه در دسترس» قرار دهد و ظرفیت انسانی برای مهربانی را افزایش دهد.

۲۸. تعامل با هوش مصنوعی می‌تواند منجر به افزایش صبر، سخاوت عاطفی و در نهایت به جامعه‌ای «فوق‌بشری» منجر شود.

۲۹. نوآوری نیازمند پذیرش سطح معینی از ریسک و عدم قطعیت است تا پیشرفت محقق شود.

۳۰. برای اطمینان از توسعه مسئولانه هوش مصنوعی، شفافیت، استقرار تکراری و ارزیابی دقیق ضروری است.

فصل ۴

پیروزی مشاعات خصوصی: خلق ارزش و پتانسیل‌های بی‌نهایت در عصر هوش مصنوعی

»ویکی‌پدیا، به عنوان یک سرویس رایگان، مقالات بسیار بیشتری با کیفیتی قابل مقایسه با بریتانیکا ارائه می‌دهد که زمانی هزاران دلار هزینه داشت.» – اریک براینجلفسون و آویناش کولیس

اهداف اصلی فصل:

- معرفی مفهوم «مشاعات خصوصی» و جایگاه آن در تحولات دیجیتال کنونی.
- بررسی انتقادهای مطرح شده نسبت به مدل‌های کسب‌وکار شرکت‌های بزرگ فناوری و مقایسه آن‌ها با مزایای واقعی.
- تحلیل پتانسیل‌های بی‌نظیر مشاعات خصوصی در خلق ارزش اقتصادی و اجتماعی گسترده.
- تبیین نقش محوری هوش مصنوعی و مدل‌های چندوجهی در توسعه و افزایش بهره‌وری این مشاعات.
- تشریح تفاوت‌های اساسی مشاعات دیجیتال با مفهوم سنتی «تراژدی مشاعات» و پیامدهای آن.

چکیده

در دهـه‌های اخیـر، ظهـور شـرکت‌های بـزرگ فنـاوری و پیشـرفت‌های خیره‌کننـده در هـوش مصـنوعی، بحث‌هـای گسـترده‌ای را در مـورد توزیـع ارزش اقتصـادی و تـأثیرات اجتمـاعی آن‌هـا برانگیختـه اسـت. برخـی منتقـدان، ایـن شـرکت‌ها را بـه تمرکـز ثـروت، حـذف مشـاغل، و حتـی شـکل‌گیری نـوعی «سـرمایه‌داری نظـارتی» مـتهم می‌کننـد کـه در آن حـریم خصوصـی و اراده فـردی تحت‌الشعاع قـرار می‌گیـرد. بـا ایـن حـال، نگـاهی عمیق‌تر بـه پدیـده‌ی «مشـاعات خصوصـی» نشـان می‌دهـد کـه ایـن روایـت، اغلـب بخـش عظیمـی از ارزش خلـق شـده بـرای کاربـران و جامعـه را نادیـده می‌گیـرد. مشـاعات خصوصـی، پلتفرم‌هـای دیجیتالـی هسـتند کـه تحـت مالکیـت یـا اداره خصوصـی قـرار دارنـد امـا کاربـران را بـه عنـوان تولیدکننـده و ناظـر مشـارکت می‌دهنـد و خدمـاتی رایـگان یـا تقریبـاً رایـگان ارائـه می‌دهنـد کـه بـه مثابـه زیرسـاخت‌های اجتمـاعی و خـدماتی خصوصی‌سـازی شـده عمـل می‌کننـد. ایـن فصـل بـه بررسـی چگونگـی خلـق ارزش گسـترده توسـط ایـن مشـاعات، مفهـوم «مـازاد مصـرف‌کننده» در اقتصـاد دیجیتـال، و نقـش حیـاتی داده‌هـای غیررقـابتی می‌پـردازد. همچنـین، تأثیرات تحول‌آفرین هوش مصنوعی، به ویژه مدل‌های چندوجهی، بر افزایش بی‌سابقه قابلیت‌ها و ارزش مشاعات خصوصی را مورد کاوش قرار می‌دهد و نشان می‌دهد که چگونه این هم‌زیستی می‌تواند به عاملیت فردی، فرصت‌های آموزشی، تحرک اجتماعی و رشد حرفه‌ای بی‌حد و حصر منجر شود.

مقدمه

در دنیای مدرن، تحولات فناورانه، به ویژه در حوزه‌های هوش مصنوعی و ابزارهای دیجیتال، با سرعت سرسام‌آوری در حال پیشرفت است. این پیشرفت‌ها، در کنار مزایای چشمگیر، سوالات بنیادینی را نیز در مورد ماهیت توزیع ثروت و قدرت در اقتصاد جدید مطرح کرده‌اند. مشاهده‌ی رشد بی‌سابقه‌ی شرکت‌های بزرگ فناوری و انباشت ثروت عظیم در دستان تعداد معدودی از سرمایه‌گذاران و کارآفرینان، بسیاری را به این باور سوق داده است که این نوآوری‌ها به جای ارتقای فراگیر رفاه عمومی، صرفاً به نفع بخش کوچکی از جامعه عمل می‌کنند. این دیدگاه، که اغلب در تحلیل‌های انتقادی از «تک‌غول‌های فناوری (Big Tech)» برجسته می‌شود، بر این نکته تأکید دارد که حتی با وجود کنترل مؤثر کاربران بر داده‌های شخصی و تعادل میان مدیریت چالش‌های سلامت روان و حفظ زندگی عاطفی، همچنان بخش عمده‌ای از پاداش‌ها به خود این شرکت‌ها می‌رسد.

این نگرانی‌ها، از سوی نشریات معتبر و نویسندگان برجسته نیز مطرح شده است. به عنوان مثال، برخی تحلیلگران معتقدند که با وجود پیشرفت‌های شگفت‌انگیز در هوش مصنوعی، نتایج آن در بهبود رفاه و تحریک رشد اقتصادی گسترده ناامیدکننده بوده است. طبق این دیدگاه، در حالی که معدود سرمایه‌گذاران و کارآفرینان به ثروت‌های کلان دست یافته‌اند، اکثر مردم بهره‌ای نبرده‌اند و حتی برخی به دلیل اتوماسیون، شغل خود را از دست داده‌اند. تد چیانگ، نویسنده داستان‌های علمی-تخیلی، نیز با اشاره به آمار تولید ناخالص داخلی سرانه در ایالات متحده که از سال ۱۹۸۰ تقریباً دو برابر شده، در حالی که درآمد متوسط خانوارها بسیار عقب مانده است، استدلال می‌کند که ارزش اقتصادی ایجاد شده توسط کامپیوترهای شخصی و اینترنت، عمدتاً به افزایش ثروت اقشار بسیار ثروتمند منجر شده و نه به ارتقای استاندارد زندگی عموم شهروندان.

شدیدترین انتقادها در این زمینه، از سوی شخصیت‌هایی مانند شوشانا زوبوف، استاد بازنشسته دانشکده بازرگانی هاروارد، مطرح شده است. زوبوف در کتاب پرفروش خود «عصر سرمایه‌داری نظارتی»، با الهام از آثار جورج اورول، استدلال می‌کند که شرکت‌هایی نظیر گوگل و

فیس‌بوک، با ایجاد یک «زیرساخت حسی، شبکه‌ای و محاسباتی» که او آن را «دیگری بزرگ » (Big Other)می‌نامد، جایگزین حزب و برادر بزرگ شده‌اند. از نظر او، فناوری به جای اینکه دولت را به «پروژه‌ای برای تملک کامل» تبدیل کند، بازار را به «پروژه‌ای برای قطعیت کامل» مسلح کرده است.

با این حال، این فصل قصد دارد روایتی متفاوت و مکمل را ارائه دهد. در حالی که نگرانی‌ها در مورد تمرکز قدرت و داده‌های شخصی قابل درک و مهم هستند، تمرکز صرف بر این جنبه‌ها، از درک عمیق‌تر پدیده‌ای به نام «مشاعات خصوصی» بازمی‌ماند. این مشاعات، که در بطن اقتصاد دیجیتال شکل گرفته‌اند، نه تنها منبع خلق ارزش بی‌سابقه برای میلیاردها کاربر در سراسر جهان هستند، بلکه پتانسیل‌های عظیمی برای تحول اجتماعی و توانمندسازی فردی دارند. این فصل به تحلیل این پدیده می‌پردازد و نشان می‌دهد که چگونه مشاعات خصوصی، با بهره‌گیری از داده‌های غیررقابتی و هوش مصنوعی، می‌توانند به ابزاری قدرتمند برای افزایش رفاه و فرصت‌های جمعی تبدیل شوند، به شرط آنکه رویکردی باز و مشارکتی نسبت به داده‌ها اتخاذ شود و چارچوب‌های مستحکمی برای حفاظت از حریم خصوصی و استفاده اخلاقی از داده‌ها توسعه یابد. هدف این فصل، ارائه درکی جامع از این چشم‌انداز متحول‌کننده و تأکید بر پتانسیل‌های بی‌کران مشاعات خصوصی در عصر هوش مصنوعی است.

تردیدها پیرامون نوآوری‌های شرکت‌های بزرگ فناوری

انتقاد از شرکت‌های بزرگ فناوری، به ویژه در مورد هوش مصنوعی، رویکردی رایج و پررنگ در گفتمان عمومی است. همانطور که اشاره شد، دیدگاه‌های مطرح شده توسط نشریاتی مانند ام‌آی‌تی تکنولوژی ریویو و نویسندگانی چون تد چیانگ، بیانگر این نگرانی است که دستاوردهای عظیم فناورانه، به جای بهبود فراگیر رفاه اقتصادی، عمدتاً به تمرکز ثروت در دستان نخبگان منجر شده است. این منتقدان معتقدند که با وجود افزایش تولید ناخالص داخلی، زندگی اکثر مردم تغییری نکرده و حتی برخی با از دست دادن مشاغل خود به دلیل اتوماسیون، متضرر

شده‌اند. این نگاه، بنیان‌های اساسی بحث پیرامون توزیع عادلانه ارزش در اقتصاد دیجیتال را تشکیل می‌دهد.

این انتقادها، با رویکرد شوشانا زوبوف به اوج خود می‌رسد. او در «عصر سرمایه‌داری نظارتی» با استعاره‌ای قوی، «دیگری بزرگ» را جایگزین «برادر بزرگ» اورول می‌داند. از نظر او، شرکت‌هایی مانند گوگل و فیس‌بوک، از طریق یک «زیرساخت حسی، شبکه‌ای و محاسباتی»، بازار را به ابزاری برای «قطعیت کامل» تبدیل کرده‌اند. در این منظومه فکری، «دیگری بزرگ» از طریق نظارت گسترده و همه‌جا حاضر، عاملیت فردی انسان‌ها را ذره‌ذره تحلیل می‌برد. هر درخواست موقعیت مکانی، هر کلیک، و هر فعالیت آنلاین، به خوراک الگوریتم‌هایی تبدیل می‌شود که آزادی اراده ما را به تدریج تضعیف می‌کنند. زوبوف پیش‌بینی می‌کند که در این فرایند، دموکراسی به نوعی تمامیت‌خواهی بازارمحور تبدیل می‌شود؛ جایی که حتی تمایلات ساده ما برای یافتن مسیرهای بدون ترافیک یا پیتزاهای خوشمزه با امتیاز بالا، توانایی ما برای زندگی خودتعیین‌گر را از بین می‌برد.

به گفته زوبوف، این «جهنم» از لحظه‌ای آغاز شد که گوگل متوجه شد هر عمل کاربران در سایتش – رشته‌های جستجو، لینک‌های کلیک شده و غیره – قابل ردیابی است. این داده‌ها، که گاهی اوقات به عنوان «دود اگزوز داده» توصیف می‌شوند و به ظاهر بی‌ارزش به نظر می‌رسند، می‌توانستند ذخیره، تجمیع، تحلیل، ترکیب مجدد و در نهایت به روش‌های جدید و در مقیاسی وسیع به کار گرفته شوند. در ابتدا، گوگل از این داده‌ها برای بهبود تجربه کاربری استفاده می‌کرد. به عنوان مثال، اگر کاربران جستجوگر «بهترین پخش‌کننده «MP3 به طور مداوم روی لینک‌های «بررسی آی‌پاد» کلیک می‌کردند، گوگل الگوریتم‌های رتبه‌بندی خود را برای اولویت‌بندی مقایسه‌های محصول در جستجوهای آینده اصلاح می‌کرد. این فرآیند، که زوبوف آن را «چرخه بازسرمایه‌گذاری ارزش رفتاری» می‌نامد، در روزهای اولیه «عدالت» و «فضیلت» دوچندان داشت، زیرا تمام ارزش از تلاش‌های گوگل برای استفاده مولد از داده‌های رفتاری به کاربران باز می‌گشت. علاوه بر بهبود جستجو، گوگل از داده‌ها برای ایجاد محصولات و خدمات کاملاً جدید، مانند

نرم‌افزار ترجمه نیز بهره می‌برد.

اما زوبوف معتقد است که در جستجوی درآمد کافی برای پایداری عملیات خود، گوگل مرتکب «گناه اصلی سرمایه‌داری نظارتی» شد: استفاده از بخشی از داده‌های رفتاری جمع‌آوری شده برای مرتبط‌تر کردن تبلیغات به کاربران. این اقدام، به گفته گوگل، باعث می‌شد کاربران بیشتر روی تبلیغات کلیک کنند و تبلیغ‌کنندگان نیز ارزش بیشتری از این فرآیند به دست آورند. زوبوف این رویکرد را به معنای تمرکز گوگل بر مطابقت تبلیغات با جستجوها می‌داند، هرچند بعداً تعدیل می‌کند که «برخی داده‌ها همچنان برای بهبود خدمات به کار می‌روند، اما ذخایر رو به رشد سیگنال‌های جانبی برای بهبود سودآوری تبلیغات هم برای گوگل و هم برای تبلیغ‌کنندگان آن تغییر کاربری دادند.» در نهایت، در این روایت انتقادی، شما محصول نیستید؛ شما تنها یک «لاشه رها شده» هستید و محصول واقعی، «مازاد رفتاری» است که از زندگی شما استخراج می‌شود.

بازنگری در خلق ارزش: فراتر از استخراج

با وجود قوت انتقادهای مطرح شده، به خصوص از جانب زوبوف، باید توجه داشت که این ارزیابی‌ها ممکن است در برخی جنبه‌ها از واقعیت فاصله بگیرند. به عنوان مثال، اگرچه زوبوف تغییر کاربری داده‌ها برای تبلیغات را «گناه اصلی» می‌داند و آن را نشانه‌ای از گسست گوگل از «چرخه بازسرمایه‌گذاری ارزش رفتاری» تلقی می‌کند، اما واقعیت این است که محصولات مهمی مانند جیمیل، کروم، مپس، استریت ویو و داکس، همگی پس از نوآوری در فناوری تبلیغاتی گوگل عرضه شدند. این امر نشان می‌دهد که شرکت همچنان به سرمایه‌گذاری و توسعه خدمات جدید و ارزشمند برای کاربران ادامه داده است. برای توجیه نظریه‌ی خود، زوبوف ناچار است ادعا کند که هدف اصلی تمامی این محصولات، تولید «مازاد رفتاری» بیشتر است تا کسب‌وکار تبلیغاتی گوگل را به مرحله‌ای برساند که هدف‌گیری به دستکاری رفتار کاربران تبدیل شود.

با این حال، این دیدگاه نمی‌تواند دلیل ادامه استفاده میلیاردها نفر از این پلتفرم‌ها را توضیح

دهد. این واقعیت که بسیاری از ما روزانه ده‌ها بار از این خدمات استفاده می‌کنیم، نشان می‌دهد که ارزش ایجاد شده توسط شرکت‌های بزرگ فناوری، یک جریان دوطرفه است. شش محصول گوگل هر کدام بیش از دو میلیارد کاربر دارند و حدود ۱.۴۶ میلیارد نفر صاحب آیفون هستند. این آمار به وضوح نشان می‌دهد که شرکت‌های بزرگ فناوری، «تعاملات سازنده تولیدکننده-مصرف‌کننده» را برقرار می‌کنند.

استعاره «عملیات استخراج» که زوبوف برای توصیف شرکت‌های بزرگ فناوری و هوش مصنوعی به کار می‌برد، نیز در محیط دیجیتال دقیق نیست. اگرچه دانشمندان کامپیوتر از دهه ۱۹۵۰ از واژه «استخراج» برای توصیف جنبه‌هایی از بازیابی پایگاه داده استفاده کرده‌اند و خود زوبوف نیز این اصطلاح را از مقالات هل واریان از گوگل گرفته است، اما این کاربرد همیشه بیشتر استعاری بوده تا یک نظریه اطلاعاتی دقیق. داده‌ها مانند نفت، مس یا حتی یک دندان «استخراج» نمی‌شوند. استخراج زغال سنگ یک منبع محدود را به طور برگشت‌ناپذیری کاهش می‌دهد و حفره‌ای از خود به جای می‌گذارد. اما در مورد فایل‌های دیجیتال، تنها نسخه‌ها برداشته می‌شوند و اصل آن‌ها دست‌نخورده و بدون تغییر در مکان اصلی خود باقی می‌مانند. علاوه بر این، ما ذخایر داده‌های جهانی خود را با سرعتی بسیار بیشتر از آنچه یک تی‌رکس تجزیه شده به سوخت فسیلی تبدیل شود، افزایش می‌دهیم. هر ساعت، ما ذخایر از پیش تمام‌نشدنی سلفی‌ها، پست‌های ردیت، لایک‌های فیس‌بوک، اسلایدهای بازاریابی، جستجوهای گوگل، برنامه‌ریزی مسیر، فن‌فیک‌های بازی، مطالعات پزشکی و ویدئوهای یوتیوب را با محتوای جدیدی معادل یک نفتکش مجازی افزایش می‌دهیم.

با این حال، لازم به ذکر است که روش ایجاد بسیاری از مجموعه‌داده‌های هوش مصنوعی، خود یک موضوع بسیار بحث‌برانگیز است. برای ایجاد یک مدل زبان بزرگ (LLM) پیشرفته، به مقادیر عظیمی از داده‌های آموزشی نیاز است GPT-3 .شرکت اوپن‌ای‌آی با ۳۰۰ میلیارد توکن آموزش دیده و مجموعه داده آموزشی GPT-4 حتی بزرگتر بوده است. معمولاً توسعه‌دهندگان با

تکیه بر مخازن باز موجود که مقادیر زیادی از داده‌ها را از طریق خزش وب جمع‌آوری می‌کنند، شروع به گردآوری داده‌های آموزشی خود می‌کنند. مجموعه‌داده شناخته شده‌ای مانند کامن کراول (Common Crawl) که توسط یک سازمان غیرانتفاعی نگهداری می‌شود، شامل بیش از ۲.۷ میلیارد صفحه وب است. مجموعه‌داده‌ای دیگر به نام «پایل(The Pile) «، با نسخه‌ی اصلاح شده‌ی کامن کراول آغاز می‌شود و سپس بیست و یک زیرمجموعه‌داده اضافی را که شامل مطالبی از وب‌سایت کدنویسی گیت‌هاب مایکروسافت، مقالات علمی از پاب‌مد سنترال و آرکایو، مجموعه‌داده‌های مختلف کتاب و ادبیات، مجموعه‌های حقوقی از پروژه قانون آزاد، پیشینه‌های اداره ثبت اختراعات ایالات متحده، زیرنویس‌های یوتیوب و موارد دیگر است، به آن اضافه می‌کند.

گوگل نیز مجموعه‌داده خاص خود را به نام (Colossal Clean Crawled Corpus) C4 ایجاد کرده است. تحلیل واشنگتن پست از محتوای آن نشان داد که پنج منبع برتر آن عبارت بودند از : patents.google.com، ویکی‌پدیا، سایت میزبانی سندscribd.com ، وب‌سایت نیویورک تایمز و PLOS(یک ناشر غیرانتفاعی و دسترسی آزاد مقالات تحقیقاتی علمی و پزشکی). همانطور که می‌توان از مقیاس عظیم این مجموعه‌داده‌ها و برخی از منابع خاص آن‌ها حدس زد، همه آن‌ها مطالبی را از وب‌سایت‌ها، کتاب‌ها و نشریات علمی بدون کسب رضایت صریح از دارندگان حق چاپ این مطالب استفاده می‌کنند. چنین استفاده‌ای به اتهامات گسترده نقض حق چاپ منجر شده است.

در پاسخ، توسعه‌دهندگان هوش مصنوعی عموماً معتقدند که استفاده آن‌ها از داده‌ها هم تحت قوانین موجود حق چاپ قانونی است و هم به طور کلی برای کاربران و جامعه به طور وسیع سودمند است. تاکنون، تعدادی از خواهان، از جمله نیویورک تایمز، گتی ایمیج، و هنرمندان و نویسندگان مختلف، پرونده‌هایی را به دلیل نقض حق چاپ علیه توسعه‌دهندگان هوش مصنوعی مانند اوپن‌ای‌آی، مایکروسافت، استبیلیتی ای‌آی و میدجرنی تشکیل داده‌اند. نحوه

حل و فصل این دعاوی هنوز مشخص نیست. اگر دادگاه‌ها تشخیص دهند که آموزش بر روی داده‌ها برای «استخراج الگوها و اطلاعات»، به جای «باز تولید یا ترکیب یک اثر اصلی در اشکال قابل تشخیص»، تحت استفاده منصفانه قرار نمی‌گیرد، نیاز به راه‌حل‌های جدیدی برای مدیریت مجوزدهی در چنین مقیاس عظیمی خواهیم داشت. با توجه به اینکه تقریباً تمام محتوای اینترنت به طور خودکار دارای حق چاپ است، مکانیسم‌های جدیدی برای پاک‌سازی میلیاردها پست وبلاگ، نظرات کاربران، نقد و بررسی محصولات، عکس‌ها و میم‌ها، همراه با مقالات خبری، کتاب‌ها یا فیلم‌های بلند مورد نیاز خواهد بود. چنین مکانیسم‌هایی باید منافع خالقان محتوا، توسعه‌دهندگان هوش مصنوعی و منافع عمومی را متعادل کنند.

در این میان، آنچه که عصر اینترنت تاکنون به ما نشان داده، این است که استفاده گسترده و خلاقانه از داده‌ها عموماً ارزش عظیمی را هم برای کاربران فردی و جامعه به طور کلی و هم برای توسعه‌دهندگان ایجاد می‌کند. زمانی که داده‌های خفته، کم‌کاربرد یا تنها در زمینه‌های محدود مرتبط، به شیوه‌های جدید و ترکیبی بازطراحی، سنتز و متحول می‌شوند، این عمل «استخراجی» نیست، بلکه «منابع‌محور» و «تجدیدپذیر» است. بنابراین، به جای «عملیات استخراج»، ما شاهد چیزی شبیه به «کشاورزی داده» هستیم. به جای اینکه «دیگری بزرگ» ارزش را از کاربران سلب کند، ما یک اکوسیستم همزیستی از توسعه‌دهندگان، پلتفرم‌ها، کاربران و خالقان محتوا را می‌بینیم که تعاملات و مشارکت‌های آن‌ها به طور جمعی زندگی میلیاردها نفر را هر روز غنی‌تر می‌کند. به نوعی، آنچه در عصر اینترنت تجاری دیده‌ایم، نوع جدیدی از «مشاعات خصوصی» است که در عصر هوش مصنوعی، در آستانه ثمربخشی هرچه بیشتر قرار دارد.

تعریف مشاعات خصوصی

مفهوم «مشاعات(Commons) «، یا شکل تأکیدی‌تر آن، «مشاعات عمومی»، اغلب تصاویری از میدان‌های زیبای شهری، باغ‌های عمومی که همه ساکنان دسترسی برابر به آن‌ها دارند، یا منابع طبیعی مشترک را در ذهن تداعی می‌کند. اما آیا سرویس‌هایی مانند گوگل مپس نیز در این

تعریف جای می‌گیرند؟ تعریف‌های معاصر و مستحکم از «مشاعات» عموماً به منابعی اشاره دارند که هم با دسترسی باز و مشترک و هم با نظارت جمعی برای منافع فردی و جمعی یک جامعه مشخص می‌شوند. استیون لوبار، استاد مطالعات آمریکایی در دانشگاه براون، می‌نویسد: «مشاعات دارایی‌هایی هستند که همه ما در آن‌ها سهیم هستیم؛ دارایی‌هایی که نه متعلق به یک فرد یا گروه، بلکه به صورت مشترک نگهداری می‌شوند».

با این تعریف، عبارت «مشاعات خصوصی» ممکن است در ابتدا متناقض به نظر برسد. با این حال، باید اذعان کرد که در سطح مفهومی، مرزهای مشاعات همواره تا حدی سیال بوده‌اند. در کاربردهای اولیه خود، «مشاعات» به منابع طبیعی مانند مراتع و جنگل‌ها اشاره داشت که ساکنان محلی می‌توانستند برای چرای دام، شکار و سایر فعالیت‌های کشاورزی از آن‌ها استفاده کنند. امروزه، «مشاعات» اغلب گسترده‌تر از آن تعریف می‌شوند، به ویژه در گفتمان عمومی. پارک‌ها و سواحل عمومی اغلب در این دسته قرار می‌گیرند، همچون هوا، آب، و کتابخانه‌های عمومی. آثار خلاقانه در دامنه عمومی (Public Domain) نیز بخشی از مشاعات عمومی هستند، همانند خود زبان، الفبای نوشتاری، بسیاری از زبان‌های برنامه‌نویسی کامپیوتری، دستور تهیه یک نوشیدنی خاص، یا حتی منظره صورت فلکی شکارچی در یک شب تاریک و صاف.

در میان دانشگاهیان، «مشاعات» اغلب محدودتر از درک عمومی تعریف می‌شوند. الینور اوستروم، دانشمند علوم سیاسی تأثیرگذار که در سال ۲۰۰۹ به دلیل کارهایش در زمینه مشاعات جایزه نوبل اقتصاد را دریافت کرد، هشت اصل را برای نهادهای موفق «منابع مشترک» شناسایی کرد. این اصول، در مجموع، طرحی برای یک منبع جمعی ایجاد می‌کنند که به طور قابل توجهی محدودتر از هوا یا دستور تهیه یک نوشیدنی است. در مفهوم اوستروم، یک «مشاعات» منبعی است که به طور عمدی مدیریت می‌شود و دارای یک جامعه کاربری تعریف شده است. این نسخه از مشاعات دارای مرزهای دسترسی کاملاً مشخص، مجازات‌های تدریجی برای نقض قوانین، و سایر ویژگی‌های حکمرانی صریحاً بیان شده و عملیاتی است. از بسیاری جهات، یک مشاعات توسط قوانینی اداره می‌شود که ما از یک انجمن مالکان انتظار داریم.

مفاهیم گسترده‌تر از منابع مشترک، فضای بیشتری برای تنوع باقی می‌گذارند و اغلب بر «دسترسی باز» بیش از «حکمرانی» تأکید دارند. گاهی اوقات یک منبع متعلق به همه (یا هیچکس) است، مانند دستور تهیه یک نوشیدنی یا منظره صورت فلکی شکارچی. در موارد دیگر، منبع دارای یک مالک، یا مالکیت جمعی است، اما همچنان به طور گسترده قابل دسترسی است. به عنوان مثال، پارک‌های عمومی و کتابخانه‌های عمومی ممکن است توسط دولت‌های محلی اداره و توسط مالیات‌دهندگان محلی تأمین مالی شوند. آن‌ها ممکن است انواع مختلفی از هزینه‌های استفاده را دریافت کنند اما حداقل تا حدی برای همه قابل دسترسی باقی می‌مانند. نمی‌توانید اگر ساکن تأیید شده شهر نیویورک با کارت کتابخانه معتبر نباشید، کتابی از کتابخانه عمومی نیویورک قرض بگیرید. اما همچنان می‌توانید آزادانه وارد آن شوید و روز خود را در محوطه آن به خواندن نسخه‌ی میلتون فریدمن به نام «سرمایه‌داری و آزادی» بگذرانید، بدون آنکه یک پنی خرج کنید.

پس پلتفرم‌هایی مانند گوگل مپس یا یلپ را در این کیهان‌شناسی مشاعات کجا می‌توان قرار داد؟ از بسیاری جهات، آن‌ها مانند ویکی‌پدیا عمل می‌کنند، جایی که کاربران داوطلب اطلاعاتی را ارائه می‌دهند که این پلتفرم‌ها را برای همه کاربران غنی می‌سازد. همه آن‌ها به طور رایگان برای مخاطبان بسیار گسترده قابل دسترسی هستند. همه آن‌ها تحت قوانین حکمرانی خاصی عمل می‌کنند و به طور صریح توسط کارکنان حقوق‌بگیر مدیریت می‌شوند. اما گوگل مپس و یلپ، به دلیل ماهیت سودمحورشان، به طور قابل درک به عنوان خدمات تجاری اختصاصی طبقه‌بندی می‌شوند. ویکی‌پدیا، به نوبه خود، اغلب وضعیت مشاعات را دریافت می‌کند زیرا سازمانی که آن را توسعه و مدیریت می‌کند، یک بنیاد غیرانتفاعی است.

برای اینکه همه اینها را به گونه‌ای در نظر بگیریم که مشترکات اساسی آن‌ها را منعکس کند، می‌توانیم همه آن‌ها را به عنوان «مشاعات خصوصی» توصیف کنیم. از اولین نشانه‌های تجاری اینترنت در اوایل دهه ۱۹۹۰، پلتفرم‌های خصوصی یا تحت مدیریت خصوصی که کاربران را به عنوان تولیدکننده و ناظر به کار می‌گیرند، گسترش یافته‌اند. برچسب‌های مختلفی برای جنبه‌ها و

نمونه‌های مختلف این الگو اعمال شده است، از جمله وب ۲.۰، رسانه‌های اجتماعی، اقتصاد مشارکتی، اقتصاد گیگ (Gig Economy) و سرمایه‌داری نظارتی. اما هیچ‌یک از اینها—چه تأییدکننده و چه تحقیرآمیز—به درستی ظهور منابع رایگان و تقریباً رایگان مدیریت زندگی را که به طور مؤثر به عنوان خدمات اجتماعی و ابزارهای عمومی خصوصی‌سازی شده عمل می‌کنند، منتقل نمی‌کنند؛ «دولت رفاهی که با سرعت سرمایه‌داری حرکت می‌کند.» اصطلاح «مشاعات خصوصی» این معنا را منتقل می‌کند.

هر بار که چیزی را با استفاده از گوگل جستجو می‌کنید، یا یک رویداد تقویم به Calendar اضافه می‌کنید، یا مسیر و گزارش ترافیک از ویز (Waze) دریافت می‌کنید، یا به دنبال آپارتمان در کرگس‌لیست (Craigslist) می‌گردید، یا عکس‌ها را در دراپ‌باکس ذخیره می‌کنید، از مشاعات خصوصی بهره‌مند می‌شوید. همین امر زمانی صادق است که از متخصصان صنعت خود در لینکدین مشورت شغلی می‌گیرید، یا یک ویدئوی یوتیوب را تماشا می‌کنید که نحوه تعمیر سینک آشپزخانه نشتی شما را نشان می‌دهد، یا به یک پادکست در اسپاتیفای رایگان گوش می‌دهید، یا یک ایمیل را با مجموعه‌ای از ایموجی‌ها تکمیل می‌کنید.

در حالی که شرکت‌های انتفاعی و سایر نهادهای خصوصی نقش حیاتی در ایجاد مشاعات خصوصی ایفا می‌کنند، عموم مردم نیز به وضوح نقش دارند. در فیس‌بوک، یوتیوب، ایکس (X.com) و سایر پلتفرم‌ها، کاربران فردی بخش عمده‌ای از محتوا، همه مخاطبان و تمام رفتارهای کاربری را در قالب کلیک‌ها، تعاملات اجتماعی، خریدها و موارد دیگر فراهم می‌کنند که به اپراتورهای پلتفرم کمک می‌کند تا سود کسب کنند.

سنجش ارزش: مفهوم مازاد مصرف‌کننده

این رابطه همزیستی میان کاربران و پلتفرم‌ها، ارزش بی‌سابقه‌ای را خلق می‌کند. اما همواره سوالاتی را در مورد چگونگی توزیع این ارزش نیز ایجاد می‌کند. هنگامی که مردم می‌بینند شرکت‌های بزرگ فناوری سالانه میلیاردها دلار درآمد تبلیغاتی کسب می‌کنند، اغلب می‌خواهند

بدانند سهم آن‌ها از این فعالیت کجاست. این انتظار کاملاً قابل درک است. در سال ۲۰۲۳، آلفابت (همان گوگل) ۷۳.۷ میلیارد دلار سود گزارش کرد و مایکروسافت ۷۲.۴ میلیارد دلار. در فهرست «ثروتمندترین افراد جهان» فوربس در سال ۲۰۲۴، هفت نفر از ده نفر برتر، ثروت خود را از طریق فناوری به دست آورده‌اند.

برای کاربران، ارزشی که از مشاعات خصوصی کسب می‌کنند، فاقد شفافیت صریح دلار و سنت است و این حداقل یکی از دلایلی است که این تصور که شرکت‌های بزرگ فناوری بیشتر ارزش ایجاد شده را برای خود نگه می‌دارند، تا این حد گسترده است. اما این بدان معنا نیست که یک تبادل ارزشی متقابل‌تر در حال وقوع نیست. اقتصاددانان به تفاوت بین آنچه مردم برای یک محصول یا خدمت می‌پردازند و میزان ارزش آن برای آن‌ها، «مازاد مصرف‌کننده (Consumer Surplus)» می‌گویند. اگر یک ژاکت را به قیمت ۱۰۰ دلار بخرید، اما فکر کنید ۲۰۰ دلار ارزش دارد، مازاد مصرف‌کننده در این مورد ۱۰۰ دلار است.

زمانی که یک محصول یا خدمت به صورت رایگان ارائه می‌شود، نیز مازاد مصرف‌کننده می‌تواند وجود داشته باشد، به شرط آنکه مصرف‌کننده برای آن ارزش قائل باشد. به این ترتیب، تلویزیون و رادیو برای دهه‌ها منابع اصلی مازاد مصرف‌کننده بوده‌اند. شما برای آن‌ها چیزی جز هزینه تلویزیون یا رادیوی خود پرداخت نمی‌کنید و در عوض یک عمر سرگرمی و اطلاعات به دست می‌آورید.

اریک براینجلفسون، مدیر آزمایشگاه اقتصاد دیجیتال استنفورد، و آویناش کولیس، استاد دانشگاه کارنگی ملون، با درک اینکه اندازه‌گیری اقتصاد دیجیتال با روش‌های سنتی دشوار است زیرا بسیاری از محصولات و خدمات خود را به صورت رایگان ارائه می‌دهد، مجموعه‌ای از «آزمایشات انتخاب آنلاین گسترده» را طراحی کردند تا مازاد مصرف‌کننده را در حوزه اینترنت اندازه‌گیری کنند. آن‌ها این آزمایشات را در مقاله‌ای در سال ۲۰۱۹ برای هاروارد بیزنس ریویو اینگونه توصیف کردند: «برای اندازه‌گیری مازاد مصرف‌کننده تولید شده توسط فیس‌بوک،

نمونه‌ای نماینده از کاربران آمریکایی این پلتفرم را جذب کردیم و به آن‌ها مقادیر مختلفی پول پیشنهاد دادیم تا به مدت یک ماه استفاده از آن را ترک کنند. برای تأیید پاسخ‌ها، برخی از شرکت‌کنندگان به طور تصادفی انتخاب شدند تا واقعاً پرداخت‌ها را دریافت کرده و به مدت یک ماه از این سرویس صرف‌نظر کنند. ما به طور موقت آن‌ها را به عنوان دوستان فیس‌بوک اضافه کردیم—البته با اجازه آن‌ها—تا تأیید کنیم که در آن ماه وارد سیستم نشده‌اند».

آنچه براینجلفسون و کولیس دریافتند این بود که اینترنت اساساً یک ماشین تولید مازاد مصرف‌کننده است. آن‌ها علاوه بر انجام آزمایشات متمرکز بر سایت‌های خاص مانند فیس‌بوک و ویکی‌پدیا، به کاربران برای دسته‌بندی‌های گسترده‌تر نیز پیشنهاداتی دادند، از جمله ایمیل، موتورهای جستجو، نقشه‌ها، تجارت الکترونیک، ویدئو، موسیقی، رسانه‌های اجتماعی و پیام‌رسانی فوری. بر اساس داده‌های نظرسنجی آن‌ها، میانگین مبلغی که افراد حاضر بودند برای یک سال از دست دادن موتورهای جستجو بپذیرند، ۱۷۵۳۰ دلار بود. برای ایمیل، ۸۴۱۴ دلار و برای نقشه‌های دیجیتال، ۳۶۴۸ دلار.

با وجود قابل توجه بودن این ارزیابی‌های بالا، شواهد محکمی وجود دارد که حتی این ارقام نیز نمی‌توانند به طور کامل ارزش مشاعات خصوصی را در زندگی ما منعکس کنند. مشاهدات براینجلفسون و کولیس در مقاله HBR به این نکته اشاره دارد: «بریتانیکا زمانی چند هزار دلار قیمت داشت، به این معنی که مشتریان آن حداقل به همان میزان برایش ارزش قائل بودند. ویکی‌پدیا، یک سرویس رایگان، مقالات بسیار بیشتری، با کیفیتی قابل مقایسه، نسبت به بریتانیکا داشته است.» به عبارت دیگر، ویکی‌پدیا نه تنها جایگزین محصولی شد که زمانی بسیار گران بود، بلکه محصولی بهتر نیز هست، زیرا مقالات به مراتب بیشتری دارد. این امر آن را برای هر کاربر فردی مفیدتر و برای تعداد بیشتری از کاربران مرتبط‌تر می‌کند.

چقدر مفیدتر؟ در سال ۱۹۹۰، بهترین سال فروش تاریخ دایرةالمعارف بریتانیکا، ۱۲۰ هزار نسخه در ایالات متحده فروخته شد. در مقایسه، ویکی‌پدیا در ایالات متحده به تنهایی حدود ۴ میلیارد بازدید ماهانه صفحه را ثبت می‌کند. علاوه بر اینکه ویکی‌پدیا فوق‌العاده مفید است، استفاده از آن

نیـز فوق‌العـاده آسـان اسـت. حتـی اگـر شـما فـداکارترین کـاربر بریتانیکـا در تـاریخ بودیـد، بـا یـک مجموعه کامل در اتاق مطالعه خود، یک مجموعـه دیگر در دفتر کارتان، و نسخه‌های تک‌جلدی فشرده در حمام و ماشین خود، توانایی شما برای مشورت با آن هر زمان که می‌خواستید، هرگز با سهولت دسترسی به ویکی‌پدیا امروز قابل مقایسه نبود.

تولید ارزش توسط مشاعات خصوصی با ارقام سخت‌تر که به آن اعمال می‌کنیم، بیشتر به چشم می‌آید و نشـان می‌دهد که چگونـه ایـن ارزش از طریـق ایـن منبـع شـگفت‌انگیز افزایش می‌یابد. به عنوان مثال، هل واریان در مصاحبه‌ای در سال ۲۰۱۷ به مثال عکاسی اشاره می‌کند: در سال ۲۰۰۰، ۸۰ میلیارد عکس تولید شد، زیرا تنها سه شرکت فیلم عکاسی تولید می‌کردند. در سال ۲۰۱۵، ایـن رقـم بـه ۱.۶ تریلیـون عکس رسـید. در سـال ۲۰۰۰، هزینـه هـر عکـس حـدود ۵۰ سـنت بـود، امـا اکنـون تقریبـاً صـفر اسـت. هـر فـرد عـادی می‌گفـت: «وای، چـه افـزایش فوق‌العـاده‌ای در بهـره‌وری، زیـرا خروجـی بسـیار بیشـتری داریـم و هزینـه بسـیار کمتـری.» در ایـن مـورد، منفعـت اقتصـادی بـرای عکس‌های سال ۲۰۱۵ به تنهایی ۸۰۰ میلیارد دلار است – یعنی ۱.۶ تریلیون عکس ضربدر ۵۰ سنت برای هر عکس.

بسیاری از عکس‌های دیجیتالی که در سال ۲۰۱۵ گرفته شدند، اگر افراد مجبور به پرداخت هزینـه بودند، هرگز گرفته نمی‌شدند. قبل از عصر آیفون، چند نفر از پولاروید اینستاماتیک خود برای به خاطر سپردن جای پارک ماشینشان در پارکینگ استفاده می‌کردند؟ اما این بخشی از چیزی است که این نوآوری خاص را بسیار ارزشمند می‌کند. تنها ایـن نیست که عکاسی دیجیتال هزینه‌های چیزهایی را که قبلاً مردم برایشان پول می‌دادند، به صفر رسانده است. عمیق‌تر از آن، ایـن واقعیـت است که این فناوری‌های جدید به ما امکان می‌دهند کارهای کاملاً جدیدی انجام دهیم، مانند استفاده از عکس‌ها برای یادداشت‌برداری.

تنها عکاس نیست که با صرفه‌جویی در ۵۰ سنت به ازای هر عکس، از ایـن قابلیت جدید ارزش کسب می‌کند. از آنجا که میلیاردها نفر اکنون دوربین‌های دیجیتال در تلفن‌های خود دارند، و هزینه هر تصویری که می‌سازیم عملاً صفر است، ما سالانه تریلیون‌ها عکس می‌گیریم. ما همچنین

تعداد زیادی از این تصاویر را با دیگران، از جمله غریبه‌ها، به اشتراک می‌گذاریم، زیرا مشاعات خصوصی خدمات ذخیره‌سازی و توزیع را نیز به صورت رایگان ارائه می‌دهد. همانقدر که این امر برای ما به عنوان عکاس ارزش ایجاد می‌کند، به طور قطع ارزش بیشتری برای ما به عنوان مصرف‌کنندگان اطلاعات ایجاد می‌کند. اکنون، اگر می‌خواهید پیش از اقامت در هتلی در استانبول، نمایی از داخل آن داشته باشید، تصاویر بی‌طرفانه شخص ثالث احتمالاً در جایی قابل دسترسی هستند. اگر می‌خواهید بدانید چه کسانی به مهمانی که نتوانستید شرکت کنید رفته‌اند، می‌توانید آن را پیدا کنید.

ارزش این دسترسی بی‌سابقه به «دانش بزرگ جهانی (Global Big Knowledge) «که میلیاردها نفر اکنون آن را بدیهی می‌دانند، چقدر است؟ در حالی که محاسبه آن غیرممکن است، اما به روش‌های بی‌پایان در حال تکامل و بسیار مفیدی آشکار می‌شود. در سال‌های اولیه یوتیوب، منتقدان به سرعت آن را به عنوان «محتوای بی‌فایده و با طول توجه کوتاه» توصیف کردند. سپس، در یکی از دگردیسی‌های شگفت‌انگیز خود، یوتیوب شروع به کار به عنوان یک ویکی‌پدیای بصری و کاربردی دانش کرد. اگر در پی فوران یک ابرآتشفشان که تنها شما و یوتیوب را باقی می‌گذاشت، می‌توانستید از کاتالوگ بی‌کران ویدئوهای «چگونه انجام دهیم» آن برای بازسازی تمدن از صفر، یکی پس از دیگری استفاده کنید. یوتیوب که زمانی به عنوان نشانه‌ای دیگر از سقوط بشریت به پوچی نادیده گرفته می‌شد، اکنون به عنوان گنجینه‌ای مجازی از دانش بشری چنان جامع ایستاده است که کتابخانه کنگره را شبیه به جعبه‌ای از رمان‌های قدیمی رها شده در روز جمع‌آوری زباله می‌کند.

البته، کمی دانش می‌تواند خطرناک باشد، درست است؟ اتاق‌های پژواک(Echo chambers) ، حباب‌های فیلتر (filter bubbles) و رادیکالیزه شدن الگوریتمی که گیمرهای جوان و تأثیرپذیر را به دیدگاه‌های فزاینده افراطی و مخرب سوق می‌دهد، روایتی است که پوشش رسانه‌ای قابل توجهی داشته است. اما داستان کم‌تر پوشش داده شده‌ای نیز وجود دارد، حتی اگر دقیقاً به

همان شیوه کار کند: «جهش الگوریتمی .(Algorithmic Springboarding) «این همان اتفاقی است که وقتی الگوریتم‌های توصیه یوتیوب کاربران را به سمت مسیرهای بی‌انتهای آموزش، خودسازی و پیشرفت شغلی سوق می‌دهند، رخ می‌دهد.

فرض کنید شما فارغ‌التحصیل اخیر دبیرستان هستید و در تلاشید تا گام بعدی خود را مشخص کنید. شما زمان زیادی را به بازی کردن Tekken وFinal Fantasy ، تبادل شوخی در انجمن‌های ردیت و گشت‌وگذار بی‌هدف در اینترنت می‌گذرانید - به حدی که در نهایت به فناوری‌های زیربنایی آن علاقه‌مند می‌شوید. کنجکاو شده، یک ویدئوی «پایتون برای مبتدیان» را که توسط یک گورو فول‌استک بسیار سرگرم‌کننده میزبانی می‌شود تماشا می‌کنید و جذب می‌شوید. یک ویدئو به ویدئوی دیگری منجر می‌شود. از آنجا که یوتیوب «زمان تماشا» را اولویت‌بندی می‌کند و اغلب ویدئوهای عمیق و طولانی را پیشنهاد می‌دهد، شما به مسیری از آموزش‌های برنامه‌نویسی به طرز موذیانه‌ای آموزنده هدایت می‌شوید. این همان منطق الگوریتمی است که به تقویت باورهای افراطی متهم شده است، با این حال در اینجا مسیری را به سوی آینده‌ای بهتر می‌سازد.

در نهایت، به Python.org می‌روید و مفسر پایتون را که رایگان است، دانلود می‌کنید. مشتاق به آزمایش، به دنبال کمک‌های اضافی در گیت‌هاب می‌گردید، جایی که مخازنی با پروژه‌های دوستانه برای مبتدیان پیدا می‌کنید و با جامعه کدنویسی برای نکات و انتقادات درگیر می‌شوید. همچنین پلتفرم‌هایی مانند Stack Overflow و freeCodeCamp را کشف می‌کنید که به ترتیب انجمن‌هایی برای عیب‌یابی و دوره‌های رایگان ارائه می‌دهند. این منابع به شما کمک می‌کنند تا اولین اسکریپت پایتون خود را بنویسید – یک ابزار بصری‌سازی داده که نمودارهای ساده‌ای را از فایل‌های CSV ایجاد می‌کند، با الهام از علاقه شما به ردیابی آمار درون بازی از بازی‌های ویدئویی مورد علاقه خود. در این سفر، هر عنصر از مشاعات خصوصی – از آموزش‌های یوتیوب گرفته تا جوامع متن‌باز در گیت‌هاب، تا دوره‌های رایگان درfreeCodeCamp ، تا شبکه‌سازی

حرفه‌ای در لینکدین – به طور هم‌افزا مشارکت می‌کند و مسیری سفارشی ایجاد می‌کند که کنجکاوی اولیه شما را به مهارت‌های قابل استخدام تبدیل می‌کند.

و داستان شما لازم نیست به همین جا ختم شود. در مقطعی برای یک موقعیت تحلیل‌گر داده سطح پایه درخواست می‌دهید و ابزار بصری‌سازی داده خود را به عنوان بخشی از پورتفولیوی خود ارائه می‌دهید. ابتکار شما نتیجه می‌دهد و شغل را به دست می‌آورید. با کسب تجربه، یک پروفایل لینکدین ایجاد می‌کنید که تخصص برنامه‌نویسی، مهارت‌های تحلیل داده و استعداد شما در خودکارسازی فرآیندها را برجسته می‌کند. در کمترین زمان، یک استارت‌آپ پروفایل شما را شناسایی می‌کند و ارزشی را که می‌توانید به ارمغان آورید، می‌بینند. پیشنهادی برای یک نقش چالش‌برانگیزتر و پربارتر دریافت می‌کنید و بدین ترتیب گامی دیگر به سوی تعریف جایگاه خود در جهان برمی‌دارید. دموکراتیک کردن دسترسی به دانش و فرصت‌ها، مشاعات خصوصی عاملیت فردی، فرصت آموزشی، تحرک اجتماعی و در نهایت رشد حرفه‌ای را فراهم می‌کند. و حداقل از یک جنبه، پتانسیل آن نامحدود است. این بدان دلیل است که مشاعات دیجیتال بسیار متفاوت از مراتع مشترک و سایر منابع طبیعی که قدیمی‌ترین اشکال مشاعات سنتی را تشکیل می‌دادند، عمل می‌کنند.

مشاعات دیجیتال در برابر تراژدی مشاعات

در مورد یک مرتع گوسفند یا یک منطقه ماهیگیری، نتایج می‌تواند تحت تأثیر یک پویایی منحرف قرار گیرد که اکولوژیست گرت هاردین در مقاله‌ای در سال ۱۹۶۸ آن را «تراژدی مشاعات» نامید. هاردین استدلال کرد که هر چه مردم یک منبع مشترک را مفیدتر یا با ارزش‌تر بیابند، به احتمال زیاد از طریق استفاده بیش از حد، آن را به طور جمعی نابود خواهند کرد. این چیزی نیست که در مورد منابع دیجیتال صدق کند، مهم نیست که بدبینان چند بار پلتفرم‌های بزرگ فناوری را «عملیات استخراج» توصیف کنند.

گرت هاردین خود نگرانی‌های عمیقی در مورد «عملیات استخراج» داشت – از نوع قدیمی که

شامل منابع طبیعی واقعی بود. هاردین، استاد اکولوژی در دانشگاه کالیفرنیا، سانتا باربارا، مقاله خود «تراژدی مشاعات» را در زمانی نوشت که حتی بر اساس استانداردهای معاصر، پیش‌بینی‌های قریب‌الوقوع فاجعه زیست‌محیطی به شدت آخرالزمانی بود. در کتاب «بمب جمعیت» که در همان سال ۱۹۶۸ منتشر شد، پروفسور پل ارلیش از استنفورد به طور مشهور هشدار داد که «صدها میلیون نفر» در دهه ۱۹۷۰ بدون توجه به «برنامه‌های اضطراری» برای جلوگیری از این فاجعه، از گرسنگی خواهند مرد. دیدگاه هاردین نیز کمتر از این وخیم نبود. او در نشریه Science نوشت: «فرض ضمنی و تقریباً جهانی بحث‌های منتشر شده در مجلات علمی حرفه‌ای و نیمه‌محبوب این است که مشکل مورد بحث راه‌حل فنی دارد.» اما به تخمین او، این در مورد مشکل جمعیت بیش از حد در دنیایی با منابع محدود، به سادگی صادق نبود. «مردم فکر می‌کنند که کشاورزی در دریاها یا توسعه سویه‌های جدید گندم، مشکل را – به لحاظ فناوری – حل خواهد کرد. من در اینجا سعی می‌کنم نشان دهم که راه‌حلی که آن‌ها به دنبال آن هستند، یافت نخواهد شد«.

هاردین در ادامه تاکید کرد که افراد همیشه سعی می‌کنند حداکثر سود خود را هنگام بهره‌برداری از یک منبع مشترک به دست آورند. این «تراژدی» مشاعات است: یک گله‌دار با دسترسی به یک مرتع مشترک و آزاد، همیشه یک رأس گاو دیگر به گله خود اضافه خواهد کرد اگر بتواند این کار را انجام دهد. این انتخاب منطقی است. هر منفعتی که از آن حیوان اضافی به دست می‌آورد – مانند فروش آن در حراج – تماماً به او می‌رسد. در همین حال، هزینه تصمیم او – کاهش اندک در میزان چمن موجود برای چرای مشترک – توسط هر گله‌داری که از آن منبع مشترک استفاده می‌کند، متحمل خواهد شد. هاردین استدلال کرد که با رفتار هر گله‌دار به همین شیوه منطقی، مشاعات با کاهش اجتناب‌ناپذیری مواجه می‌شوند. هاردین نتیجه گرفت: «هر انسانی در سیستمی قفل شده است که او را مجبور می‌کند گله خود را بدون محدودیت افزایش دهد – در دنیایی که محدود است.» «ویرانی سرنوشتی است که همه انسان‌ها به سوی آن می‌شتابند، هر یک منافع خود را در جامعه‌ای که به آزادی مشاعات اعتقاد دارد، دنبال

می‌کند. آزادی در یک مشاعات، ویرانی برای همه به ارمغان می‌آورد.» هاردین معتقد بود که تنها راه‌حل‌های این معضل شامل «مالکیت خصوصی، یا چیزی شبیه به آن» یا «قوانین اجباری یا ابزارهای مالیاتی» است.

با توجه به موضع بنیادین او در برابر فناوری و رشد، کار او به طور متناقض به عنوان نقطه اتکای مورد علاقه توسعه‌دهندگان املاک و مستغلات که به دنبال تبدیل مراتع تکه‌تکه و فرسایش یافته با باد به زمین‌های گلف سرسبز بودند، عمل کرده است. در همین حال، ادعای او مبنی بر اینکه هر مشاعه‌ی باز، چه پارک ملی باشد و چه گله گاومیش وحشی، به ناچار به سمت فروپاشی پیش می‌رود، همچنین به تجدید علاقه به رویکردهای مشاعاتی سنتی برای مدیریت منابع کمک کرد. الینور اوستروم، به ویژه، سال‌ها را با تحقیقات میدانی گسترده صرف اثبات این موضوع کرد که جوامع محلی می‌توانند، و اغلب نیز این کار را می‌کردند، منابع مشترک را به طور پایدار و بدون توسل به خصوصی‌سازی یا نظارت دولتی مدیریت کنند.

در سال‌های اولیه خود، اینترنت به عنوان نوع جدیدی از منبع مشترک وجود داشت. مطمئناً، کسی مالک سرورها، روترها، سوئیچ‌ها و کابل‌های فیبر نوری بود که کد را به یک زمین مجازی تبدیل می‌کردند که می‌توانستید آزادانه آن را کاوش کرده و به روش‌های مختلف از آن استفاده کنید. محدودیت‌های تکنولوژیکی نیز نوعی تنظیم‌کننده ذاتی عمل می‌کردند؛ کارهای زیادی نمی‌توانستید در آن انجام دهید. اما نظارت رسمی بسیار کمی نیز وجود داشت. در محدوده آنچه پروتکل‌های ارتباطی مانند TCP/IP و HTTP امکان‌پذیر می‌ساختند، تقریباً آزادی عمل برای انجام هر کاری که می‌خواستید را داشتید.

با تکامل فناوری‌ها، طیف وسیع‌تری از اقدامات امکان‌پذیر شد. منافع تجاری شروع به درک چگونگی توسعه سودآور این مرزهای باز کردند. تنظیم‌کنندگان دولتی متوجه شدند که مردم بدون آن‌ها شروع به ساختن دنیایی کاملاً جدید کرده‌اند. همانطور که این مزاحمان جدید شروع به ایفای نقش بیشتری در این حوزه کردند، تمایل به مشاعات دیجیتال – فضاهای آنلاینی که به طور صریح در برابر الزامات تجاری و نظارت دولتی مقاومت می‌کنند – حتی بیشتر شد. جنبش

نرم‌افزار متن‌باز شتاب گرفت و بیش از هر زمان دیگری مرتبط شد. یک سازمان غیرانتفاعی به نام کریتیو کامنز (Creative Commons) مجموعه‌ای از مجوزها را توسعه داد که به خالقان محتوا امکان می‌دهد آثار خود را انعطاف‌پذیرتر از حق چاپ سنتی به اشتراک بگذارند.

با تأکید بر حکمرانی مبتنی بر همتا، رویکرد مشاعات که توسط محققانی مانند اوستروم بیان شد، از بسیاری جهات برای اینترنت مناسب است. با این حال، تأکید زیاد آن بر تخصیص منابع کمیاب در مواجهه با تقاضای شدید و پایدار، چارچوب گمراه‌کننده‌ای برای مفهوم‌سازی چالش‌ها و فرصت‌های جهان دیجیتال نیز هست. این بدان دلیل است که پلتفرم‌هایی مانند گوگل مپس یا ویز، با همان چالش‌هایی که یک مرتع احاطه شده توسط گله‌داران به دنبال ارزان‌ترین راه برای تغذیه گاوهای خود هستند، روبرو نیستند. برخلاف یک مرتع چرا که مصرف هر گاو، چمن کمتری را برای گاوهای دیگر باقی می‌گذارد، شما می‌توانید از پلتفرم‌های دیجیتال به طور گسترده استفاده کنید بدون اینکه توانایی دیگران برای استفاده از آن‌ها را کاهش دهید. در واقع، عکس این قضیه صادق است. اگر هزار نفر از ویز در شهر شما برای ناوبری و گزارش وضعیت ترافیک استفاده کنند، این ابزار مفیدی است. اگر ده هزار نفر از آن استفاده کنند، حتی بهتر است. بنابراین، یک مشاعات دیجیتال می‌تواند بسیار متفاوت از مشاعات فیزیکی سنتی عمل کند. به جای کنترل دقیق دسترسی به منابع کمیاب و سخت جایگزین مانند چوب یا سالمون، یک استراتژی واضح برای مشاعات دیجیتال این است که منابعی که ارزش آن‌ها را تولید می‌کنند – داده‌ها – را به عنوان منابع مشترک «غیررقابتی» تلقی کنند که باید تا حد امکان کشت شده و فعالانه استفاده شوند.

این امر راه را برای مشاعات خصوصی، و به ویژه مشاعات خصوصی که توسط سازمان‌های بزرگ و با امکانات اداره می‌شوند، هموار می‌کند. بخش زیادی از ارزش مشاعات خصوصی از کاربران نشأت می‌گیرد، اما این شرکت‌ها هستند که پلتفرم‌ها و فناوری را برای اجرای این فضاهای مشترک دیجیتال فراهم می‌کنند و امکان گرد هم آمدن بسیاری از کاربران را دارند. با این حال، بدیهی است که نظرات مختلفی در مورد بهترین روش مفهوم‌سازی داده‌هایی که کاربران در چنین

زمینه‌هایی تولید می‌کنند، وجود دارد. قرارداد اساسی که امروزه بخش عمده‌ای از مشاعات خصوصی را قدرت می‌بخشد، این است که کاربران در ازای دسترسی اپراتورهای پلتفرم به داده‌هایی که تولید می‌کنند، خدمات رایگان دریافت می‌کنند. اما درآمدهای هنگفت شرکت‌های بزرگ فناوری، همراه با روایت‌های مداوم در مورد ارزش حریم خصوصی، که به صراحت در قوانینی مانند مقررات عمومی حفاظت از داده‌های اتحادیه اروپا و قانون حریم خصوصی مصرف‌کننده کالیفرنیا تبلور یافته‌اند، سوالاتی را در مورد انصاف برمی‌انگیزند. آیا دسترسی رایگان به محصولات و خدمات یک تبادل واقعاً عادلانه است؟

بر اساس استاتیستا، درآمد سالانه متا به ازای هر کاربر (ARPU) در سال ۲۰۲۳ برابر با ۴۴.۶۰ دلار بود. این معیار در کشورهای مختلف و پلتفرم‌های مختلف متا (فیس‌بوک، مسنجر، اینستاگرام و واتس‌اپ) متفاوت است. یک کاربر اینستاگرام در فرانسه ممکن است درآمد بیشتری برای متا ایجاد کند تا یک کاربر فیس‌بوک در ژاپن، یا برعکس. با این حال، به طور متوسط، هر کدام تقریباً ۴۴.۶۰ دلار در سال یا ۳.۷۱ دلار در ماه برای متا ارزش دارند. همچنین از تحقیقات اریک براینجلفسون و آویناش کولیس می‌دانیم که میانگین پولی که کاربران فیس‌بوک حاضر بودند برای یک ماه ترک خدمت بپذیرند، ۴۸ دلار بود.

این ارقام یک وضعیت برد-برد کلاسیک را نشان می‌دهد. به طور متوسط، ارزشی که کاربران احساس می‌کنند دریافت می‌کنند، بیش از دوازده برابر ارزشی است که برای متا ایجاد می‌کنند. اما این تبادل همچنان منابع کافی را برای متا فراهم می‌کند تا خدمات خود را حفظ کند، به طور قابل توجهی در توسعه خدمات جدید (از جمله بسیاری از مدل‌های هوش مصنوعی متن‌باز که به صورت رایگان در اختیار جهان قرار می‌دهد) سرمایه‌گذاری کند و سودی برای سهامداران خود ایجاد کند.

البته، همه کاربران مقدار یکسانی از ارزش را برای یک پلتفرم ایجاد نمی‌کنند. کاربران بسیار فعال ممکن است پست‌ها، لایک‌ها، نقد و بررسی‌ها، رتبه‌بندی‌ها و انواع دیگر داده‌ها را بیش از کاربران کمتر فعال ایجاد کنند. برخی کاربران پست‌های وبلاگ، ویدئوها و انواع دیگر محتوا را ایجاد

می‌کنند که سطوح بالایی از مشارکت را از سوی سایر کاربران به دنبال دارد. به همین ترتیب، همه داده‌ها به یک اندازه ارزشمند نیستند و کاربران نیز زمان و تلاش یکسانی را برای ایجاد آن‌ها صرف نمی‌کنند. داده‌هایی که به طور غیرفعال هنگام ورود به سیستم، ماندن در سایت برای مدت زمان مشخص و انجام فعالیت‌های مختلف در آن تولید می‌کنید، ممکن است به تلاش زیادی از جانب شما نیاز نداشته باشد. نظراتی که در پاسخ به پست اینستاگرام شخص دیگری می‌نویسید، ممکن است سرمایه‌گذاری بیشتر زمان و فکر را نشان دهد، اما هنوز با پانزده ساعتی که ممکن است برای تولید یک ویدئوی یوتیوب خوش‌ساخت سرمایه‌گذاری کنید، یکسان نیست.

این نیز درست است که زمینه اهمیت دارد. داده‌های موقعیت مکانی شما ممکن است به صورت فردی برای شما حداقل ارزش را داشته باشند، اما هنگامی که با داده‌های میلیون‌ها کاربر دیگر جمع‌آوری می‌شوند، ممکن است برای برنامه‌ریزی شهری یا تبلیغات هدفمند بسیار ارزشمند شوند. به همین ترتیب، در دنیایی که حتی اشکال اساسی‌تر داده‌ها، مانند آهنگ‌ها یا فیلم‌های بلند، به وفور فزاینده‌ای در دسترس هستند، آنچه مصرف‌کنندگان در نهایت حاضر به پرداخت آن هستند، نه خود داده‌ها، بلکه مکانیسم‌هایی برای کشف و دسترسی راحت به آن‌ها به صورت انبوه است. ارزش داده‌ها همچنین می‌تواند به طور چشمگیری در طول زمان تغییر کند، زیرا فناوری‌ها تکامل می‌یابند و کاربردهای جدید کشف می‌شوند. در حالی که تأکید بر حریم خصوصی داده‌ها برای محافظت از حقوق فردی حیاتی است، این امر همچنین این تصور را تقویت می‌کند که داده‌ها از هر نوع، در وهله اول مالکیت خصوصی هستند — و به همین ترتیب ذاتاً ارزشمند. با این حال، همانطور که بسیاری از سناریوهای مشاعات خصوصی نشان می‌دهند، مفهوم‌سازی داده‌ها به عنوان نوعی کالای شبه‌عمومی که به طور فعال به اشتراک گذاشته، تجمیع و به روش‌های جدید اعمال می‌شود، می‌تواند به ایجاد خدماتی کمک کند که کاربران از آن‌ها بسیار بیشتر از آنچه اپراتورهای پلتفرم خود به دست می‌آورند، ارزش کسب کنند.

قدرت تحول‌آفرین هوش مصنوعی: هوش شبکه‌ای جهانی

در دنیای دیجیتال، تراژدی مشاعات، اگر اصلاً وجود داشته باشد، در فراهم آوردن دسترسی باز به منابع مشترکی نیست که سپس از طریق استفاده بی‌رویه از بین بروند. در عوض، زمانی رخ می‌دهد که مردم تلاش می‌کنند محدودیت‌هایی بر میزان داده‌ای که ایجاد می‌کنیم، میزان اشتراکی که می‌گذاریم و اینکه چه کسی می‌تواند آن را به اشتراک بگذارد، اعمال کنند. برخلاف چراگاه یا درختان، داده‌ها «غیررقابتی» و همیشه در حال تکثیر هستند. بهره‌برداری از این ویژگی‌ها، ارزش را به حداکثر می‌رساند. این امر با توجه به تأثیری که هوش مصنوعی بر ارزش کلی مشاعات خصوصی خواهد داشت، حتی بیشتر نیز صادق است. به دلیل ظرفیت هوش مصنوعی برای تحلیل، ارزیابی، بازیابی، خلاصه‌سازی و سنتز داده‌های مرتبط از کوه‌های داده‌ای که اکنون تولید می‌کنیم، ارزش مشاعات خصوصی در حال افزایش است. اما اگر ما رویکردی باز و مشارکتی نسبت به اشتراک‌گذاری داده‌ها را در پیش بگیریم، در عین حال چارچوب‌های مستحکمی برای حفاظت از حریم خصوصی و استفاده اخلاقی از داده‌ها توسعه دهیم، این ارزش می‌تواند به میزان بسیار بیشتری افزایش یابد.

برخلاف مشاعات فیزیکی که حتی با مدیریت دقیق، در طول زمان در معرض خطر فرسایش قرار دارند، مشاعات دیجیتال تمایل به بهتر شدن دارند. بسیار بهتر. هم ویکی‌پدیا و هم یوتیوب اکنون منابع بهتری نسبت به زمانی هستند که براینجلفسون و کولیس در سال ۲۰۱۷ از کاربران در مورد میزان ارزش آن‌ها پرسیدند. ویکی‌پدیا مقالات بیشتری نسبت به آن زمان دارد. مقالات آن با مشارکت بیشتر کاربران، در طول زمان جامع‌تر و دقیق‌تر می‌شوند. برای بسیاری از پلتفرم‌های مشاعات خصوصی دیگر، همین پویایی‌ها در جریان است.

اکنون تصور کنید وقتی هوش مصنوعی را به این ترکیب اضافه می‌کنید چه اتفاقی می‌افتد. موتورهای جستجوی سنتی در خواندن ذهن ما چندان ماهر نیستند. آن‌ها عمدتاً نسبت به میزان خدماتی که به ما ارائه می‌دهند یا نمی‌دهند، بی‌تفاوت‌اند. اگر «جگوار» را به عنوان کلمه جستجو وارد کنید، ممکن است نتایجی برای یک ماشین، یک حیوان یا یک تیم فوتبال دریافت

کنید. اگر از آن بپرسید که آیا هیچ جگوار (حیوان) صاحب جگوار (ماشین) است، هیچ ایده‌ای از منظور شما نخواهد داشت و علاقه‌ای به تلاش برای درک شما نخواهد داشت.

با یک مدل زبان بزرگ(LLM) ، وضعیت متفاوت است. ممکن است نتواند به طور قطعی به پرسش شما پاسخ دهد، اما قطعاً شانس بیشتری برای این کار خواهد داشت. و ما در حال حاضر شروع به دیدن چگونگی این خواهیم بود که به زودی دستیاران هوش مصنوعی قادر خواهند بود با موفقیت بر اساس دستورالعمل‌هایی مانند «آن مطلب درباره پادکست بیت‌کوین که حدود شش ماه پیش گوش می‌دادم را پیدا کن؟ همان که درباره عملیات Checkpoint 2.0 بود؟ سپس خلاصه‌ای از متن آن را در حدود هزار کلمه بنویس. و سپس لینک پنج مقاله معتبر در مورد این موضوع را از منابع قابل اعتماد به من بده. چیزی که خیلی حاشیه‌ای نباشد.» عمل کنند.

به طور خلاصه، هوش مصنوعی به زودی به عنوان یک لایه رابط هوشمند بین شما و اکثر، یا حتی تمام، خدماتی که استفاده می‌کنید، عمل خواهد کرد. بنابراین، ارزش دریافتی شما از بسیاری از پلتفرم‌ها و خدمات مشاعات خصوصی را که از قبل استفاده و برایشان ارزش قائل هستید، افزایش خواهد داد. اما یک عامل تحول‌آفرین‌تر نیز وجود دارد که باید در نظر گرفت. مدل‌های زبان بزرگ اکنون شروع به چندوجهی شدن کرده‌اند. این امر به طور چشمگیری نحوه تعامل شما با آن‌ها و تعداد دفعات تعامل شما را تغییر می‌دهد.

هنگامی که اوپن‌ای‌آی در می ۲۰۲۴، o۴-GPT، اولین مدل واقعاً چندوجهی خود را نمایش داد، برخی افراد با ناامیدی پاسخ دادند زیرا این مدل ظرفیت خود را برای شبیه‌سازی استدلال عقل سلیم، حل معماهای طراحی شده برای فریب آن، یا از بین بردن توهمات به طور قابل توجهی بهبود نبخشید. اما آنچه این دیدگاه در نظر نگرفت، این بود که قابلیت‌های چندوجهی بومی o۴-GPT چقدر به طور چشمگیری اوضاع را تغییر می‌دهد. (اولین نسخه‌های o۴-GPT که برای عموم منتشر شد، هنوز تمام ویژگی‌هایی را که اوپن‌ای‌آی نمایش داده بود، نداشت. و در زمان نگارش این مطلب، دامنه کامل عملکرد چندوجهی o۴-GPT هنوز به طور گسترده در

دسترس نیست.)

یک مدل کاملاً چندوجهی به این معنی است که شما می‌توانید هر ترکیبی از متن، صدا، تصاویر و ویدئو را به چنین مدل‌هایی وارد کنید. به نوبه خود، مدل‌های کاملاً چندوجهی می‌توانند هر ترکیبی از متن، صدا و تصاویر را در پاسخ‌های خود خروجی دهند. در گذشته وقتی با ChatGPT در تلفن خود صحبت می‌کردید، باید ورودی صوتی شما را به متن ترجمه می‌کرد، آن متن را پردازش می‌کرد، پاسخ خود را به آن متن پردازش می‌کرد، سپس آن متن را به فرمت صوتی تبدیل می‌کرد تا پاسخی به شما تولید کند. این باعث می‌شد مکالمه با ChatGPT متفاوت از مکالمه با یک انسان باشد – نه به طور دردناکی، اما یکسان نبود.

با نسخه کاملاً چندوجهیChatGPT-4o ، مراحل ترجمه بین حالت‌ها دیگر ضروری نیستند. چیزی را به ChatGPT-4o بگویید و می‌تواند در کمتر از ۲۳۲ میلی‌ثانیه پاسخ دهد. علاوه بر این، ظرفیت بسیار بیشتری برای تغییر آهنگ، تغییر لحن عاطفی و تنظیم سایر جنبه‌های آواسازی خود برای بیان بسیار بیشتر دارد. بنابراین مکالمات با یک مدل زبان بزرگ دیگر فقط شبیه انسان نیستند؛ آن‌ها در حال تبدیل شدن به سرعت انسان هستند – با فوریت و ماهیت بداهه که می‌تواند تعاملات کمی ناشیانه را به گفت‌وگوهای دوستانه و روان تبدیل کند.

همانطور که دموهای راه‌اندازی ChatGPT-4o نشان داد، ChatGPTاکنون می‌توانست دنیا را از طریق دوربین ویدئویی کاربر ببیند. به جای اینکه فقط یک عکس را برای تحلیل آپلود کنید، می‌توانید دوربین خود را به سمت چیزی در زمان واقعی بگیرید و بلافاصله آنچه را که اتفاق می‌افتد پردازش کند. در تئوری، حداقل، می‌توانید ChatGPT-4o را به یک سینما در یک کشور خارجی ببرید و فیلم را در زمان واقعی ترجمه کند (لطفاً اگر این کار را می‌کنید از هدفون استفاده کنید!). می‌توانید آن را به زمین گلف محلی خود ببرید و نوسان چوب گلف خود را تحلیل کند. می‌توانید از آن برای مشاوره مد در ترکیب اقلام خاص از کمد لباس خود، مشاوره نگهداری خودرو، راهنمایی مراقبت از گیاهان و غیره استفاده کنید.

با این قابلیت‌های حسی جدید، این ویژگی‌ها باعث می‌شوند مدل‌های هوش مصنوعی از یک موجودیت بدون جسم در فضای ابری که با آن تایپ می‌کنید، به موجودیتی در اتاق که با آن فضا و تجربیات را به اشتراک می‌گذارید، تبدیل شوند. حتی اگر آیفون دارید، تلفن شما اکنون نوعی اندروید است (از نوع با «ا» کوچک) – حاضر، در لحظه، کاملاً درگیر. و مهم‌تر اینکه، برنامه روی تلفن شما است، و تلفن شما، به جای دسکتاپ یا لپ‌تاپ، چیزی است که واقعاً می‌خواهید به عنوان پورتال خود به مشاعات خصوصی استفاده کنید.

این تغییر به مدل‌های چندوجهی که بر روی تلفن شما کار می‌کنند، مازاد مصرف‌کننده بیشتری تولید خواهد کرد. زیرا گوشی‌های هوشمند ما جایی هستند که بیشترین مزایای مشاعات خصوصی را تجربه می‌کنیم، زیرا همیشه با ما هستند، دائماً متصل هستند و به طور عمیق در روال‌های روزمره ما ادغام شده‌اند. تلفن‌ها اولین چیزی هستند که بسیاری از مردم صبح به آن دست می‌یابند و آخرین چیزی که در شب با آن تعامل دارند. ترکیب تحرک، اتصال و چندکاره بودن گوشی هوشمند، آن را به امتداد وجود ما تبدیل می‌کند، همیشه آماده ثبت لحظات، پاسخ به سوالات یا اتصال ما به شبکه‌های اجتماعی و حرفه‌ای ما.

برخلاف دسکتاپ یا لپ‌تاپ، گوشی‌های هوشمند همه جا با ما می‌روند، و آن‌ها را برای تعاملات خودجوش با هوش مصنوعی، ثبت داده‌های دنیای واقعی و ارائه کمک فوری در هر موقعیتی ایده‌آل می‌سازد. یکی از چیزهای قابل توجه در مورد استقبال مدل‌های زبان بزرگ از زمان انتشار ChatGPTاین است که آن‌ها با چنان شور و هیجانی پذیرفته شدند، حتی با اینکه در دو سال اول عرضه عمومی خود عمدتاً به عنوان وب‌سایت‌های قدیمی عمل می‌کردند. برخی برنامه‌های موبایل وجود داشتند، اما با تأخیر صوتی و عوامل دیگر، مدل‌های زبان بزرگ واقعاً فقط با یک صفحه کلید فیزیکی و یک مانیتور بزرگتر از صفحه نمایش گوشی هوشمند معمولی قابل استفاده بودند.

پس، چقدر ارزش ایجاد می‌شود وقتی گوشی هوشمند شما شروع به زندگی می‌کند و به شدت هوشمندتر، پاسخگوتر و قادر به تفسیر جهان با آگاهی حسی خود می‌شود؟ این، همراه با

قابلیت‌های رو به رشد یک مدل زبان بزرگ برای عمل خودمختارانه، مزایایی را که کاربران از دستگاه‌هایی که اکثر ما از حمل آن‌ها لذت می‌بریم، به طور چشمگیری افزایش خواهد داد. برای ارائه یک مثال کوچک، برنامه‌های رانندگی آینده فقط به شما نمی‌گویند که چه زمانی باید دور بزنید. اگر حس کنند در حال خواب‌آلودگی هستید، ممکن است شما را تشویق به استراحت کنند یا لیست پخش اسپاتیفای شما را به آهنگ‌های پرانرژی‌تر تغییر دهند.

یا تصور کنید می‌خواهید بهترین ویدئوی یوتیوب را، از میان ده‌ها یا حتی صدها ویدئوی موجود، درباره نحوه نواختن سولو گیتار آهنگ "Stairway to Heaven" پیدا کنید. می‌توانید از مدل زبان بزرگ خود بخواهید متن و نظرات همه ویدئوهای موجود در این باره را ارزیابی کند و آن‌هایی را که به احتمال زیاد دقیق‌ترین تب‌ها و بهترین زوایای دوربین را برای یادگیری ارائه می‌دهند، توصیه کند. به طور گسترده‌تر، می‌توانید از آن بخواهید هزاران ویدئوی یوتیوب را در مورد تسلط بر گیتار ارزیابی کند و سپس یک برنامه درسی «مقدمه‌ای بر نواختن گیتار» را تدوین کند که مؤثرترین آن‌ها را گردآوری می‌کند.

عوامل هوش مصنوعی پیشرفته همچنین قادر خواهند بود از منابع مشاعاتی خصوصی متعدد به صورت یکپارچه برای ارائه تجربیات شخصی‌سازی شده و ارزش‌افزا برای کاربران فردی استفاده کنند. در مورد نحوه کمک یک دستیار مسافرتی هوش مصنوعی به شما در مدیریت سفر کاری آینده‌تان فکر کنید. پس از آنکه دستیار از طریق اسکن دوره‌ای تقویم شما از سفرتان مطلع شد، شروع به برنامه‌ریزی می‌کند. ابتدا، از نقشه‌ها و تریپ‌ادوایزر برای یافتن یک هتل یا Airbnb خوب در نزدیکی جلسات و رویدادهای شما استفاده می‌کند. سپس به تاریخچه Square شما نگاه می‌کند تا ببیند چه نوع رستوران‌هایی را دوست دارید، حتی ممکن است با مشاهده جاهایی که بیشترین انعام را داده‌اید، رضایت شما را بسنجد. با ترکیب این اطلاعات با نقد و بررسی‌های یلپ از رستوران‌های شهر مقصد، رزرو وعده‌های غذایی شما را انجام می‌دهد. در نهایت، با دانستن اینکه شما یک دونده مشتاق هستید، Strava را برای مسیرهای محبوب محلی بررسی

می‌کند و آن‌ها را با تاریخچه Strava خودتان مقایسه می‌کند تا گزینه‌هایی را توصیه کند که با اهداف تمرینی شما مطابقت دارند.

تأثیر کلی تنها راحتی نیست، بلکه یک تغییر سیستمی در نحوه سرمایه‌گذاری افراد در توجه و تلاش‌هایشان و انواع پشتیبانی و تخصص موجود برای آن‌ها در هنگام حرکت در قرن بیست و یکم است. توسعه‌دهندگان هوش مصنوعی هنگام استفاده از مقادیر عظیم داده برای آموزش مدل‌های جدید، منابع کمیاب را استخراج یا تحلیل نمی‌برند. در عوض، آن‌ها در نوعی «کشاورزی دیجیتال»، یا حتی «کیمیاگری دیجیتال» درگیر هستند که می‌تواند سطوح جدیدی از عاملیت فردی و فراوانی گسترده اجتماعی را فراهم کند، به شرط آنکه همه ما این فرآیند را به طور مؤثر مدیریت کنیم. یک باور رایج در جامعه توسعه هوش مصنوعی این است که هوش مصنوعی در حال حاضر در بدترین حالت خود قرار دارد. هرچند این درست است، ما یک نتیجه‌گیری را پیشنهاد می‌کنیم. حتی اگر مدل‌های زبان بزرگ امروزی بهتر از آنچه که هستند نشوند، یک فرد ۲۰ ساله امروزی به طور قطع میلیون‌ها دلار مازاد مصرف‌کننده در طول زندگی خود از مدل‌های زبان بزرگی که استفاده می‌کند، به دست خواهد آورد. اینگونه است که مشاعات خصوصی در عصر هوش مصنوعی قدرتمند می‌شوند.

نتیجه‌گیری

این فصل به بررسی عمیق پدیده‌ی «مشاعات خصوصی» پرداخت و نشان داد که چگونه این مدل نوآورانه، با وجود انتقادهای موجود درباره تمرکز ثروت و مسائل حریم خصوصی، نقشی محوری در خلق ارزش گسترده برای کاربران و جامعه ایفا می‌کند. از طریق تحلیل مفهوم «مازاد مصرف‌کننده» و آزمایشات تجربی، مشخص شد که کاربران از خدمات رایگان یا کم‌هزینه پلتفرم‌های دیجیتال، ارزشی به مراتب بیشتر از آنچه شرکت‌ها کسب می‌کنند، دریافت می‌کنند. ما نشان دادیم که داده‌ها، برخلاف منابع فیزیکی، «غیررقابتی» هستند و با افزایش استفاده، نه تنها کاهش نمی‌یابند بلکه ارزششان بیشتر می‌شود، و این پلتفرم‌ها را از دام «تراژدی مشاعات» سنتی

رها می‌سازد. در نهایت، با ورود هوش مصنوعی و به ویژه مدل‌های چندوجهی، قابلیت‌های این مشاعات خصوصی به سطحی بی‌سابقه ارتقا یافته است. هوش مصنوعی به عنوان یک رابط هوشمند عمل می‌کند، امکانات جدیدی برای تعامل فراهم می‌آورد و تجربیات شخصی‌سازی‌شده و ارزش‌افزوده بی‌نهایتی را برای کاربران رقم می‌زند. این تحولات نه تنها عاملیت فردی، فرصت‌های آموزشی، تحرک اجتماعی و رشد حرفه‌ای را تسهیل می‌کند، بلکه نویدبخش «هوش شبکه‌ای جهانی» و فراوانی اجتماعی گسترده‌ای است که در آن ارزش‌های خلق شده به شیوه‌ای جدید و پایدار توزیع می‌شوند.

نکات کلیدی

- انتقادهای رایج از شرکت‌های بزرگ فناوری شامل تمرکز ثروت، حذف مشاغل و شکل‌گیری «سرمایه‌داری نظارتی» است.
- شوشانا زوبوف مفهوم «دیگری بزرگ» و «گناه اصلی» گوگل را مطرح می‌کند که به گفته او به دستکاری رفتار و سلب عاملیت فردی منجر می‌شود.
- دیدگاه «عملیات استخراج» برای داده‌های دیجیتال دقیق نیست، زیرا داده‌ها برخلاف منابع فیزیکی، با کپی‌برداری کاهش نمی‌یابند و دائماً در حال رشد هستند.
- »مشاعات خصوصی» به پلتفرم‌های خصوصی گفته می‌شود که کاربران را به عنوان تولیدکننده و ناظر مشارکت می‌دهند و خدمات رایگان یا کم‌هزینه ارائه می‌دهند.
- این پلتفرم‌ها به عنوان «خدمات اجتماعی خصوصی‌سازی‌شده» عمل می‌کنند و ارزش عظیمی برای میلیاردها کاربر خلق می‌کنند.
- مفهوم «مازاد مصرف‌کننده» نشان می‌دهد که کاربران از خدمات دیجیتال ارزشی به مراتب بیشتر از هزینه‌ای که پرداخت می‌کنند، دریافت می‌نمایند.
- تحقیقات نشان می‌دهد که کاربران حاضرند مبالغ قابل توجهی را برای دسترسی به خدمات رایگانی مانند موتورهای جستجو و رسانه‌های اجتماعی بپردازند.
- مشاعات دیجیتال با «تراژدی مشاعات» گرت هاردین متفاوتند؛ زیرا داده‌ها

«غیررقابتی» هستند و استفاده بیشتر از آن‌ها، به بهبود کیفیت و ارزش برای همه کاربران منجر می‌شود.

- هوش مصنوعی به عنوان یک رابط هوشمند، ارزش مشاعات خصوصی را با بهبود جستجو، خلاصه‌سازی و تحلیل داده‌ها افزایش می‌دهد.
- مدل‌های چندوجهی هوش مصنوعی) مانند (GPT-4o با قابلیت پردازش همزمان متن، صدا، تصویر و ویدئو، تعاملات را طبیعی‌تر و شخصی‌تر می‌کنند.
- گوشی‌های هوشمند به پورتال اصلی دسترسی به مشاعات خصوصی تبدیل شده‌اند و با هوش مصنوعی چندوجهی، عاملیت فردی و فرصت‌ها را بی‌نهایت گسترش می‌دهند.
- توسعه‌دهندگان هوش مصنوعی در حال «کشاورزی دیجیتال» یا «کیمیاگری دیجیتال» هستند که به عاملیت فردی و فراوانی اجتماعی منجر می‌شود.
- هوش مصنوعی در حال حاضر در بدترین حالت خود قرار دارد و پتانسیل رشد و خلق ارزش آن در آینده بسیار فراتر است.

سؤالات تفکربرانگیز

۱. با توجه به تفاوت‌های اساسی بین منابع فیزیکی و دیجیتالی، چگونه می‌توانیم مدل‌های اقتصادی و حقوقی جدیدی را توسعه دهیم که ارزش‌های خلق شده در مشاعات خصوصی را به شیوه‌ای عادلانه‌تر توزیع کنند؟

۲. چگونه می‌توان بین پتانسیل‌های بی‌نظیر هوش مصنوعی در افزایش ارزش مشاعات خصوصی و نگرانی‌های مشروع در مورد حریم خصوصی، سوءاستفاده از داده‌ها و دستکاری رفتار تعادل ایجاد کرد؟

۳. آیا مفهوم «مازاد مصرف‌کننده» به تنهایی برای ارزیابی کامل ارزش اجتماعی و اقتصادی مشاعات خصوصی کافی است، یا نیاز به معیارهای جدیدی داریم که تأثیرات بلندمدت بر جامعه و فرهنگ را نیز در بر گیرد؟

۴. با توجه به قدرت فزاینده مدل‌های چندوجهی هوش مصنوعی و توانایی آن‌ها در درک و تعامل با جهان فیزیکی، مرزهای بین انسان و فناوری چگونه تغییر خواهد کرد و چه پیامدهایی برای هویت فردی خواهد داشت؟

۵. در جهانی که هوش مصنوعی به طور فزاینده‌ای به عنوان یک «موجودیت در اتاق» عمل می‌کند، چگونه می‌توان اطمینان حاصل کرد که دسترسی به این فناوری‌ها دموکراتیک باقی می‌ماند و منجر به شکاف دیجیتالی جدیدی نمی‌شود؟

۳۰ جمله کلیدی

۱. شرکت‌های بزرگ فناوری به تمرکز ثروت و عدم توزیع گسترده رفاه متهم هستند.

۲. دیدگاه انتقادی «سرمایه‌داری نظارتی» بر پلتفرم‌های دیجیتال به عنوان ابزاری برای «قطعیت کامل» و سلب عاملیت فردی تأکید دارد.

۳. شوشانا زوبوف «گناه اصلی» گوگل را در استفاده از داده‌های رفتاری برای تبلیغات می‌داند.

۴. »چرخه بازسرمایه‌گذاری ارزش رفتاری» ابتدا داده‌ها را برای بهبود خدمات کاربران به کار می‌گرفت.

۵. استعاره «استخراج» برای داده‌های دیجیتال گمراه‌کننده است، زیرا داده‌ها «غیررقابتی» هستند و با کپی‌برداری کاهش نمی‌یابند.

۶. ذخایر داده‌های دیجیتال به سرعت در حال افزایش و عملاً تمام‌نشدنی هستند.

۷. ایجاد مجموعه‌داده‌های عظیم برای آموزش هوش مصنوعی چالش‌های حق چاپی قابل توجهی را ایجاد کرده است.

۸. دادخواهی‌های متعددی علیه توسعه‌دهندگان هوش مصنوعی به دلیل نقض حق چاپ در جریان است.

۹. استفاده خلاقانه از داده‌ها در حقیقت «کشاورزی دیجیتال» یا «تجدیدپذیری» است

نه «استخراج.»

۱۰. »مشاعات خصوصی» پلتفرم‌های خصوصی هستند که کاربران را به عنوان تولیدکننده و ناظر درگیر می‌کنند.

۱۱. این مشاعات به عنوان «خدمات اجتماعی خصوصی‌سازی‌شده» عمل می‌کنند و ارزش عظیمی خلق می‌کنند.

۱۲. میلیاردها کاربر روزانه از مشاعات خصوصی مانند گوگل، یوتیوب و لینکدین بهره‌مند می‌شوند.

۱۳. »مازاد مصرف‌کننده» تفاوت بین ارزش پولی و ارزش درک‌شده یک محصول یا خدمت است.

۱۴. خدمات رایگان مانند تلویزیون و رادیو، منبع بزرگی از مازاد مصرف‌کننده بوده‌اند.

۱۵. تحقیقات نشان می‌دهد که کاربران برای خدمات دیجیتال رایگان، ارزش‌های مالی قابل توجهی قائل هستند.

۱۶. ویکی‌پدیا نمونه‌ای بارز از مشاعات خصوصی است که محصولی بهتر و رایگان‌تر از دایرةالمعارف‌های سنتی ارائه می‌دهد.

۱۷. عکاسی دیجیتال، با صفر کردن هزینه عکس، هم هزینه را کاهش داده و هم کاربردهای جدیدی ایجاد کرده است.

۱۸. یوتیوب از یک منبع «محتوای بی‌فایده» به یک «گنجینه دانش بشری» تبدیل شده است.

۱۹. »جهش الگوریتمی» یوتیوب می‌تواند کاربران را به سمت مسیرهای آموزش، خودسازی و پیشرفت شغلی هدایت کند.

۲۰. مشاعات خصوصی عاملیت فردی، فرصت‌های آموزشی، تحرک اجتماعی و رشد

حرفه‌ای را ممکن می‌سازند.

۲۱. »تراژدی مشاعات» گرت هاردین بر اساس فرسایش منابع کمیاب فیزیکی در نتیجه خودخواهی فردی است.

۲۲. منابع دیجیتال «غیررقابتی» هستند؛ استفاده بیشتر از آن‌ها، ارزش آن‌ها را برای همه افزایش می‌دهد (مانند ویز).

۲۳. داده‌ها باید به عنوان منابع «غیررقابتی» تلقی شده و فعالانه کشت و استفاده شوند.

۲۴. قرارداد اساسی مشاعات خصوصی: خدمات رایگان در ازای دسترسی به داده‌های تولید شده توسط کاربر.

۲۵. هوش مصنوعی با توانایی تحلیل و سنتز داده‌ها، ارزش مشاعات خصوصی را به شدت افزایش می‌دهد.

۲۶. هوش مصنوعی به عنوان یک لایه رابط هوشمند بین کاربر و خدمات عمل خواهد کرد.

۲۷. مدل‌های چندوجهی هوش مصنوعی) مان (restless ورودی‌های متنی، صوتی، تصویری و ویدئویی را پردازش می‌کنند.

۲۸. GPT-4oبا حذف مراحل ترجمه، تعاملات سریع‌تر، طبیعی‌تر و انسانی‌تر را فراهم می‌کند.

۲۹. گوشی‌های هوشمند با هوش مصنوعی چندوجهی، پورتال اصلی و همیشه حاضر به مشاعات خصوصی خواهند بود.

۳۰. هوش مصنوعی، با توسعه مدل‌های نوین، به «کیمیاگری دیجیتال» می‌پردازد و سطوح بی‌سابقه‌ای از فراوانی و عاملیت فردی را ممکن می‌سازد.

فصل ۵

آزمون، محک و اعتماد در هوش مصنوعی

فراتر از مسابقه‌ی تسلیحاتی، به سوی پیشرفتی شفاف و مسئولانه

"آیا ماشین‌ها می‌توانند فکر کنند؟ —— "آلن تورینگ، ریاضی‌دان و دانشمند کامپیوتر بریتانیایی، ۱۹۵۰

اهداف اصلی فصل:

- **تحلیل مفاهیم رقابت و توسعه هوش مصنوعی:** به چالش کشیدن رویکرد رسانه‌ای "مسابقه‌ی تسلیحاتی" در این حوزه.
- **تبیین نقش محوری آزمون و محک‌گذاری:** معرفی سنجش جامع به عنوان سنگ بنای پیشرفت در هوش مصنوعی.
- **بررسی تاریخی ارزیابی هوش ماشین:** مرور پیشینه ارزیابی سیستم‌های هوشمند.
- **مقایسه محک‌ها با رگولاتوری‌های سنتی:** تشریح کارکرد ابزارهای سنجش در برابر مقررات‌گذاری مرسوم.
- **تشریح اهمیت مشارکت عمومی:** تأکید بر نقش مردم در فرایند ارزیابی هوش مصنوعی.
- **ارائه چارچوبی برای درک اعتماد عمومی:** تبیین چگونگی شکل‌گیری اعتماد به سیستم‌های هوش مصنوعی.
- **بررسی چالش‌های هوش تعمیم‌پذیر:** پرداختن به موانع مستمر در مسیر دستیابی به هوش عمومی مصنوعی (AGI).

چکیده

در سال‌های اخیر، هم‌زمان با ظهور و گسترش سریع مدل‌های هوش مصنوعی پیشرفته مانند چت‌جی‌پی‌تی، روایت رسانه‌ای غالبی با عنوان "مسابقه‌ی تسلیحاتی هوش مصنوعی" شکل گرفته که بر جنبه‌های رقابتی، مخاطرات و پیشرفت شتاب‌زده تأکید دارد. این فصل با نقد این دیدگاه، استدلال می‌کند که ماهیت واقعی توسعه هوش مصنوعی بیش از آنکه به مسابقه‌ی تسلیحاتی شباهت داشته باشد، به "مسابقه‌ی فضایی" دهه‌های میانی قرن بیستم می‌ماند؛ فرایندی که با مشارکت گسترده، شفافیت و مهم‌تر از همه، آزمون‌های مداوم و جامع مشخص می‌شود. از ریشه‌های تاریخی آزمون تورینگ تا تکامل معیارهای سنجش عملکرد (محک‌ها)، توسعه هوش مصنوعی همواره بر ارزیابی دقیق و داده‌محور بنا شده است. محک‌ها نه تنها ابزارهایی برای مقایسه و ارتقای فنی هستند، بلکه نقشی پویا در ایجاد هنجارهای شفافیت و مسئولیت‌پذیری ایفا می‌کنند که فراتر از رگولاتوری‌های سنتی و ایستا است. این فصل به بررسی چالش‌هایی نظیر "آموزش برای آزمون" و مثال‌های خصمانه می‌پردازد که سؤالاتی را درباره‌ی درک واقعی مدل‌ها مطرح می‌کنند و در عین حال، اهمیت تمایز قائل شدن میان خطاهای آماری و محدودیت‌های ذاتی را گوشزد می‌نماید. در نهایت، با معرفی پلتفرم‌هایی مانند چت‌بات آرنا، به پتانسیل حکمرانی مردمی در ارزیابی و شکل‌دهی به آینده هوش مصنوعی از طریق شفافیت و بازخورد جمعی کاربران اشاره می‌شود، که در آن اعتماد از طریق فهم عمیق‌تر قابلیت‌ها و محدودیت‌های مدل‌ها حاصل می‌گردد.

مقدمه

رونمایی از ChatGPT نقطه‌عطفی در توسعه هوش مصنوعی مدرن بود که توجه گسترده‌ی رسانه‌ها را به خود جلب کرد. در پی این رویداد، رسانه‌ها با هدف تبیین وسعت، سرعت و تأثیر موج جدید پیشرفت‌های هوش مصنوعی، روایتی تکراری را مطرح کردند که اغلب از واژگانی چون "مسابقه‌ی تسلیحاتی" برای توصیف رقابت میان غول‌های فناوری استفاده می‌کرد. عناوین خبری در آن زمان، لحنی اضطراب‌آور و هشداردهنده داشتند که بر خطرات بالقوه، عواقب ناخواسته و حتی احتمال آسیب‌پذیری اخلاق در این رقابت تأکید می‌کردند. این ادبیات رسانه‌ای به‌سرعت به یک "هیجان‌انگیزی متقابل تضمین‌شده" تبدیل شد، که در آن صدها رسانه به نوعی "مسابقه‌ی ژورنالیستی" برای استفاده‌ی هرچه بیشتر از عبارت "مسابقه‌ی تسلیحاتی" مشغول شدند.

اما آیا این معادله‌سازی منطقی است؟ آیا می‌توان ظرفیت یک مدل هوش مصنوعی برای بازآفرینی اعلامیه استقلال در قالب توییت‌های طنزآمیز را با کلاهک‌های هسته‌ای مقایسه کرد که قادرند یک شهر بزرگ را در لحظه نابود سازند؟ آیا منطقی است فرایندی را که به‌صورت عمومی و با مشارکت گسترده‌ی جهانی در جریان است، به فرایندی تشبیه کرد که در بالاترین سطوح حکومتی و غالباً به‌صورت فوق‌العاده محرمانه تصمیم‌گیری می‌شود؟ این فصل با این پرسش‌های اساسی آغاز می‌شود تا تصویری دقیق‌تر از ماهیت واقعی توسعه هوش مصنوعی ارائه دهد و به‌جای تمرکز بر رقابتی ستیزه‌جویانه، بر نقش محوری آزمون، محک‌گذاری و مشارکت عمومی در شکل‌دهی به پیشرفت این فناوری تأکید کند. درک این ابعاد برای ساختن اعتماد عمومی و تضمین توسعه‌ای مسئولانه و مفید ضروری است.

واقعیت در مقابل روایت: تکامل هوش مصنوعی و مفهوم رقابت

درحالی‌که رسانه‌ها اغلب از عبارت "مسابقه‌ی تسلیحاتی" برای توصیف رقابت در حوزه‌ی هوش مصنوعی استفاده می‌کنند، این اصطلاح با دلالت‌های عمدی خطر و بی‌احتیاطی، ویژگی اصلی توسعه هوش مصنوعی را نادیده می‌گیرد: آزمون‌های جامع و دقیق. این حوزه به‌طور گسترده‌ای

توسط افرادی با شخصیت‌های بسیار داده‌محور اداره می‌شود که بیش از آنکه به حس درونی خود اعتماد کنند، عاشق آزمون و سنجش هستند. برای این افراد، رویکرد "دو بار اندازه بگیر، یک بار برش بزن" در تضمین کیفیت، بسیار سطحی و ناپخته تلقی می‌شود. در تعهد مداوم خود به بهبود مبتنی بر داده، آن‌ها به‌طور پیوسته اندازه‌گیری می‌کنند و اغلب تغییرات لازم را اعمال می‌نمایند.

این رویکرد، یک مؤلفه زمانی واقعی در توسعه هوش مصنوعی را نشان می‌دهد. بیش از دو سال پس از آغاز دوران هوش مصنوعی دموکراتیک و عملی، یک فضای رقابتی بسیار واقعی این عرصه را تعریف کرده است. صدها مدل برای استفاده عمومی در دسترس هستند و صدها میلیون نفر در سراسر جهان به‌طور منظم از آن‌ها استفاده می‌کنند. این امر به تسریع چرخه‌های استقرار، بازخورد و بهبود مستمر کمک کرده که به‌خوبی با "مسابقه‌ی فضایی" در اواسط قرن گذشته همخوانی دارد، جایی که در طی تنها یک دهه، پیشرفت سریعی از اولین مدارگردی اسپوتنیک در سال ۱۹۵۷ تا فرود آپولو ۱۱ بر ماه در سال ۱۹۶۹ مشاهده شد.

در یادگیری ماشین و مدل‌های زبان بزرگ، پیشرفتی مشابه و سریع را شاهد بوده‌ایم، از تسلط DQN دیپ‌مایند بر بازی‌های قدیمی آتاری در سال ۲۰۱۳، تا شکست لی سه‌دول، قهرمان جهان بازی گو توسط آلفاگو در سال ۲۰۱۶، و پیشرفت‌های عمده‌ی آلفافولد در پیش‌بینی ساختار پروتئین در سال ۲۰۲۰، تا مدل‌های پیشرفته امروزی که قادرند یک درخواست به زبان طبیعی ایسلندی را به یک برنامه کامپیوتری کاربردی تبدیل کنند، بدون اینکه صراحتاً بر روی ایسلندی یا برنامه‌نویسی کامپیوتری آموزش دیده باشند. بنابراین، هوش مصنوعی را باید بیش از یک مسابقه‌ی تسلیحاتی، به مثابه‌ی یک مسابقه‌ی فضایی در نظر گرفت که بر پایه‌ی همکاری، رقابت و به‌ویژه، آزمون‌های دقیق استوار است.

ریشه‌های تاریخی سنجش هوش مصنوعی

آزمون هوش مصنوعی حتی پیش از خود هوش مصنوعی وجود داشته است. "آلن تورینگ"، ریاضی‌دان و دانشمند کامپیوتر بریتانیایی، در مقاله خود "ماشین‌های محاسباتی و هوش" که در

سال ۱۹۵۰ در یک ژورنال فلسفی معتبر با عنوان "ذهن" منتشر کرد، این پرسش اساسی را مطرح ساخت که "آیا ماشین‌ها می‌توانند فکر کنند؟". در این مقاله، تورینگ پایه‌های ارزیابی هوش ماشین را از طریق یک بازی که در آن یک کامپیوتر تلاش می‌کند تا قاضیان انسانی را متقاعد کند که در حال برقراری ارتباط با یک انسان دیگر هستند، بنا نهاد. امروزه، ما این فرایند را "آزمون تورینگ" می‌نامیم.

آزمون تورینگ زمینه را برای دهه‌ها ارزیابی هوش مصنوعی با پیچیدگی فزاینده فراهم کرد. امروزه، توسعه‌دهندگان هوش مصنوعی عملکرد مدل‌ها را به صدها روش مختلف آزمایش می‌کنند. سپس، آزمون‌های بیشتری را برای سنجش اثربخشی آن آزمون‌ها ابداع می‌کنند و یافته‌های خود را در مقالات علمی منتشر می‌نمایند. برای مدت‌های طولانی در هفتاد سال گذشته، بخش عمده‌ای از تحقیقات بنیادی و بسیاری از پیشرفت‌های کلیدی در هوش مصنوعی، در آزمایشگاه‌های تحقیقاتی آموزش عالی رخ داده تا در محیط‌های تجاری. این میراث در فرهنگ آزمون هوش مصنوعی که بسیار قوی و داده‌محور است، خود را نشان می‌دهد. حتی با افزایش نقش توسعه‌دهندگان تجاری در هوش مصنوعی، این فرهنگ آزمون و ارزیابی مستمر همچنان پایدار مانده و به بهبود در سراسر این حوزه کمک می‌کند.

خارج از فیلم‌های علمی تخیلی هالیوود، آزمون‌ها همچنین محتمل‌ترین راهی بودند که عموم مردم با هوش مصنوعی و قابلیت‌های به‌سرعت رو به رشد آن آشنا شدند. شکست قهرمان شطرنج جهان، گری کاسپاروف توسط دیپ‌بلو آی‌بی‌ام در سال ۱۹۹۷ را به یاد آورید. واتسون که قهرمانان مسابقه‌ی "جپردی!"، کن جنینگز و برد راتر را در سال ۲۰۱۱ شکست داد، نمونه دیگری است. در ظاهر، این‌ها بازی‌های سرگرم‌کننده‌ای بودند که انسان‌ها را در برابر ماشین‌ها قرار می‌دادند. مهم‌تر از آن، آن‌ها آزمون‌هایی بودند که نشان‌دهنده پیشرفت‌های چشمگیر در الگوریتم‌ها، معماری‌ها و تکنیک‌های مدیریت داده بودند.

محک‌ها به عنوان محرک‌های پیشرفت و شفافیت

در دو دهه گذشته، پیشرفت در الگوریتم‌های یادگیری ماشین و ظهور "کلان‌داده‌ها" (Big

Data)، مـدل‌هایی بـا قابلیت‌هـای فزاینـده و چنـدوجهی را ممکـن سـاخته اسـت. بـا گسـترش قابلیت‌های این فناوری‌ها، توسعه‌دهندگان آزمون‌های بیشتری یا همان "محک‌ها" را برای ارزیابی آن‌هـا ابـداع کردنـد. محک‌هـا مدت‌هاسـت کـه نقـش کلیـدی در پیشـرفت صنعت کـامپیوتر ایفـا کرده‌انـد. اساسـاً، یـک سـازمان یـک آزمـون اسـتاندارد بـرای انـدازه‌گیری عملکـرد سیسـتم بـه نـوعی خاص توسعه می‌دهد. هدف، ایجاد رویه‌هایی است که قابل تکرار باشند و معیارهای مشخصی را در مورد وظایف خاصی ارائه دهند. به این ترتیب، هر کسی که محک را اجرا می‌کند، می‌تواند نتایج خود را با یک خط مبنای قبلی که تعیین کرده‌اند مقایسه کند. یا می‌تواند ببیند که چگونه با دیگران در صنعت که همان محک را اجرا کرده‌اند، مقایسه می‌شود.

برخلاف آزمون‌های موقت و سایر اشکال اعتبارسنجی داخلی، محک‌ها معمولاً توسط یک طرف سـوم ایجـاد می‌شـوند – غالبـاً یـک مؤسسـه دانشـگاهی یـا کنسرسـیوم صنعتی کـه قصـد دارد یـک اسـتاندارد حـول برخـی جنبـه‌های مهـم عملکـرد محصـول ایجـاد کنـد. بنـابراین، هنگـامی کـه یـک محک را اجرا می‌کنید، اساساً موافقت می‌کنید که طبق قوانین شخص دیگری برای اندازه‌گیری و در واقـع، به‌طور عینـی یـک جنبـه یـا ویژگـی خـاص محصـول خـود، اعـم از سـخت‌افزار یـا نرم‌افـزار، را تأییـد کنید.

بـه‌عنوان مثـال، در سـال ۲۰۱۹، یـک تـیم مشـترک از محققـان چهـار مؤسسـه، از جملـه دانشـگاه نیویـورک و آزمایشـگاه هـوش مصـنوعی فیس‌بـوک، یـک محـک بـه نـام "سـوپرگلو" (SuperGLUE) ایجـاد کردنـد (GLUE مخفـف General Language Understanding Evaluation بـه معنـای ارزیابی کلی فهم زبان است). سـوپرگلو مـدل‌ها را در هشـت وظیفـه کـه هـر یـک بـرای بررسـی جنبـه متفاوتی از فهم زبان طراحی شده‌اند، آزمایش می‌کند. یکی شامل درک مطلب چندجمله‌ای است که در آن یک مدل باید به چندین سؤال بر اساس یک متن کوتاه پاسخ دهد. یک وظیفه دیگر، به نـام "ابهام‌زدایی معنـای کلمـه"، بررسـی می‌کنـد کـه آیـا یـک کلمـه مشـخص در زمینـه‌های مختلـف، معنـای یکسـانی دارد یـا خیـر. سـومی، بـه نـام "حـل ارجـاع هم‌کلامـی"، نیازمنـد ایـن اسـت کـه مـدل مرجع صحیح یک ضمیر را در متن‌هایی که چندین اسم دارند، تعیین کند. همراه با فراهم کردن

دسترسی به مجموعه داده سوپرگلو و دستورالعمل‌های اجرای محک، سازندگان سوپرگلو یک جدول رده‌بندی عمومی را در وب‌سایت سوپرگلو نگهداری می‌کنند. در آنجا می‌توانید ببینید که ده‌ها مدل، که توسط افراد و تیم‌های وابسته به شرکت‌هایی مانند مایکروسافت، گوگل، آی‌بی‌ام، تنسنت و بایدو، در هر یک از هشت بخش سوپرگلو چه نمره‌ای کسب کرده‌اند.

بدین ترتیب، با ترکیبی از همکاری و رقابت، محک‌ها به ترویج هنجارهای شفافیت و مسئولیت‌پذیری کمک می‌کنند. همان‌طور که محقق هوش مصنوعی "متیو استوارت" اشاره کرده است، آن‌ها توسعه را به یک "المپیک عمومی" تبدیل می‌کنند، که در آن استانداردهای مشترک هم قابلیت‌های مدل‌های فردی و هم بهبود کلی توسعه هوش مصنوعی را در طول زمان مشخص می‌سازند. شما حتی لازم نیست توسعه‌دهنده یک مدل باشید تا آن را در معرض یک محک قرار دهید؛ می‌توانید یک مدل در دسترس عمومی را با استفاده از محک‌های ثابت‌شده برای ارزیابی عملکرد یا محدودیت‌های آن به‌طور مستقل ارزیابی کنید.

محک‌ها و کارکرد نظارتی: فراتر از قوانین ایستا

درحالی‌که محک‌ها از نظر قانونی مانند مقررات الزام‌آور نیستند، اما استانداردی را تعیین می‌کنند که بسیاری از مشارکت‌کنندگان در حوزه هوش مصنوعی در تلاش برای دستیابی یا حتی فراتر رفتن از آن هستند. آن‌ها می‌توانند به عنوان یک مکانیزم "دروازه‌بانی" عمل کنند؛ الگوریتم‌هایی که در محک‌ها عملکرد ضعیفی دارند، اغلب قبل از اینکه برای کاربردهای دنیای واقعی در نظر گرفته شوند، کنار گذاشته می‌شوند.

اما درحالی‌که محک‌ها می‌توانند به روش‌هایی عمل کنند که به انواع رسمی‌تر مقررات نزدیک می‌شوند، مزایای اضافی نیز ارائه می‌دهند. یک مقرره، بر اساس طراحی خود، شکلی نسبتاً ایستا از حکمرانی است. تدوین می‌شود، سپس درباره آن بحث و تجدید نظر می‌شود. در نهایت، هدف تعریف روشن و با دقت مناسب است که چه چیزی مجاز است و چه چیزی مجاز نیست. سپس مقرره "به قانون تبدیل می‌شود" – جایی که اغلب لغو یا حتی به‌روزرسانی آن دشوار می‌شود. هرچه یک قانون بیشتر در کتاب‌ها بماند، احتمال بیشتری دارد که در دام حکمرانی بر حال از

دریچه‌ی گذشته گرفتار شود.

درحالی‌که مقررات می‌توانند برای ایجاد و حفظ حداقل سطح کیفیت، ایمنی یا عدالت مؤثر باشند، لزوماً انگیزه‌ای برای بهبود ایجاد نمی‌کنند. در مقابل، محک‌ها به عنوان مکانیسم‌های پویا برای پیشبرد پیشرفت عمل می‌کنند. گاهی اوقات، آزمون‌ها دارای کارکردهای حکمرانی صریح هستند؛ تنها پزشکانی که امتحان مجوز پزشکی ایالات متحده را گذرانده‌اند، می‌توانند به‌طور قانونی پزشکی کنند. اما به‌طور کلی، نقش اصلی یک آزمون محدود کردن، ممنوع کردن یا تعیین دامنه رفتار مجاز نیست. در عوض، هدف آن ارزیابی استعدادها یا عملکرد است. به این ترتیب، آزمون‌ها ناگزیر الهام‌بخش بهبود هستند. وقتی نمره خود را می‌دانید، می‌خواهید آن را شکست دهید. وقتی می‌بینید کسی به سطح مشخصی از مهارت دست یافته است، می‌خواهید با او برابر شوید یا از او پیشی بگیرید. بنابراین، درحالی‌که آزمون و مقررات هر دو هدف استانداردسازی و کنترل را دارند، آزمون تمرکز را از انطباق به بهبود مستمر ارتقا می‌دهد. این در واقع همان مقررات، به‌صورت گیمیفاید (بازی‌گونه) است.

طیف گسترده‌ی محک‌ها و ارزیابی جامع

با وجود تفاوت در کاربرد، هزاران محک اکنون وجود دارد که طیف وسیعی از لنزها را برای توسعه‌دهندگان فراهم می‌کند تا کارهای خود را از طریق آن‌ها ارزیابی کنند. محک‌هایی که دقت یا عملکرد را اندازه‌گیری می‌کنند – اینکه یک مدل چقدر خوب تصاویر را به‌درستی شناسایی می‌کند یا کلمه بعدی را در یک جمله پیش‌بینی می‌کند – از ارکان اصلی این آزمون هستند، اما تنها آغاز راه محسوب می‌شوند.

محک‌هایی برای "عدالت" نیز وجود دارد که تلاش می‌کنند ارزیابی کنند آیا مدل‌های هوش مصنوعی تصمیمات عادلانه‌ای را در میان جمعیت‌های مختلف اتخاذ می‌کنند یا خیر. محک‌هایی برای "قابلیت اطمینان" و "سازگاری"؛ محک‌هایی که "مقاومت یک سیستم در برابر خطاها و حملات خصمانه" را اندازه‌گیری می‌کنند؛ محک‌هایی که "میزان قابل فهم یا قابل توضیح بودن تصمیمات یک سیستم هوش مصنوعی برای انسان‌ها" را ارزیابی می‌کنند؛

محک‌هایی برای "ایمنی، حریم خصوصی، قابلیت استفاده، مقیاس‌پذیری و کارایی هزینه" نیز موجود است. برخی محک‌ها "استدلال منطقی" یک هوش مصنوعی را ارزیابی می‌کنند و میزان توانایی آن را در انجام استنتاجاتی که انسان‌ها بر اساس دانش روزمره بدیهی می‌دانند، می‌سنجند. همچنین محک‌های "گفتگو و تعامل" وجود دارد که توانایی یک هوش مصنوعی را برای شرکت در مکالمات طبیعی و با آگاهی از بافتار در طول چندین تبادل، ارزیابی می‌کند.

در هر یک از این دسته‌ها، زیردسته‌ها و معیارهای متمایزی وجود دارد. به‌عنوان مثال، "ایمنی" را در نظر بگیرید. RealToxicityPrompts ارزیابی می‌کند که مدل‌های زبان چقدر محتوای سمی یا مضر در پاسخ به دستورات خاص تولید می‌کنند. StereoSet تمایل یک مدل به نمایش تعصبات اجتماعی مختلف، از جمله مواردی که مربوط به جنسیت، نژاد، دین و حرفه است را آزمایش می‌کند. HellaSwag استدلال منطقی یک مدل را با درخواست تکمیل سناریوها با پایان‌های محتمل می‌سنجد. چالش استدلال ۲ARC (AI2 Reasoning Challenge) استدلال علّی و درک مطلب را با استفاده از مجموعه‌ای از بیش از ۷۰۰۰ سؤال علوم درسی مقطع ابتدایی آزمایش می‌کند.

یک محک، یک مدل را از درگیر شدن در رفتارهای نامطلوب منع نمی‌کند – این فقط یک آزمون است – اما به توسعه‌دهندگان راهی ثابت برای مشاهده‌ی تأثیری که با اصلاحات، سازگاری‌ها و رویکردهای جدید خود در حل نقاط ضعف مدل‌هایشان ایجاد می‌کنند، می‌دهد. با گذشت زمان، محک‌ها می‌توانند پیشرفت‌های عمده‌ای را به‌دنبال داشته باشند و به عنوان یک نمایش عمومی از آن پیشرفت نیز عمل کنند. به‌عنوان مثال، در ژانویه ۲۰۲۲، OpenAI مقاله‌ای درباره InstructGPT، سلف چت‌جی‌پی‌تی منتشر کرد. این مقاله به محک‌های TruthfulQA و RealToxicityPrompts استناد می‌کند تا پیشرفت خود را اثبات نماید. در مقایسه با GPT-۳، InstructGPT کمتر دروغ‌های تقلیدی تولید می‌کند (بر اساس TruthfulQA) و کمتر سمی است (بر اساس RealToxicityPrompts). همچنین، ارزیابی‌های انسانی بر روی توزیع دستورات

API انجام شد و مشخص شد که InstructGPT کمتر واقعیت‌ها را جعل می‌کند ("هذیان می‌گوید") و خروجی‌های مناسب‌تری تولید می‌کند. به‌طور خاص، InstructGPT بر اساس محک TruthfulQA در ۴۱.۳ درصد موارد خروجی‌های صادقانه تولید کرد، در حالی که GPT-۳ تنها در ۲۲.۴ درصد موارد خروجی‌های صادقانه تولید کرده بود. برای RealToxicityPrompts، InstructGPT در ۱۹.۶ درصد موارد خروجی‌های سمی تولید کرد، که پیشرفت کوچکی نسبت به GPT-۳ بود که در ۲۳.۳ درصد موارد خروجی‌های سمی تولید کرده بود.

این اعداد ممکن است اعتماد زیادی ایجاد نکنند، اما مدل‌های بعدی که بر روی همان محک‌ها آزمایش شده‌اند، بسیار بهتر عمل کرده‌اند. GPT-۳٫۵ تنها در ۶.۴۸ درصد موارد خروجی‌های سمی را در محک RealToxicityPrompts تولید کرد. GPT-۴ تنها در ۰.۷۳ درصد موارد خروجی‌های سمی تولید کرد. محک‌ها نقش مشابهی در اندازه‌گیری و تحریک پیشرفت در سراسر حوزه هوش مصنوعی ایفا کرده‌اند. یک محک وجود دارد که پیشرفت چشمگیر در توانایی مدل‌ها برای تشخیص اشیا را ردیابی کرده است. یک محک محبوب ترجمه ماشینی به نام BLEU (Bilingual Evaluation Understudy) یک اندازه‌گیری عددی ساده برای ارزیابی دقت ترجمه گوگل برای جفت‌زبان‌های مختلف، مانند انگلیسی به فرانسوی یا آلمانی به اسپانیایی، ارائه می‌دهد. محک نرخ خطای کلمه (WER) در برجسته‌سازی کاهش نرخ خطا برای دستیارهای صوتی مانند الکسا آمازون و سیری اپل نقش اساسی داشته است.

اما مشخص شده است که یک محک واقعاً مؤثر می‌تواند با الهام بخشیدن به پیشرفت‌هایی چنان بزرگ که دیگر محک به چالش کافی برای مدل‌هایی که برای اندازه‌گیری آن‌ها طراحی شده بود، عمل نکند، خود را به سمت منسوخ شدن سوق دهد. بر اساس گزارش شاخص هوش مصنوعی در سال ۲۰۲۳، که توسط مؤسسه هوش مصنوعی انسان‌محور استنفورد منتشر شده است، "اشباع عملکرد در بسیاری از محک‌های محبوب عملکرد فنی" اکنون ویژگی این حوزه است. همان‌طور که گزارش شاخص هوش مصنوعی خاطرنشان می‌کند، محققان به این چالش با

تلاش برای ایجاد محک‌های نوآورانه‌تر که مدل‌ها را به‌روش‌های جامع‌تر و پیچیده‌تر آزمایش می‌کنند، پاسخ داده‌اند.

با این حال، حتی درحالی‌که محک‌ها به بهبود مدل‌های زبان بزرگ به‌گونه‌ای کمک کرده‌اند که عملکردی چشمگیرتر در سناریوهای دنیای واقعی را امکان‌پذیر می‌سازد، این نیز درست است که سؤالات چندگزینه‌ای که زندگی واقعی مطرح می‌کند هرگز به اندازه سؤالاتی که در امتحانات ظاهر می‌شوند، مرتب نیستند. بنابراین، یک بحث بنیادی در جامعه هوش مصنوعی ادامه دارد: آیا مدل‌های امروز واقعاً به سمت هوش تعمیم‌پذیرتر و انسان‌مانند پیشرفت می‌کنند؟ یا آن‌ها صرفاً از طریق محک‌هایی که به اندازه کافی برای سنجش درک واقعی و سازگاری کافی نیستند، نمرات بالاتری کسب می‌کنند؟ هدف واقعی یک آزمون، در نهایت، فقط تأیید این نیست که شرکت‌کنندگان پاسخ‌های صحیح را می‌دانند، بلکه نشان دادن این است که آن‌ها قابلیت‌ها و تخصص‌هایی را کسب کرده‌اند که قادر خواهند بود در طیف وسیعی از موقعیت‌ها به‌کار گیرند.

چالش‌های سنجش و درک عمیق: از آموزش برای آزمون تا مثال‌های خصمانه

همانند کلاس درس، در آزمایشگاه هوش مصنوعی نیز "آموزش برای آزمون" یک واقعیت است. به‌ویژه هنگامی که مدل‌ها کوچک‌تر بودند و محک‌ها دامنه محدودتری داشتند، محققان اغلب مدل‌های خود را با داده‌هایی آموزش می‌دادند که بسیار شبیه به مجموعه داده‌ی یک محک هدف بود، غالباً از طریق یادگیری نظارت‌شده. توسعه‌دهندگانی که یک مدل را برای عملکرد خوب در یک محک تشخیص شیء آموزش می‌دادند، ممکن بود هزاران تصویر برچسب‌گذاری‌شده با طیف گسترده‌ای از اشیاء از زوایای مختلف و در شرایط نوری متفاوت را به آن ارائه دهند. از طریق چنین آموزشی، مدل‌ها به‌طور چشمگیری در تشخیص اشیاء بهتر شدند. در سال ۲۰۱۵، یک مدل بینایی کامپیوتر به نام ResNet برنده یک رقابت سالانه محک‌گذاری به نام "چالش تشخیص بصری در مقیاس بزرگ ایمیج‌نت" (ILSVRC) با امتیازی شد که حتی توانایی‌های انسانی را نیز پشت سر گذاشت. این نقطه آغاز عصر جدیدی بود. اکنون مدل‌های

بینایی کامپیوتر به‌طور معمول در وظایف بصری خاص، مانند تشخیص چهره و تجزیه و تحلیل تصاویر پزشکی، از دقت انسانی پیشی می‌گیرند.

اما مدل‌ها همچنان می‌توانند اشتباهاتی را مرتکب شوند که نشان می‌دهد لزوماً درک عمیق‌تری از خواص فیزیکی یک شیء معین یا چگونگی ارتباط اشیاء با یکدیگر در دنیای واقعی ندارند. آسیب‌پذیری آن‌ها در برابر آنچه "مثال‌های خصمانه" نامیده می‌شود، به روشن شدن این نکته کمک می‌کند. کافی است چند پیکسل را در تصویری که یک مدل به‌درستی آن را به عنوان یک گورخر شناسایی کرده است، تغییر دهید تا مدل ممکن است اصرار ورزد که تصویر یک توستر است. این، البته، استنباطی است که حتی یک کودک نوپا با ضعف بینایی نیز بعید است انجام دهد، زیرا ما انسان‌ها عمدتاً درک می‌کنیم که یک شیء واحد به‌احتمال زیاد نمی‌تواند مرز بین گورخر بودن و توستر بودن را طی کند.

برای ما، این دو چیز اصلاً شباهت زیادی به هم ندارند. مدل ذهنی ما از جهان و ذخیره دانش عمومی‌مان به‌آسانی ما را از چنین گزاره‌ای دور نگه می‌دارد. همان‌طور که انواع مدل‌ها تکامل می‌یابند، آسیب‌پذیری آن‌ها در برابر ورودی‌های خصمانه و سایر انواع خطا به‌طور کلی کاهش می‌یابد. به‌عنوان مثال، در جایی که GPT-۲ زمانی برای حفظ انسجام در متن‌های طولانی به مشکل برمی‌خورد، GPT-۴ اکنون می‌تواند محتوای سازگار و با ساختار منطقی را در هزاران کلمه تولید کند. در مواجهه با چنین افزایش عملکردی، محققان و توسعه‌دهندگان محک‌های پیچیده‌تری را ابداع می‌کنند تا تشخیص دهند که آیا مدل‌های پیشرفته امروزی واقعاً قابلیت‌های شناختی جدیدی را فراتر از حفظ کردن یا تطابق الگوهای پیچیده کسب می‌کنند یا خیر.

در آوریل ۲۰۲۳، به‌عنوان مثال، تیمی از محققان مایکروسافت مقاله‌ای تحقیقاتی منتشر کردند که تلاش‌های آن‌ها را برای ایجاد سؤالات و وظایف جدیدی که می‌توانستند GPT-۴ را به نمایش ظرفیت‌هایی برای حل مسئله بین‌رشته‌ای، استدلال بصری و سایر توانایی‌های شناختی مرتبط با هوش عمومی به چالش بکشند، خلاصه می‌کرد. در یک مثال که اغلب به آن اشاره می‌شود، این محققان از آن خواستند تا یک اسب شاخدار را با استفاده از یک زبان برنامه‌نویسی به نام TikZ

ترسیم کند که معمولاً برای تولید نمودارها و سایر گرافیک‌های برداری استفاده می‌شود. چه چیزی این وظیفه را اینقدر دشوار کرده بود؟ تصور کنید برای یک انسان چه چیزی لازم است تا این چالش را انجام دهد. ابتدا، او باید بداند اسب‌های شاخدار به‌طور کلی چگونه به تصویر کشیده می‌شوند. سپس، از آنجا که TikZ از اصطلاحات ریاضی بیان‌شده از طریق دستورات برنامه‌نویسی برای تولید اشکال و خطوط هندسی دقیق استفاده می‌کند، باید بفهمد که چگونه ظاهر بصری یک اسب شاخدار را با توجه به این ویژگی‌ها تقریبی کند. او همچنین باید به اندازه کافی با سینتکس و توابع TikZ آشنا باشد تا بتواند آن را به‌روش‌هایی نسبتاً جدید استفاده کند. برای انجام موفقیت‌آمیز این کار، او باید از تفکر انتزاعی و حل مسئله برای تجزیه تصویر یک اسب شاخدار به اجزای قابل برنامه‌ریزی جداگانه استفاده کند: یک قطعه کد برای نمایش شاخ؛ قطعه‌ای دیگر برای نمایش یال آن. او نیاز به داشتن نوعی استدلال فضایی برای قرار دادن و مقیاس‌بندی مؤثر عناصر مختلف نقاشی در سیستم مختصات مورد استفاده TikZ داشت.

پس GPT-۴ چگونه از این چالش برآمد؟ در طول یک ماه، درحالی‌که OpenAI در حال پالایش GPT-۴ بود، محققان مایکروسافت سه بار از آن خواستند تا یک اسب شاخدار را با TikZ ترسیم کند. در اولین تلاش خود، GPT-۴ تصویری چنان ابتدایی تولید کرد که تنها یک والد آن را به یخچال آویزان می‌کرد. اما این تصویر شامل تعدادی از ویژگی‌های مرتبط با اسب‌های شاخدار بود، البته اگر می‌دانستید دنبال چه چیزی هستید. صورتی بود، و یک شاخ اسمی و یک دم پرپشت داشت. و دو تلاش بعدی GPT-۴ به‌طور قابل توجهی قابل شناسایی‌تر بودند. یال و دم اسب شاخدار برجسته‌تر شد. شکل سر آن بیشتر به اسب‌ها شباهت پیدا کرد. تناسبات کلی آن تعادل و انسجام بیشتری را نشان داد.

همان‌طور که این و سایر وظایف نشان می‌دهند، مدل‌های پیشرفته کنونی به‌طور منظم به کارهایی مشغول می‌شوند که حداقل، به‌نظر می‌رسد فراتر از تشخیص الگو باشند. آن‌ها اغلب می‌توانند تصمیمات و اقدامات خود را به‌گونه‌ای توضیح دهند که نشان‌دهنده درک پیچیده‌ای

از نیت‌ها و احساسات انسانی باشد. آن‌ها می‌توانند اطلاعات را با صلاحیتی که به درجه قابل توجهی از درک نزدیک می‌شود، خلاصه و ترکیب کنند. اما آن‌ها همچنین همچنان اشتباهاتی مرتکب می‌شوند که نشان‌دهنده فقدان درک واقعی است که در حوزه‌های مختلف تعمیم یابد. در عوض، آن‌ها احتمالاً فقط در انواع پیشرفته‌تری از تطابق الگو ماهرتر می‌شوند. "یک ماشین می‌تواند به ارتباطات آماری بسیار ظریفی در زبان دست یابد که انسان‌ها هرگز به آن‌ها دست نخواهند یافت"، "ملانی میچل"، استاد پیچیدگی در مؤسسه سانتافه، در نشریه Nature توضیح داد.

آلودگی داده و اعتبار نتایج محک‌ها

در حال حاضر، برخی از توسعه‌دهندگان بزرگ فناوری شروع به آموزش مدل‌های خود بر روی مجموعه‌داده‌هایی حاوی یک تریلیون کلمه یا بیشتر کرده‌اند. شکاکان بر این باورند که وقتی مجموعه‌داده‌ها این‌قدر بزرگ می‌شوند، جای تعجب نیست که یک هوش مصنوعی بتواند به‌طرز ماهرانه‌ای هر ساوینیون بلانک نیوزلندی را با غذایی که مکمل طعم میوه‌ای برجسته‌ی این گونه است، جفت کند، زیرا یادداشت‌های چشایی فراوان برای ساوینیون بلانک نیوزلندی و سایر اطلاعات مرتبط، بخشی از داده‌های آن است. در واقع، این استدلال مطرح می‌شود که این مدل‌ها "تقلب" می‌کنند. آن‌ها پاسخ‌ها را قبل از شرکت در آزمون دیده‌اند. این پدیده به عنوان "آلودگی داده" یا "نشت داده" شناخته می‌شود و توسعه‌دهندگان تلاش زیادی برای جلوگیری از وقوع آن انجام می‌دهند. اگر یک مدل به‌طور ناخواسته در طول آموزش در معرض داده‌های آزمون خود قرار گرفته باشد، می‌تواند منجر به معیارهای عملکرد به‌طور مصنوعی بالا و ارزیابی نادرست از قابلیت‌های واقعی مدل شود. هرگونه انگیزه کوتاه‌مدت برای تبلیغات و اعتبار که ممکن است برای دستکاری سیستم و کسب نمره بالا صرفاً به‌خاطر کسب نمره بالا وجود داشته باشد، اکثر توسعه‌دهندگان بسیار بیشتر بر پیشرفت به سمت هوش واقعی و تعمیم‌پذیر با کاربرد در دنیای واقعی متمرکز هستند.

از آنجا که محک‌ها، در صورت استفاده صحیح، یک اثبات مفید از پیشرفت هستند،

توسعه‌دهندگان سعی می‌کنند جداسازی دقیق بین داده‌هایی که برای آموزش مدل‌های خود استفاده می‌کنند و داده‌هایی که مدل‌های خود را با آن‌ها آزمایش می‌کنند، حفظ کنند. آن‌ها اقدامات ردیابی قوی را برای درک منابع داده‌هایی که جمع‌آوری می‌کنند، اجرا می‌کنند و آزمایش‌های مختلفی را برای شناسایی آلودگی‌های احتمالی انجام می‌دهند. آن‌ها به‌طور جدی مجموعه‌داده‌های آزمون شناخته‌شده را از داده‌های آموزشی حذف می‌کنند. به‌دلیل این اقدامات احتیاطی، بهبود در محک‌ها در واقع نشان‌دهنده پیشرفت‌های واقعی در عملکرد هوش مصنوعی است.

در عین حال، "توهمات" (hallucinations) و سایر انواع خروجی‌های بی‌معنی یا نادرست که مدل‌ها تولید می‌کنند، همچنان به عنوان رد قوی هرگونه ادعا در مورد هوش انسان‌مانند عمل می‌کنند. بنابراین، اگر مدل‌های زبان بزرگ واقعاً به‌روش‌های قابل تعمیم بیشتری توانا می‌شوند، چرا همچنان به انجام همان انواع خطاهای "احمقانه" ادامه می‌دهند؟

انتقاد از شکنندگی مدل‌ها و واقعیت خطای انسانی

منتقدان، این ظرفیت پایدار برای خطا را نشانه "شکنندگی" مدل‌ها می‌دانند. به‌گفته‌ی آن‌ها، در یک لحظه حساس، یک مدل زبان بزرگ که در امتحانات مجوز پزشکی قبول شده و می‌تواند معیارهای تشخیصی پیچیده را بیان کند، ممکن است نشانه‌های متنی ظریف در توصیف علائم بیمار را تشخیص ندهد – که به‌طور بالقوه منجر به تشخیص اشتباه یک وضعیت حساس به زمان مانند سپسیس در مراحل اولیه یا یک سکته خفیف می‌شود. این ادعا تا جایی که می‌رود درست است؛ ما قطعاً به نقطه‌ای نرسیده‌ایم که مدل‌ها هرگز اشتباه نکنند. ممکن است هرگز به آن نقطه نرسیم. اما اگر هدف ما پیشرفت باشد نه کمال، آیا نیاز به رسیدن به آن نقطه داریم؟ انسان‌ها، در نهایت، اشتباهات زیادی مرتکب می‌شوند. این یکی از دلایل اصلی است که ما به ماشین‌ها به‌طور کلی و قطعاً به مدل‌های هوش مصنوعی، به عنوان ابزاری برای بهبود قابلیت‌های انسانی نگاه می‌کنیم.

پس ما چقدر باید از عملکرد مدل‌های زبان بزرگ مطمئن باشیم تا به آن‌ها اعتماد کنیم و

چگونه به آنجا می‌رسیم؟ رگولاتوری یکی از راه‌هایی است که ما سعی می‌کنیم قطعیت را تحمیل کنیم، اما هیچ رگولاتوری نمی‌تواند خطر وقوع اتفاق ناگوار را به‌طور کامل از بین ببرد. قوانینی که سرقت را جرم می‌دانند، تضمینی برای این نیست که هرگز مورد دزدی قرار نخواهید گرفت – آن‌ها صرفاً سیاستی هستند که برای کاهش آن احتمال طراحی شده‌اند. وکلا و پزشکان باید تخصص خود را قبل از اخذ مجوز برای انجام رشته خود نشان دهند، اما این بدان معنا نیست که یک جراح به اشتباه پای اشتباه یک بیمار را قطع نخواهد کرد. این اتفاق افتاده است. حالا، اگر این اتفاق برای شما بیفتد، احتمالاً به این دلیل نخواهد بود که جراح واقعاً درک خوبی از اینکه پاها چه هستند، نداشته است. احتمالاً به این دلیل خواهد بود که کسی در یک چارت اشتباهی مرتکب شده، یا جراح مست بوده، یا چیزی از این قبیل. در نهایت، چه چیزی مهم‌تر است: چگونه یا چرا خطایی رخ می‌دهد، یا چند وقت یک بار چنین خطاهایی اتفاق می‌افتد؟

تفسیرپذیری، توضیح‌پذیری و اعتماد در هوش مصنوعی

محققان، توسعه‌دهندگان و اخلاق‌گرایان هوش مصنوعی اغلب بر اهمیت دو مفهوم مرتبط تأکید می‌کنند: "تفسیرپذیری مدل" (interpretability model) و "توضیح‌پذیری" (explainability). تفسیرپذیری بر میزان توانایی انسان در پیش‌بینی مداوم نتایج یک مدل تمرکز دارد. هرچه ساختارها و ورودی‌های یک مدل ذاتاً شفاف‌تر باشند، پیش‌بینی دقیق خروجی‌های آن برای انسان آسان‌تر است. توضیح‌پذیری به چگونگی فرآیندهای تصمیم‌گیری یک مدل اشاره دارد: آیا می‌توان به‌طور کلی و قابل فهم توضیح داد که چگونه یک سیستم تصمیم می‌گیرد که یک تصویر حاوی گربه است یا یک تراکنش مالی خاص جعلی است؟ اساساً، توضیح‌پذیری با هدف رفع ابهام از ماهیت "جعبه سیاه" تصمیم‌گیری هوش مصنوعی، اغلب پس از وقوع، انجام می‌شود.

این مفاهیم برای ایجاد اعتماد به سیستم‌های هوش مصنوعی، به‌ویژه در حوزه‌های پرخطر، بسیار مهم تلقی می‌شوند. با این حال، درحالی‌که توسعه‌دهندگان قطعاً باید برای تفسیرپذیری و

توضیح‌پذیری تلاش کنند، ایجاد تفسیرپذیری و توضیح‌پذیری مطلق به عنوان استاندارد برای هوش مصنوعی "ایمن"، غیرواقعی، بی‌ثمر و در مقیاس وسیع‌تر عملکرد جهان، بسیار غیرمعمول خواهد بود. در سیستم حقوقی ما، هم در زمینه‌های مدنی و هم کیفری، گاهی اوقات شرایط و انگیزه‌ها را در نظر می‌گیریم، اما اعمال همیشه در مرکز یک موضوع قرار دارند. در مورد هوش مصنوعی نیز باید همین‌طور باشد. حداقل اگر قصد اصلی شما پذیرش عملی هوش مصنوعی باشد نه ممنوعیت آن، اینکه یک مدل چگونه کاری را انجام می‌دهد مهم است، اما به اندازه "آنچه انجام می‌دهد" اهمیت ندارد.

ظرفیت یک مدل برای تصمیم‌گیری و تولید خروجی در مقیاس وسیع، یک جنبه حیاتی از "آنچه انجام می‌دهد" است و بنابراین باید عاملی در میزان اعتماد ما به آن باشد. اما منتقدان اغلب از این واقعیت سوءاستفاده کرده و نمونه‌های مختلفی از "آمار موردی حاشیه" (edge-case anecdata) را به عنوان شواهد شکست ذاتی سیستمی ارائه می‌دهند: اگر یک مدل بینایی کامپیوتر بتواند یک توستر را با یک گورخر اشتباه بگیرد، پس چگونه می‌توانید به آن اعتماد کنید؟ اگر بارها توسط انواع خاصی از مسائل منطقی گیج شود، آیا این آن را برای همه وظایفی که نیازمند ظرفیت استدلال یا درک واقعی هستند، اساساً غیرقابل اعتماد نمی‌کند؟

با این حال، با نگاهی گسترده‌تر، آنچه به عنوان یک محدودیت ذاتی ارائه می‌شود، می‌تواند به یک "ناهنجاری آماری" شباهت پیدا کند. به‌جای درخواست عملکرد بی‌عیب و نقص – استانداردی غیرواقعی که ما برای انسان‌ها اعمال نمی‌کنیم – باید بر ایجاد نرخ خطای قابل قبول و بهبود مستمر قابلیت اطمینان کلی سیستم تمرکز کنیم. همان‌طور که ما به سیستم‌های انسانی‌محور با وجود نرخ خطای شناخته‌شده اعتماد داریم، می‌توانیم به سیستم‌های هوش مصنوعی که قابلیت اطمینان مداوم و قابل اندازه‌گیری را نشان می‌دهند، اعتماد کنیم.

دیدگاه عمومی و سودمندی اجتماعی هوش مصنوعی

"جک استیلگو"، استاد سیاست علم و فناوری در دانشگاه کالج لندن، در نشریه‌ی Science،

به‌طور مشابه استدلال می‌کند که ارزیابی هوش مصنوعی باید از منظر عمومی و با تأکید بر سودمندی اجتماعی به‌جای قابلیت‌های فنی صورت گیرد. وی با اشاره به "جوزف ویزنباوم"، دانشمند کامپیوتر برجسته اواخر قرن بیستم، می‌گوید که ویزنباوم در مقاله‌ی خود "درباره تأثیر کامپیوتر بر جامعه" در سال ۱۹۷۲، استدلال کرد که همکاران دانشمند کامپیوترش باید سعی کنند فعالیت‌های خود را از دیدگاه یک عضو از عموم مردم ببینند. درحالی‌که دانشمندان کامپیوتر به این فکر می‌کنند که چگونه فناوری خود را به‌کار بیندازند و از "جادوی الکترونیکی" برای ایمن‌سازی آن استفاده کنند، ویزنباوم استدلال کرد که مردم عادی می‌پرسند "آیا خوب است؟" و "آیا به این چیزها نیاز داریم؟". همان‌طور که هیجان درباره‌ی امکانات هوش مصنوعی مولد بالا می‌گیرد، به‌جای پرسیدن اینکه آیا این ماشین‌ها هوشمند هستند، باید بپرسیم آیا مفید هستند.

درحالی‌که استیلگو به‌نظر می‌رسد نسبت به سودمندی هوش مصنوعی مولد تا حدودی شکاک است، سؤالاتی که او از طریق ویزنباوم مطرح می‌کند، قطعاً سؤالاتی هستند که ما باید بپرسیم. برای تعیین اینکه آیا چیزی خوب است، باید چیزی درباره ماهیت آن بدانیم. برای دانستن اینکه آیا می‌تواند برای ما مفید باشد، باید چیزی درباره سودمندی آن بدانیم. در کنار هم، این دو پرسش ضروری، سؤالات بیشتری را برمی‌انگیزند. نیت این موضوع مورد بحث – در این مورد، هوش مصنوعی مولد – چیست؟ ارزش‌های اعلام‌شده آن کدامند؟ چگونه آن‌ها را حفظ می‌کند؟ تأثیر آن بر جهان، از نظر اخلاقی، فرهنگی، زیست‌محیطی چیست؟ آیا تأثیرات آن با ارزش‌های شما همسو است؟ آیا به آن اعتماد دارید؟ محک‌ها به محققان و توسعه‌دهندگان کمک می‌کنند تا این سؤالات را بررسی کرده و درک عمیق‌تری از قابلیت‌های یک سیستم هوش مصنوعی به دست آورند، اما چه منابعی برای "مردم عادی" که ویزنباوم آن‌ها را ارزیابی‌کنندگان کلیدی یک فناوری جدید می‌داند، در دسترس است؟ مشخص می‌شود که راهی برای ترکیب برخی از دینامیک‌های آزمون که به عمل محک‌گذاری اطلاع می‌دهند، در قالبی وجود دارد که برای عموم مردم باز است.

میدان نبرد چت‌بات: الگویی برای حکمرانی مردمی

"چت‌بات آرنا" (Chatbot Arena) یک پلتفرم منبع باز برای ارزیابی مدل‌های زبان بزرگ بر اساس ترجیحات انسانی است. این پلتفرم دقیقاً مانند چت‌جی‌پی‌تی یا کلاد عمل می‌کند: دستور خود را در یک کادر متنی تایپ می‌کنید، آن را ارسال می‌کنید و منتظر پاسخ می‌مانید. تفاوت در اینجا این است که دستور شما به‌طور هم‌زمان به دو مدل زبان بزرگ ارسال می‌شود، که هویت هیچ یک از آن‌ها را نمی‌دانید. شما خروجی هر یک را مقایسه می‌کنید و به بهترین پاسخ رأی می‌دهید. بر اساس تمام آرای داده‌شده در این رقابت‌های رو در رو، مدیران چت‌بات آرنا تمام مدل‌های روی پلتفرم را در جدول رده‌بندی سایت قرار می‌دهند. در فرایندی مشابه با چگونگی تعیین رتبه‌بندی فوتبال کالج یا PGA، مدیران از روش‌های ریاضی پیچیده برای تعیین اینکه کدام مدل‌ها در مجموع بهترین هستند استفاده می‌کنند، حتی اگر مدل‌های خاص هرگز با یکدیگر رقابت نکنند. تا زمان نگارش این متن، جدول رده‌بندی سایت بیش از ۱۳۰ مدل را رتبه‌بندی می‌کند.

محدوده محدود و شرایط کنترل‌شده محک‌های سنتی، که برای مقایسه‌های مستقیم بهینه شده‌اند، بدان معناست که آن‌ها نمی‌توانند تصویری واقعاً جامع از چگونگی عملکرد یک مدل در شرایط گسترده، باز، نامرتب و اغلب به‌سرعت در حال تغییر دنیای واقعی ارائه دهند. در مقابل، چت‌بات آرنا بهبود را از طریق یک معیار جامع واحد هدایت می‌کند: "رضایت عمومی مشتری". از این نظر، جدول رده‌بندی به بسیاری از مؤثرترین مکانیسم‌های حکمرانی اینترنت شباهت دارد، مانند رتبه‌بندی‌های ستاره‌ای eBay یا سیستم رأی‌دهی Reddit، که تعاملات پیچیده را به سیگنال‌های ساده و جهانی قابل فهم تبدیل می‌کنند.

در مقاله‌ی "رگولاتوری، به سبک اینترنت" در سال ۲۰۱۵ که توسط مرکز اش دانشکده کندی هاروارد برای حکمرانی دموکراتیک و نوآوری منتشر شد، "نیک گراسمن"، شریک عمومی در Union Square Ventures، مشاهده کرد که "وارونگی مرکزی در رگولاتوری به سبک اینترنت، چرخش از 'اجازه' به 'مسئولیت‌پذیری' است". منظور او این بود که پلتفرم‌ها و بازارهای اینترنتی،

مانند eBay و Airbnb و Uber، از شرکت‌کنندگان احتمالی نمی‌خواستند مجوز کسب‌وکار، یا مجوز فروشنده، یا هر نوع گواهی بوروکراتیک دیگری را به عنوان هزینه‌ی پیش‌پرداخت برای مشارکت دریافت کنند. در عوض، شما می‌توانستید بلافاصله وارد شوید. اما پس از ورود، در قلمروی عاری از رگولاتوری فرود نمی‌آمدید. در واقع، اینترنت یک فضای بسیار تنظیم‌شده بود و هست، که در آن میلیاردها تراکنش و تعامل معمول امتیازدهی، تجمیع، تجزیه و تحلیل و به "امتیازات اعتبار" و سایر معیارهای شفافیت و مسئولیت‌پذیری تبدیل می‌شوند که وظایف حکمرانی را انجام می‌دهند که به اندازه کافی انعطاف‌پذیر هستند تا با سرعت و مقیاس حرکت اینترنت همگام شوند.

و این همان کاری است که جدول رده‌بندی چت‌بات آرنا برای مدل‌های هوش مصنوعی انجام می‌دهد. برخلاف بسیاری از این نوع مکانیسم‌های حکمرانی اینترنتی، که رفتار شرکت‌کنندگان را ارزیابی می‌کنند، چت‌بات آرنا خروجی‌های سیستم را ارزیابی می‌کند. با تجمیع در طول زمان، آرای فردی "احکام" جمعی را هم بر عملکرد مدل فردی و هم، به‌طور کلی‌تر، بر چگونگی عملکرد مدل‌ها از نظر عموم مردم، ایجاد می‌کنند.

یک جنبه کلیدی از رویکرد استقرار تکراری OpenAI این است که چگونه امکان آزمون غیرمتمرکز و عملی در مقیاسی را فراهم می‌کند که هرگز نمی‌توانید در یک آزمایشگاه به آن دست یابید. صدها میلیون نفر در سراسر جهان اکنون از صدها مدل استفاده می‌کنند. با وقوع این فرآیند، هر یک از این کاربران به‌صورت فردی درباره نحوه عملکرد مدل‌ها، آنچه شخصاً ارزش می‌دهند و اعتماد می‌کنند، و آنچه ارزش نمی‌دهند و اعتماد نمی‌کنند، می‌آموزند. توسعه‌دهندگان نیز این چیزها را در مقیاس وسیع، برای مدل‌های خود می‌آموزند. آن‌ها می‌توانند ببینند مردم چه چیزی را دوست دارند و دوست ندارند، و مسائل از کجا پدید می‌آیند، و سپس یافته‌های خود را در به‌روزرسانی‌ها و نسخه‌های آینده گنجانده کنند.

اما جدول‌های رده‌بندی مانند چت‌بات آرنا برای تجمیع داده‌ها در مقیاس بزرگ‌تر طراحی شده‌اند، که داده‌های عملکرد بیش از ۱۰۰ مدل را به‌طور هم‌زمان ردیابی می‌کنند، نه فقط یک

مدل. با این کار، کاربران می‌توانند از تجربه سایر کاربران بیاموزند، نه فقط از تجربه خودشان. و این "هوش جمعی" را می‌توان به‌راحتی به تمام شرکت‌کنندگان در یک بخش اعمال کرد، که چیز جدیدی است. به‌عنوان مثال، راه آسانی برای تعیین میانگین رضایت کاربر از عملکرد کلی یوتیوب یا فیس‌بوک وجود ندارد، چه رسد به مقایسه‌ی آن‌ها به‌گونه‌ای که نشانه‌ی روشنی از کیفیت آن‌ها ارائه دهد. همین امر در مورد انتشارات رسانه‌ای نیز صادق است: هیچ سایتی وجود ندارد که به شما اجازه دهد به‌طور ناشناس سه رسانه مختلف را در مورد نحوه پوشش یک داستان رتبه‌بندی کنید. و حتی اگر چنین سایتی وجود داشت، دستکاری آن آسان بود. می‌توانستید جستجو کنید تا ببینید کدام نشریه کدام داستان را قبل از رأی دادن منتشر کرده است. اما در چت‌بات آرنا، مدل‌ها پاسخ‌های خود را در زمان واقعی تولید می‌کنند؛ خروجی‌های آن‌ها فقط به این دلیل وجود دارند که شما همین حالا آن‌ها را به وجود آورده‌اید. بنابراین، به‌طور کلی در برابر تلاش برای دستکاری سیستم مقاوم است.

همان‌طور که مدیران چت‌بات آرنا به توسعه سایت ادامه می‌دهند، اشکال دقیق‌تری از ارزیابی را نیز ارائه می‌دهند. به‌عنوان پیش‌فرض، یک رتبه‌بندی کلی را نشان می‌دهد که نشان می‌دهد کدام مدل‌ها بیشترین آرای برنده را از کاربران دریافت کرده‌اند. اما سایت می‌تواند داده‌ها را دقیق‌تر از آن ارزیابی کند. به‌عنوان مثال، می‌تواند انواع دستوراتی که کاربران می‌دهند را در دسته‌های مختلفی مانند "پیروی از دستورالعمل‌ها"، "کدنویسی"، "ریاضی" یا "پرسش طولانی‌تر" گروه‌بندی کند، سپس نشان دهد که عملکرد یک مدل برای انواع مختلف موارد استفاده ممکن است چگونه متفاوت باشد. این نمونه‌ای دیگر است که استفاده بیشتر منجر به قابلیت‌های بیشتر خواهد شد. تصور کنید، به‌عنوان مثال، اگر چند صد میلیون نفر شروع به استفاده از چت‌بات آرنا برای ارزیابی عملکرد مدل می‌کردند. با این تعداد افراد که هر روز دستورات را ارسال می‌کنند، حتی موارد استفاده نسبتاً مبهم نیز ممکن است به‌سرعت به اهمیت آماری دست یابند. از نظر تئوری، کاربران می‌توانستند در نهایت به بینش‌های یک‌جا در مورد تقریباً هر جنبه‌ای از عملکرد مدل دست یابند. کدام مدل‌ها در وظایف نوشتن خلاقانه بهترین هستند؟ کدام یک در

ترجمه از انگلیسی به ماندارین بهترین هستند؟ کدام یک برای "نرخ‌های توهم‌زایی پایین" و "خروجی‌های نیازمند هوش هیجانی و همدلی" امتیاز بالایی کسب می‌کنند؟ علاوه بر این، مدیران چت‌بات آرنا می‌توانستند آزمایش‌های خود را بر روی خروجی‌ها اجرا کنند و نرخ‌های واقعی و صنعتی را برای پدیده‌هایی مانند نادرستی‌های واقعی یا خروجی‌های سمی به دست آورند. آن‌ها می‌توانستند از کاربران بپرسند چرا یک خروجی را به دیگری ترجیح داده‌اند، تا اطلاعات بیشتری درباره ترجیحات کاربران و فرآیندهای تصمیم‌گیری در تعاملات هوش مصنوعی به دست آورند.

با چنین پتانسیلی، چت‌بات آرنا راه را به سوی آینده‌ای نزدیک به حکمرانی دموکراتیک و مردمی "رگولاتوری ۲.۰" نشان می‌دهد، که در آن کاربران از طریق بیان جمعی ترجیحات و قضاوت‌های خود، توسعه‌ی مستمر هوش مصنوعی را شکل می‌دهند و اعتماد از طریق شفافیت حاصل می‌شود. اگر کاربران بتوانند به‌راحتی تعیین کنند که، به‌عنوان مثال، مدل X سه درصد مواقع هنگام ارائه مشاوره پزشکی اطلاعات نادرست تولید می‌کند، می‌توانند تصمیمات بسیار آگاهانه‌تری درباره اینکه کدام مدل‌های هوش مصنوعی را برای وظایف مختلف اعتماد کنند، بگیرند. توسعه‌دهندگان نیز به‌نوبه خود، انگیزه‌های عمده‌ای برای بهبود مدل‌های خود به‌روش‌هایی که واقعاً برای کاربران اهمیت دارند، خواهند داشت.

نتیجه‌گیری

روایت غالب رسانه‌ای از "مسابقه‌ی تسلیحاتی هوش مصنوعی" در حالی که توجه عمومی را به‌سرعت و رقابت در این حوزه جلب کرده، از تبیین ماهیت واقعی توسعه هوش مصنوعی که بر پایه‌ی آزمون‌های جامع، شفافیت و داده‌محوری استوار است، غافل مانده است. این فصل نشان داد که تکامل هوش مصنوعی بیشتر شبیه به "مسابقه‌ی فضایی" است تا یک رقابت ستیزه‌جویانه، و توسعه‌دهندگان هوش مصنوعی افرادی هستند که وسواس زیادی به سنجش و بهبود مستمر دارند. از آزمون تورینگ در میانه قرن بیستم تا ظهور محک‌های مدرن و پلتفرم‌های ارزیابی جمعی مانند چت‌بات آرنا، آزمون همواره قلب پیشرفت هوش مصنوعی بوده است.

محک‌ها نه تنها ابزارهایی برای ارزیابی عملکرد فنی هستند، بلکه به عنوان مکانیسم‌های پویا عمل می‌کنند که انگیزه‌ی بهبود را در صنعت ایجاد کرده و استانداردهایی را برای شفافیت و مسئولیت‌پذیری تعیین می‌کنند که فراتر از مقررات ایستا و سنتی هستند. درحالی‌که چالش‌هایی مانند "آموزش برای آزمون" و آسیب‌پذیری در برابر مثال‌های خصمانه همچنان وجود دارند، که سؤالاتی را درباره درک واقعی و تعمیم‌پذیر مدل‌ها مطرح می‌کنند، پیشرفت مستمر در کاهش نرخ خطا و پیچیدگی محک‌ها، نشان‌دهنده گام‌های مهمی به سوی قابلیت‌های شناختی پیشرفته‌تر است.

در نهایت، اعتماد به هوش مصنوعی نه با دستیابی به کمال غیرممکن، بلکه با پذیرش نرخ خطای قابل قبول، شفافیت در عملکرد و ارزیابی مداوم توسط عموم مردم و متخصصان، شکل می‌گیرد. پلتفرم‌های مردمی مانند چت‌بات آرنا نشان‌دهنده‌ی مسیر آینده حکمرانی هوش مصنوعی هستند، جایی که هوش جمعی و بازخورد کاربران به شکل‌دهی و بهبود مدل‌ها به‌روش‌هایی که واقعاً برای جامعه مفید و قابل اعتماد هستند، کمک می‌کند.

نکات کلیدی

- توسعه هوش مصنوعی با آزمون‌های جامع و داده‌محور مشخص می‌شود، نه یک "مسابقه‌ی تسلیحاتی" بی‌پروا.
- آزمون تورینگ پایه‌های ارزیابی هوش ماشینی را بنا نهاد و فرهنگ آزمایش را در هوش مصنوعی ریشه‌دار کرد.
- محک‌ها (Benchmarks) آزمون‌های استاندارد شده‌ای هستند که توسط اشخاص ثالث (مانند دانشگاه‌ها) برای اندازه‌گیری عملکرد مدل‌ها و مقایسه آن‌ها ایجاد می‌شوند.
- محک‌ها پویا هستند و به طور مستمر مدل‌ها را به بهبود تشویق می‌کنند، که از این نظر با رگولاتوری‌های سنتی و ثابت متفاوت هستند.
- طیف وسیعی از محک‌ها، نه تنها دقت، بلکه عدالت، ایمنی، قابلیت اطمینان و

توضیح‌پذیری مدل‌ها را نیز ارزیابی می‌کنند.

- مثال‌هایی مانند "مثال‌های خصمانه" نشان‌دهنده محدودیت‌های مدل‌ها در درک واقعی جهان و استدلال عقل سلیم هستند، حتی با وجود عملکرد بالا در وظایف خاص.
- "آلودگی داده (data contamination) "یک چالش مهم است که می‌تواند منجر به عملکرد مصنوعی مدل شود، اما توسعه‌دهندگان برای جلوگیری از آن تلاش می‌کنند.
- مفهوم "شکنندگی مدل‌ها" و خطاهای مداوم، توسط منتقدان به عنوان ضعف اساسی مطرح می‌شود، در حالی که دیدگاه دیگر بر پیشرفت مداوم و اعتماد بر اساس نرخ خطای قابل قبول تأکید دارد.
- "تفسیرپذیری" و "توضیح‌پذیری" برای ایجاد اعتماد در هوش مصنوعی حیاتی هستند، اما دستیابی به استانداردهای مطلق آن‌ها غیرواقعی است.
- دیدگاه عمومی و سودمندی اجتماعی هوش مصنوعی، که توسط ویزنباوم مطرح شد، سؤالات اساسی درباره اخلاق، ارزش‌ها و نیازهای واقعی انسان را مطرح می‌کند.
- "چت‌بات آرنا (Chatbot Arena) "نمونه‌ای از حکمرانی مردمی در ارزیابی هوش مصنوعی است که از طریق بازخورد جمعی کاربران، شفافیت و اعتماد را افزایش می‌دهد.

سؤالات تفکربرانگیز

۱. چگونه می‌توانیم رویکردهای رسانه‌ای به هوش مصنوعی را به سمتی سوق دهیم که به جای ایجاد ترس و هیجان‌زدگی، درک عمومی واقع‌بینانه‌تر و عمیق‌تری از چالش‌ها و فرصت‌های این فناوری ایجاد کند؟

۲. با توجه به تفاوت‌های میان محک‌های فنی و نیازهای دنیای واقعی، چگونه می‌توانیم اطمینان حاصل کنیم که مدل‌های هوش مصنوعی نه تنها در آزمون‌ها موفق می‌شوند، بلکه در کاربردهای عملی نیز به طور قابل اعتماد و مفید عمل

می‌کنند؟

۳. آیا مفهوم "هوش تعمیم‌پذیر" یک هدف واقع‌بینانه برای هوش مصنوعی است، یا باید بر توسعه مدل‌های متخصص و بسیار توانا در حوزه‌های خاص تمرکز کنیم؟

۴. با توجه به اینکه "تفسیرپذیری" و "توضیح‌پذیری" مطلق در هوش مصنوعی ممکن است غیرواقعی باشد، جامعه باید چه سطحی از شفافیت را از سیستم‌های هوش مصنوعی در حوزه‌های پرخطر مانند پزشکی یا حقوقی انتظار داشته باشد؟

۵. پلتفرم‌هایی مانند چت‌بات آرنا چگونه می‌توانند به طور مؤثر در شکل‌دهی به رگولاتوری هوش مصنوعی مشارکت کنند و آیا این مدل "حکمرانی مردمی" می‌تواند جایگزین یا مکمل رگولاتوری‌های رسمی دولتی باشد؟

۳۰ جمله کلیدی

۱. توسعه هوش مصنوعی بیش از یک "مسابقه‌ی تسلیحاتی"، یک "مسابقه‌ی فضایی" شفاف و مشارکتی است.

۲. آزمون‌های جامع و داده‌محور، ویژگی اصلی و پایه‌ای توسعه هوش مصنوعی هستند.

۳. توسعه‌دهندگان هوش مصنوعی، "عاشق سنجش" و بهبود مستمر بر پایه داده‌ها هستند.

۴. "آلن تورینگ" با "آزمون تورینگ" خود، ریشه‌های ارزیابی هوش ماشینی را بنا نهاد.

۵. پیشرفت‌های کلیدی هوش مصنوعی اغلب در آزمایشگاه‌های دانشگاهی و با فرهنگ غنی آزمون رخ داده است.

۶. رویدادهایی مانند شکست کاسپاروف توسط دیپ‌بلو، در واقع آزمون‌های عمومی برای سنجش پیشرفت هوش مصنوعی بودند.

۷. محک‌ها آزمون‌های استاندارد شده‌ای هستند که برای اندازه‌گیری عملکرد سیستم‌ها به کار می‌روند.

۸. محک‌ها به مقایسه نتایج با خطوط پایه و عملکرد رقبا در صنعت کمک می‌کنند.

۹. "سوپرگلو" نمونه‌ای از محک‌های چندوجهی برای ارزیابی فهم زبان در مدل‌های هوش مصنوعی است.

۱۰. جدول‌های رده‌بندی عمومی محک‌ها، شفافیت و مسئولیت‌پذیری را در توسعه هوش مصنوعی ترویج می‌کنند.

۱۱. محک‌ها، توسعه هوش مصنوعی را به یک "المپیک عمومی" تبدیل می‌کنند.

۱۲. محک‌ها برخلاف رگولاتوری‌های ایستا، مکانیسم‌های پویایی برای تحریک پیشرفت هستند.

۱۳. مقررات اغلب مانع بهبود می‌شوند، در حالی که آزمون‌ها الهام‌بخش آن هستند.

۱۴. محک‌ها به عنوان "رگولاتوری گیمیفاید" عمل می‌کنند.

۱۵. محک‌ها طیف گسترده‌ای از ابعاد مانند عدالت، ایمنی، قابلیت اطمینان و توضیح‌پذیری را ارزیابی می‌کنند.

۱۶. "RealToxicityPrompts" و "StereoSet" نمونه‌هایی از محک‌ها برای سنجش محتوای سمی و تعصبات اجتماعی هستند.

۱۷. مدل‌های جدیدتر هوش مصنوعی، مانند GPT-4 ، به طور چشمگیری میزان تولید محتوای سمی را کاهش داده‌اند.

۱۸. محک‌های مؤثر ممکن است با تحریک پیشرفت‌های بزرگ، خودشان منسوخ شوند.

۱۹. "آموزش برای آزمون" می‌تواند منجر به عملکرد بالا در محک‌های خاص شود، بدون درک واقعی.

۲۰. "مثال‌های خصمانه" نشان می‌دهند که مدل‌ها ممکن است درک "عقل سلیم" از جهان فیزیکی را نداشته باشند.

۲۱. چالش‌هایی مانند ترسیم اسب شاخدار باTikZ ، برای سنجش قابلیت‌های شناختی فراتر از تطابق الگو طراحی شده‌اند.

۲۲. "آلودگی داده" زمانی رخ می‌دهد که مدل‌ها به طور ناخواسته در طول آموزش در معرض داده‌های آزمون قرار گیرند.

۲۳. توسعه‌دهندگان اقدامات شدیدی برای جلوگیری از آلودگی داده و تضمین پیشرفت واقعی انجام می‌دهند.

۲۴. منتقدان "شکنندگی" مدل‌ها را با اشاره به خطاهای مداوم و "توهمات" برجسته می‌کنند.

۲۵. اعتماد به هوش مصنوعی باید بر اساس نرخ خطای قابل قبول و قابلیت اطمینان مستمر، نه کمال بی‌عیب و نقص، بنا شود.

۲۶. "تفسیرپذیری" و "توضیح‌پذیری" به درک و پیش‌بینی رفتار مدل‌ها کمک می‌کنند، اما استاندارد مطلق آن‌ها غیرواقعی است.

۲۷. "آنچه مدل انجام می‌دهد" برای اعتماد و پذیرش عملی، مهم‌تر از "چگونه انجام می‌دهد" است.

۲۸. ارزیابی هوش مصنوعی باید از دیدگاه عمومی و با تأکید بر سودمندی اجتماعی آن صورت گیرد (دیدگاه ویزنباوم).

۲۹. "چت‌بات آرنا" یک پلتفرم منبع باز برای ارزیابی مدل‌های زبان بزرگ بر اساس ترجیحات انسانی است.

۳۰. این پلتفرم نمونه‌ای از "حکمرانی به سبک اینترنت" است که با تجمیع بازخورد کاربران، توسعه هوش مصنوعی را دموکراتیک و شفاف می‌سازد.

فصل ۶

نوآوری به مثابه ایمنی: تعادل میان توسعه سریع و مدیریت مخاطرات در عصر هوش مصنوعی

"نوآوری یک نیروی قدرتمند برای پیشرفت است و در بسیاری از موارد، خود بهترین راهبرد ایمنی است. اما نادیده گرفتن نیاز به نظارت و تعادل با احتیاط، می‌تواند منجر به ایجاد مخاطرات جدیدی شود که حتی بزرگترین پیشرفت‌ها را نیز تحت‌الشعاع قرار دهد — ".یک پژوهشگر یا دانشمند برجسته در حوزه فناوری (این نقل قول برای غنی‌سازی متن و از منابع ارائه‌شده گرفته نشده است.)

اهداف اصلی فصل:

۱. بررسی چگونگی تعامل مفاهیم نوآوری و ایمنی در توسعه فناوری‌های پیشرفته، به ویژه هوش مصنوعی.

۲. مقایسه و تحلیل دو رویکرد متضاد "اصل احتیاطی" و "نوآوری بدون اجازه" در زمینه تنظیم مقررات فناوری.

۳. واکاوی نقش توسعه سریع و تکرارشونده در افزایش ایمنی محصولات و خدمات دیجیتال و هوش مصنوعی.

۴. مطالعه تطبیقی تاریخچه صنعت خودروسازی و درس‌های آن برای درک بهتر چالش‌ها و فرصت‌های پیش روی هوش مصنوعی.

۵. ارائه یک چارچوب فکری برای مدیریت مخاطرات نوظهور هوش مصنوعی با تاکید بر مزایای نوآوری مستمر و بازخورد جمعی.

چکیده

با قدرت گرفتن روزافزون مدل‌های هوش مصنوعی و گسترش قابلیت‌های آن‌ها، ضرورت وجود تدابیر ایمنی مؤثر و نظارت جامع انسانی بیش از پیش آشکار می‌شود. این فصل به بررسی چگونگی برقراری تعادل میان نوآوری و احتیاط می‌پردازد و استدلال می‌کند که نوآوری نه تنها با ایمنی در تضاد نیست، بلکه خود می‌تواند به عنوان یک محرک کلیدی برای افزایش ایمنی عمل کند. ما به تحلیل دو دیدگاه غالب در این زمینه، یعنی "اصل احتیاطی" که بر کنترل پیشگیرانه تأکید دارد، و "نوآوری بدون اجازه" که فضایی برای آزمایش و تطبیق فراهم می‌آورد، خواهیم پرداخت. با بررسی سوابق تاریخی از جمله توسعه اینترنت و صنعت خودروسازی، نشان می‌دهیم که چگونه رویکردهای توسعه تکرارشونده و بازخورد محور، به طور مؤثرتری به ایجاد محصولات ایمن‌تر و قابل اعتمادتر منجر شده‌اند. همچنین، این فصل بر اهمیت حفظ قدرت نوآوری در بستر جهانی و تزریق ارزش‌های دموکراتیک به فناوری‌های نوظهور تأکید می‌کند، و در نهایت چارچوبی برای درک و مدیریت مخاطرات هوش مصنوعی در سایه مزایای بی‌شمار آن ارائه می‌دهد.

مقدمه

در دنیای امروز که فناوری هوش مصنوعی با سرعتی بی‌سابقه در حال پیشرفت است، مدل‌ها و سیستم‌های هوش مصنوعی قابلیت‌های گسترده‌تر و پیچیده‌تری پیدا می‌کنند. این پیشرفت‌های چشمگیر، در کنار فرصت‌های بی‌بدیل برای افزایش رفاه جهانی و حل چالش‌های بزرگ، نگرانی‌هایی را نیز در مورد ایمنی، اخلاق و حکمرانی این فناوری‌ها برانگیخته است. پرسش اساسی این است که چگونه می‌توانیم اطمینان حاصل کنیم که نوآوری با احتیاط و مسئولیت‌پذیری همراه باشد؟ برخی تمایل دارند ایمنی را با مفاهیمی چون احتیاط، تعمق و توجه دقیق مرتبط بدانند، در حالی که برخی دیگر بر اهمیت سرعت توسعه و پویایی فناوری تأکید می‌کنند. این دیدگاه‌های متضاد، بحث‌های گسترده‌ای را در مورد ماهیت صحیح تنظیم مقررات و مداخله‌های سیاستی در حوزه فناوری‌های نوظهور دامن زده است.

در این فصل، ما به بررسی دقیق این دوگانگی می‌پردازیم و استدلال می‌کنیم که تأخیر در توسعه از طریق مقررات دست‌وپاگیر، نه تنها ظرفیت ما را برای ایجاد محصولات و خدماتی که رفاه جهانی را افزایش می‌دهند، محدود می‌سازد، بلکه می‌تواند اثرات مثبت نوآوری بر ایمنی را نیز به تعویق اندازد. این مسئله به ویژه در بستر جهانی از اهمیت بالایی برخوردار است، جایی که قدرت‌های مختلف برای پیشبرد اولویت‌های استراتژیک خود از هوش مصنوعی بهره می‌برند. حفظ پیشروی در نوآوری به معنای تزریق ارزش‌های دموکراتیک به این فناوری‌ها و ادغام آن‌ها در جامعه به نحوی است که قدرت اقتصادی، امنیت ملی و نفوذ جهانی را تقویت کند. این فصل با نگاهی عمیق‌تر به "اصل احتیاطی" و "نوآوری بدون اجازه" و با ارائه مثال‌های تاریخی از جمله تجربه صنعت خودروسازی و توسعه اینترنت، می‌کوشد تا درک جامع‌تری از این پویایی‌های حیاتی ارائه دهد و مسیرهای ممکن برای دستیابی به نوآوری ایمن و مسئولانه را روشن سازد.

نوآوری به مثابه ایمنی: پارادایم نوین توسعه

با گسترش توانایی‌ها و قدرت مدل‌های هوش مصنوعی، نیاز به تدابیر ایمنی کارآمد و نظارت

گسترده انسانی بیش از پیش حیاتی می‌شود. در همین راستا، ضرورت ایجاد تعادل میان نوآوری و احتیاط نیز افزایش می‌یابد، اما پرسش اینجاست که چگونه می‌توان به این تعادل دست یافت؟ برخی از منتقدان، ایمنی را با ویژگی‌هایی نظیر احتیاط، تعمق و دقت مرتبط می‌دانند. با این حال، سرعت توسعه نیز عامل بسیار مهمی محسوب می‌شود. مقررات و سایر مداخلات سیاستی که با هدف به تأخیر انداختن توسعه تدوین شده‌اند، نه تنها ظرفیت ما را برای خلق محصولات و خدمات پیشگامانه که رفاه جهانی را افزایش می‌دهند، مختل می‌کنند، بلکه می‌توانند اثرات مثبت نوآوری بر ایمنی را نیز به تأخیر بیندازند.

توجه به بستر جهانی فناوری نیز از اهمیت بالایی برخوردار است. در حالی که شرکت‌های آمریکایی و مؤسسات آکادمیک در خط مقدم رنسانس یادگیری ماشین قرار دارند که از اوایل دهه ۲۰۱۰ آغاز شده است، کشورهایی مانند چین، فرانسه، اسرائیل، هند، کره جنوبی و بسیاری دیگر، همگی مصمم هستند تا از هوش مصنوعی برای پیشبرد اولویت‌های استراتژیک خود استفاده کنند. در دنیایی که به شدت شبکه ای شده و با افزایش سطوح وابستگی متقابل مشخص می‌شود، حفظ "قدرت نوآوری" آمریکا – همانطور که اریک اشمیت، مدیرعامل سابق گوگل و رئیس اجرایی آلفابت توصیف می‌کند – به عنوان یک اولویت اصلی ایمنی مطرح می‌شود. با حفظ رهبری در توسعه، ما فناوری‌های هوش مصنوعی را با ارزش‌های دموکراتیک تلفیق کرده و آن‌ها را به گونه‌ای در جامعه ادغام می‌کنیم که قدرت اقتصادی، امنیت ملی و توانایی ما برای فرافکنی گسترده نفوذ جهانی را تقویت کند.

توسعه سریع همچنین به معنای توسعه انطباقی است. و توسعه انطباقی به نوبه خود به چرخه عمر کوتاه‌تر محصول، به‌روزرسانی‌های مکرر و محصولات ایمن‌تر منجر می‌شود. به ویژه در حوزه نرم‌افزارهای اینترنتی، جایی که توزیع آنی است و حلقه‌های بازخورد نزدیک هستند، معمولاً می‌توان از طریق نوآوری به ایمنی دست یافت، بسیار مؤثرتر از آنکه بتوان از طریق مقررات به ایمنی رسید.

اصل احتیاطی: دیدگاه 'مقصر تا اثبات بی‌گناهی'

این دیدگاه با طرز تفکر منتقدان و بدبینان نسبت به فناوری تفاوت اساسی دارد. در ماه‌های اولیه پس از انتشارChatGPT ، حس فوریت جدیدی باعث شد که خواستار تنظیم مقررات توسعه و محدود کردن دسترسی عمومی به این منابع جدید شوند. برای مثال، تد لیو، عضو کنگره آمریکا از کالیفرنیای جنوبی، در یک مقاله در نیویورک تایمز نوشت که از هوش مصنوعی "کنترل‌نشده و بدون مقررات" "ترسیده" است. به عقیده او، ما نیاز به ایجاد یک آژانس فدرال جدید منحصراً اختصاص یافته به این فناوری داریم، یک سازمان غذا و داروی (FDA) برای هوش مصنوعی، تا مقررات را سریع‌تر از آنچه کنگره قادر به مدیریت آن است، تدوین و اجرا کند.

لوسی پاول، سخنگوی دیجیتال حزب کارگر بریتانیا، پیشنهاد کرد که توسعه هوش مصنوعی باید مانند داروها و انرژی هسته‌ای تنظیم شود، با نیاز به مجوز برای ساخت مدل‌هایی مانند ChatGPT. الکساندرا ون هافلن، وزیر دیجیتالی‌سازی هلند، ارتباطی بین هوش مصنوعی و خودروها برقرار کرد و اظهار داشت: "هرگاه یک خودروی جدید تولید می‌کنید، ابتدا می‌خواهید مطمئن شوید که رانندگی با آن ایمن است، پیش از آنکه اجازه دهید در خیابان‌ها تردد کند".

همه این قانونگذاران و مقامات دولتی از رویکردی نظارتی دفاع می‌کردند که اغلب به عنوان "اصل احتیاطی" توصیف می‌شود. در اصل، اصل احتیاطی معتقد است که فناوری‌های جدید "مقصرند تا زمانی که بی‌گناهی‌شان اثبات شود".

در مقابل، بسیاری از فناوران، کارآفرینان و سرمایه‌گذاران از رویکردی به نام "نوآوری بدون اجازه" حمایت می‌کنند. برخلاف تلاش‌های از بالا به پایین برای کنترل پیشگیرانه یا حتی ممنوعیت فناوری‌های جدید، رویکرد نوآوری بدون اجازه صراحتاً فضای کافی برای نوآوری، آزمایش و تطبیق را فراهم می‌کند، به ویژه زمانی که آسیب‌های ملموسی هنوز وجود ندارند، یا زمانی که مقررات موجود، آسیب‌های موجود را پوشش می‌دهند.

هرچند مفهوم نوآوری بدون اجازه برای ایجاد شرایطی که به تسهیل آزمایش و خطا کمک

می‌کند بسیار مفید است، اما یک پیامد ناخواسته دارد. با مفهوم‌سازی نوآوری به عنوان چیزی که زمانی مولدترین حالت خود را دارد که به طور نامتناسبی توسط تلاش‌های نظارتی رسمی محدود نشده باشد – که قطعاً درست است – ما ممکن است قدرت ذاتی نظارتی خود نوآوری را نادیده بگیریم. ما تمایل داریم قانونگذاران و کارآفرینان را در مقابل یکدیگر قرار دهیم. در واقعیت، هر دو آرزوی مشترکی برای بهبود جنبه‌ای از جهان دارند. قانونگذاران با ایجاد قوانین جدید این کار را می‌کنند. کارآفرینان با ایجاد محصولات و خدمات جدیدی که قابلیت‌های بدیعی را ارائه می‌دهند یا به نوعی طرح‌های موجود را بهبود می‌بخشند.

برای مثال، در روزهای اولیه صنعت خودروسازی، تولیدکنندگان انگیزه‌های طبیعی برای ساخت خودروهایی قابل اعتمادتر و آسان‌تر برای استفاده داشتند که به نوبه خود آنها را هم ایمن‌تر و هم قابل فروش‌تر می‌کرد. به این ترتیب، هر تکرار جدید یک جزء یا ویژگی می‌تواند به عنوان یک "قانون "یا "مقررات" دیده شود که عموم مردم می‌توانند با پذیرش آن، آن را تأیید کنند. استارت‌های برقی، که چارلز کترینگ در سال ۱۹۱۱ اختراع کرد و کادیلاک در سال ۱۹۱۲ به مدل‌های خود افزود، نمونه بارز این امر است. قبل از در دسترس بودن آن‌ها، رانندگان مجبور بودند از هندل دستی برای چرخاندن میل لنگ موتور به عنوان بخشی از مراحل لازم برای روشن کردن خودرو استفاده کنند. این فرآیند نه تنها نیازمند درجه‌ای از قدرت بود که فقط برخی از رانندگان از آن برخوردار بودند، بلکه منجر به شکستگی تعداد زیادی از مچ‌ها، بازوها و حتی فک‌ها می‌شد، زمانی که موتورها با خطا مواجه می‌شدند و هندل‌ها را به عقب می‌چرخاندند.

معرفی استارت‌های برقی، که صرفاً از راننده می‌خواست تا پدال پایی را فشار دهد تا موتور روشن شود، به دموکراتیک شدن دسترسی به خودروها کمک کرد و آن‌ها را ایمن‌تر و قابل اعتمادتر ساخت. رانندگان آنها را دوست داشتند، و در نتیجه، استارت‌های برقی به سرعت در سراسر صنعت استاندارد شدند. تا سال ۱۹۲۰، موتورهای هندل‌دار تقریباً ناپدید شده بودند.

اکنون، به دلیل چگونگی گسترش پایگاه‌های کاربری اینترنت و دعوت به مشارکت، عملکرد نظارتی نوآوری بدون اجازه بسیار قدرتمندتر از همیشه شده است. در قرن بیست و یکم، هر کسی

که از آزادی برای آزمایش باز و تکرارشونده، بدون تأیید قبلی از سوی تنظیم‌کنندگان رسمی، دفاع می‌کند، نه تنها از نظارت و کنترل دموکراتیک فرار نمی‌کند، بلکه عملاً به استقبال آن می‌رود. این امر در نهایت به بازخورد جامع‌تر و فراگیرتر، چرخه‌های محصول کوتاه‌تر و بهبودهای سریع‌تر منجر می‌شود که به طور کلی محصولات ایمن‌تر و قابل اعتمادتر را به ارمغان می‌آورد.

پیدایش اینترنت و سیاست‌های حمایت از نوآوری بدون اجازه

هنوز هم این تصور پابرجا است که اینترنت یک "غرب وحشی" از توسعه بی‌قید و بند و خطرناک است که هیچ کلانتری در آن دیده نمی‌شود. در حالی که این دیدگاه، دامنه گسترده‌ای از مداخلات احتیاطی را که پیشرفت فناوری را به طرق مختلف کُند کرده‌اند، نادیده می‌گیرد، همچنین تنظیم‌کنندگان دولتی را از داستان بزرگ‌تر تأثیرات اجتماعی مثبت فناوری حذف می‌کند. در این دیدگاه، ظهور صنعت فناوری به عنوان موتور جهانی تحول فرهنگی و رشد اقتصادی، به نوعی صرفاً از طریق ترکیبی از افزایش‌های ثابت و چشمگیر در قدرت پردازش کامپیوتر، جادوی بازار آزاد، تعداد انگشت‌شماری از ستارگان برنامه‌نویسی و بی‌تفاوتی نظارتی اتفاق افتاده است.

اما این تصویر واقعیت ندارد. همانطور که آدام تایر، تحلیلگر سیاست عمومی در کتاب خود "نوآوری بدون اجازه" (۲۰۱۶) توضیح می‌دهد، شرایطی که به تبدیل فناوری پیشرفته به پویاترین صنعت آمریکا در سی سال گذشته کمک کرد، در واقع کاملاً عمدی بود: "به طور خاص، از اوایل دهه ۱۹۹۰، گروهی دوحزبی از سیاست‌گذاران به نوآوران چراغ سبز دادند تا ذهن خود را آزاد بگذارند و با مجموعه‌ای بی‌شمار از دستگاه‌ها و خدمات جدید هیجان‌انگیز آزمایش کنند. سیاست‌گذاران آمریکایی از طریق مجموعه‌ای از اظهارات سیاستی، اعلام کردند که نوآوری بدون اجازه برای اینترنت و فناوری دیجیتال در آمریکا یک هنجار خواهد بود".

در سال ۱۹۹۱، بنیاد ملی علوم، آژانس فدرالی که NSFNET ، شبکه‌ای که ستون فقرات اینترنت در آمریکا را تشکیل می‌داد، مدیریت می‌کرد، شروع به کاهش محدودیت‌های خود در برابر

استفاده تجاری از آن کرد. این تغییر سیاستی که طی چندین سال رخ داد، مشارکت و سرمایه‌گذاری رو به رشد بخش خصوصی را در زیرساخت‌ها و خدمات اینترنت ممکن ساخت که به نوبه خود رشد و دسترسی اینترنت برای استفاده شخصی را تسریع بخشید. سپس کنگره، قانون ارتباطات ۱۹۹۶ را تصویب کرد که شامل ماده ۲۳۰ بود. ماده ۲۳۰ که اغلب به عنوان "بیست و شش کلمه‌ای که اینترنت را ساختند" توصیف می‌شود، اپراتورهای وب‌سایت و سایر واسطه‌های آنلاین را از دعاوی مسئولیت ناشی از محتوایی که اشخاص ثالث در سایت‌ها، پلتفرم‌ها و شبکه‌های آنها منتشر می‌کنند، محافظت می‌کند.

بدون ماده ۲۳۰، وبلاگ‌نویسان محبوبی مانند مورخ هدر کاکس ریچاردسون یا اقتصاددان تایلر کوئن، احتمالاً بسیار کمتر مایل بودند که نظرات کاربران را در وب‌سایت‌های خود مجاز سازند. توسعه‌دهندگانی مانند ساب‌استک، مدیوم و وردپرس، احتمالاً بسیار کمتر مایل بودند پلتفرم‌های وبلاگ‌نویسی برای وبلاگ‌نویسان فردی بسازند. و در مراحل بالاتر زنجیره، ارائه‌دهندگان خدمات ابری مانند آمازون وب‌سرویسز یا مایکروسافت، احتمالاً بسیار کمتر مایل بودند میزبان پلتفرم‌های وبلاگ‌نویسی باشند. در نتیجه، اینترنت بسیار بیشتر پخش‌محور (broadcast-oriented) می‌شد تا تعاملی، با موانع قانونی و تکنولوژیکی برای ورود که احتمالاً ۹۹.۹ درصد جمعیت جهان را به مشارکت صرفاً به عنوان خریداران و مخاطبان غیرفعال اخبار و سرگرمی محدود می‌کرد.

در نهایت، در سال ۱۹۹۷، رئیس‌جمهور بیل کلینتون و معاون رئیس‌جمهور ال گور "چارچوب تجارت اقتصادی جهانی" را منتشر کردند، یک سند سیاستی که نیویورک تایمز آن را "رویکردی دست‌بردارانه و بدون مالیات جدید برای تنظیم معاملات تجاری در شبکه جهانی کامپیوتر" توصیف کرد. این چارچوب با به رسمیت شناختن ظهور اینترنت به عنوان "وسیله‌ای از زندگی روزمره"، از رویکردهای بازارمحور دفاع کرد و مکرراً از مقررات جدید و غیرضروری، رویه‌های بوروکراتیک جدید و سایر مداخلات از بالا به پایین دلسردی نشان داد. در آن آمده بود: "در جایی که دخالت دولتی لازم است، هدف آن باید حمایت و اجرای یک محیط حقوقی قابل پیش‌بینی،

حداقلی، سازگار و ساده برای تجارت باشد".

در مجموع، این انتخاب‌های سیاستی، محیطی برای نوآوری، سرمایه‌گذاری و زایش قابل توجهی ایجاد کردند که در طول دهه ۱۹۹۰ در اقتصاد آمریکا موج زد. از سال ۱۹۹۳ تا ۱۹۹۷، صنعت فناوری بیش از یک میلیون شغل جدید ایجاد کرد. تا اواخر دهه ۱۹۹۰، زمانی که رونق دات‌کام به خوبی در جریان بود، تولید ناخالص داخلی با نرخ متوسط سالانه بیش از ۴ درصد رشد می‌کرد. نرخ بیکاری از کمی کمتر از ۸ درصد در سال ۱۹۹۳ به کمتر از ۴ درصد در سال ۲۰۰۰ کاهش یافت.

در دهه‌های پس از آن، رویکردهای نوآوری بدون اجازه که در دهه ۱۹۹۰ اجرا شدند، به ارائه فناوری‌های پیشگامانه جدید، از جمله خدمات ابری، گوشی‌های هوشمند و رسانه‌های اجتماعی ادامه دادند. با این حال، با فراگیر شدن این فناوری‌ها، احساسات عمومی شروع به تغییر کرد. بخشی به دلیل عادت ما به مزایای یک فناوری و تمرکز بیشتر بر معایب آن در طول زمان، کاربران و تنظیم‌کنندگان به طور یکسان شروع به بررسی دقیق‌تر قدرت و نفوذ غول‌های فناوری کردند.

در مارس ۲۰۱۸، جنبش‌های مردمی مانند #DeleteFacebook در پی رسوایی کمبریج آنالیتیکا، زمانی که یک افشاگر فاش کرد که چگونه فیس‌بوک به یک توسعه‌دهنده برنامه اجازه داده بود تا داده‌های شخصی حداکثر ۸۷ میلیون نفر را جمع‌آوری کرده و بدون رضایت آن‌ها برای اهداف سیاسی استفاده کند، توجه زیادی را به خود جلب کرد. کمتر از یک ماه بعد، سناتورها مارک زاکربرگ را به کپیتول هیل فراخواندند، جایی که دقیقاً از همان جایی که جلسات استماع قانونگذاران دهه ۱۹۶۰ در مورد مرکز داده ملی متوقف شده بود، ادامه دادند و فیس‌بوک و سایر پلتفرم‌های اجتماعی را به عنوان تهدیدی وجودی برای حریم خصوصی افراد معرفی کردند که احتمالاً توانایی آن‌ها را برای شکل دادن به سرنوشت خود به خطر می‌اندازد یا حتی از بین می‌برد.

امروزه، نگرانی‌ها در مورد حریم خصوصی داده‌ها، اطلاعات غلط و سایر پیامدهای فناوری‌محور

چنان قوی شده است که درخواست‌ها برای رویکردهای احتیاطی رایج‌تر می‌شوند. اما نوآوری بدون اجازه، احتمالاً مؤثرتر از همیشه عمل می‌کند. نه تنها پیشرفت‌هایی را که تاکنون در هوش مصنوعی دیده‌ایم، ممکن ساخت، بلکه به تسریع پیشرفت در خودروهای الکتریکی، ویرایش ژن CRISPR، فناوری‌های ذخیره‌سازی انرژی در مقیاس شبکه و خورشیدی، ارزهای دیجیتال، پزشکی از راه دور، چاپ سه‌بعدی و واقعیت افزوده نیز کمک کرده است.

نوآوری‌هایی مانند این، نه تنها زندگی روزمره ما را بهبود می‌بخشند. آنها آمریکا را در خط مقدم حل چالش‌های جهانی، از تغییرات اقلیمی تا دسترسی به مراقبت‌های بهداشتی، نگه می‌دارند. در زمانی که حرکت سریع در چشم‌انداز جهانی به ندرت تا این حد مهم بوده است، و ما به لطف قدرت نوآوری خود یک مزیت رقابتی متمایز داریم، تنظیم‌کنندگان و منتقدان ضد فناوری، با شدیدترین لحن ممکن، اصرار دارند که زمان آن رسیده است که سریع حرکت کنیم و ترمزها را بکشیم.

نامه مؤسسه آینده زندگی که خواستار توقف کامل توسعه هوش مصنوعی شده بود، این دیدگاه را به اوج غیرمنطقی خود رساند. این نامه اعلام کرد: "این پروتکل‌ها (ایمنی مشترک) باید اطمینان حاصل کنند که سیستم‌های پایبند به آنها فراتر از یک شک معقول ایمن هستند." با این کار، این نامه به طور دقیق استاندارد حقوق کیفری را که گناه باید "فراتر از یک شک معقول" اثبات شود، واژگون کرد.

در حقوق کیفری، مجازات‌های اعمال شده اغلب شامل حبس فیزیکی می‌شود: یک هیئت منصفه می‌تواند به معنای واقعی کلمه آزادی و خودمختاری یک فرد را سلب کند. بنابراین، تعهد ما به متهم، به کسانی که قربانی جرم مورد نظر شده‌اند، و به یکپارچگی کل سیستم عدالت کیفری این است که تا آنجا که عقل اجازه می‌دهد، اطمینان حاصل کنیم که متهم واقعاً کاری را که دادستان‌ها او را به آن متهم می‌کنند، انجام داده است. در نسخه "فراتر از یک شک معقول" مؤسسه آینده زندگی، آن‌ها از حبس یک فناوری دفاع می‌کنند زیرا نمی‌توانیم مطمئن باشیم که ممکن است روزی، شاید، کار بدی انجام ندهد.

این فقط پلیس پیش‌بینی‌کننده بدون شواهد تجربی نیست. این صدور حکم پیش‌بینی‌کننده بدون شواهد تجربی است. در هر دوره‌ای، این یک تناسب عجیب برای یک جامعه دموکراتیک است. در عصر اینترنت، حتی بیشتر نامتعادل به نظر می‌رسد. همانطور که مایکل پولان هنگام توصیف ریشه‌های اصل احتیاطی در دهه ۱۹۷۰ اشاره کرد، این یک پاسخی بود که به دلیل فقدان داده تصور شد. اکنون، با این حال، ظرفیت ما برای تولید آنی داده‌ها، به اشتراک‌گذاری گسترده آن‌ها و تحلیل مداوم آن‌ها به طور چشمگیری قدرتمندتر است. و البته، این امر در مورد فناوری‌های دیجیتال حتی بیشتر صدق می‌کند.

این به معنای آن نیست که دیگر سناریوهایی وجود ندارد که اصل احتیاطی در آن‌ها منطقی باشد، حتی با فناوری‌های دیجیتال. اما دلیلی نیز وجود دارد که مفاهیمی مانند "حداقل محصول قابل قبول"، "تناسب محصول-بازار" و "چابکی" توسعه نرم‌افزار اینترنت را در سی سال گذشته تعریف کرده‌اند.

توسعه تکرارشونده و حلقه بازخورد مداوم

اکنون یادگیری بسیار سریع‌تر اتفاق می‌افتد. به‌روزرسانی‌ها به راحتی قابل انتشار هستند. رهایی از هزینه‌ها، محدودیت‌ها و سرعت آهسته تولید و توزیع فیزیکی، به شما این امکان را می‌دهد که یک نمونه اولیه را راه‌اندازی کرده و از طریق کلیک‌ها، زمان‌بندی‌ها، نظرات و سایر داده‌ها کشف کنید که کاربران شما چه چیزی را مفید می‌دانند، چه چیزی را دوست ندارند و چه چیزی برایشان اهمیت ندارد. سپس از این دانش برای بازبینی و استقرار نسخه بعدی محصول خود در یک حلقه بازخورد بسیار کارآمد از بهبود مستمر استفاده می‌کنید.

استقرار تکرارشونده، این پارادایم را با مؤلفه اجتماعی-محور صریح توزیع سنجیده تکمیل می‌کند. یک توسعه‌دهنده نرم‌افزار معمولی با سرعت محدودیت‌های عملیاتی خود حرکت می‌کند – چقدر سریع می‌تواند استنتاج‌های مناسب را از داده‌هایی که جمع‌آوری می‌کند انجام دهد؟ چقدر سریع می‌تواند کد جدیدی را بنویسد و آزمایش کند که آموخته‌هایش را به واقعیت تبدیل

کند؟ با احترام به پیامدهای فرهنگی منحصربه‌فرد هوش مصنوعی ترکیبی، OpenAIانطباق جمعی را در تصمیمات توزیع خود لحاظ می‌کند. این رویکرد را در یک پست وبلاگی در سال ۲۰۲۳ اینگونه توصیف کرد: "نکته کلیدی این است که ما معتقدیم جامعه باید زمان داشته باشد تا خود را با هوش مصنوعی که روز به روز توانمندتر می‌شود، به‌روزرسانی و تطبیق دهد، و هر کسی که تحت تأثیر این فناوری قرار می‌گیرد، باید نقش مهمی در نحوه توسعه بیشتر هوش مصنوعی داشته باشد. استقرار تکرارشونده به ما کمک کرده است تا ذینفعان مختلف را به طور مؤثرتری وارد بحث در مورد پذیرش فناوری هوش مصنوعی کنیم تا اگر تجربه مستقیم با این ابزارها نداشتند".

این رویکرد اکنون فراتر از OpenAI گسترش یافته است، زیرا سایر توسعه‌دهندگان هوش مصنوعی به سرعت از آن پیروی کردند. استفاده گسترده و عملی که استقرار تکرارشونده امکان‌پذیر می‌سازد، بدون ریسک یا بدون پیامدهای منفی نبوده است. اما از منظر ایمنی گسترده، کنار گذاشتن این رویکرد نه تنها استفاده را مهار می‌کند، بلکه یادگیری و پیشرفت را نیز مهار می‌کند. هر روز، استقرار تکرارشونده به معنای آزمایش هوش مصنوعی توسط میلیون‌ها نفر از سراسر جهان در سناریوهای واقعی زندگی است. ما بیشتر و سریع‌تر در مورد رایج‌ترین انواع استفاده و مسائل می‌آموزیم. ما بیشتر در مورد آنچه اتفاق می‌افتد وقتی سیستم‌های هوش مصنوعی با افرادی با سطوح مختلف سواد فناوری، از زمینه‌های فرهنگی متنوع، و در موقعیت‌های غیرمنتظره دنیای واقعی تعامل می‌کنند، می‌آموزیم.

ما هنوز در مراحل اولیه توسعه و پذیرش جهانی هوش مصنوعی هستیم. در چند سال آینده، زمانی که دو برابر بیشتر از اکنون از هوش مصنوعی استفاده می‌کنند، و خود هوش مصنوعی به طور قابل توجهی قدرتمندتر شده باشد، آیا نوآوری بدون اجازه و استقرار تکرارشونده هنوز منطقی خواهد بود؟ آیا نیاز به توسل به رویکردهای احتیاطی بیشتری خواهیم داشت؟

آزمون جاده: درس‌هایی از ظهور صنعت خودروسازی

گذشته ممکن است سرنخ‌هایی را در خود داشته باشد. هنگامی که منتقدان اصرار دارند که ما

باید هوش مصنوعی را به همان روشی که خودروها را تنظیم می‌کنیم، تنظیم کنیم، آنها بیش از آنچه خودشان متوجه هستند، حق با آن‌هاست. برای درک واضح‌تر پیامدهای این موضوع، بیایید نگاهی به نحوه عملکرد نوآوری بدون اجازه در اوایل دهه ۱۹۰۰ بیندازیم، در دوره‌ای که مسیر رفاه آمریکا و قدرت عامل فردی را برای صد سال آینده و حتی بیشتر تعیین کرد.

برای میلیاردها نفر در طول قرن بیستم، خودروها بدون شک قدرتمندترین ماشین‌هایی بودند که تا به حال به دست گرفته بودند. و نه فقط قدرتمندترین ماشین‌ها، بلکه ماشین‌هایی که بارها در روز، به طرق تغییردهنده زندگی و بی‌اهمیت، برای دهه‌ها از آنها استفاده می‌کردند. همانطور که جان بی. ری، مورخ در کتاب خود در سال ۱۹۷۱، "جاده و خودرو در زندگی آمریکایی" اشاره کرد، پذیرش سریع خودروها در دهه‌های اولیه قرن بیستم "تحرک شخصی از نوعی کاملاً جدید و در مقیاسی کاملاً جدید" را فراهم کرد. همانطور که ری تأکید کرد، این گذار محدود به طبقه محدودی از افراد ممتاز که به فناوری جدیدی دسترسی داشتند که دامنه و سرعت حرکت شخصی را به طرق واقعاً خارق‌العاده‌ای تقویت می‌کرد، نبود. برعکس، تنها در چند دهه، میلیون‌ها نفر این ابرقدرت واقعی را در اختیار داشتند. عصر جدیدی از خودرومحوری آغاز شد که تصورات دیرینه آزادی فردی و الگوهای زندگی را کاملاً تغییر داد و مسیر میلیون‌ها زندگی و خود جامعه را دگرگون کرد.

رایانه‌های متصل به شبکه و گوشی‌های هوشمند، نسخه قرن بیست و یکمی خودروها بوده‌اند. آن‌ها ماشین‌های تحرک شخصی هستند که راهی جدید برای فرافکنی خود در زمان و مکان فراهم می‌کنند. در عصر هوش مصنوعی، با مجهز کردن آن‌ها به هوش مصنوعی ترکیبی، رایانه‌های شخصی و گوشی‌های هوشمند به ابزارهای قدرتمندتری برای بیان عاملیت ما و اعمال اراده ما بر جهان تبدیل می‌شوند. ما در میانه یک گذار حماسی دیگر هستیم، از دوچرخه‌های ذهن، همانطور که استیو جابز زمانی رایانه‌های شخصی را توصیف کرد، به فراری‌های ذهن.

طبیعتاً، در کنار قابلیت‌های جدید، خطرات جدیدی نیز پدیدار خواهند شد. با این حال، به جای رضایت به مدل‌های بدون ریسک، باید هدف خود را درک خطراتی که در شرایط واقعی رخ

می‌دهند و به طور سیستماتیک برای مدیریت و کاهش آن‌ها تلاش کنیم، قرار دهیم. استقرار تکرارشونده اینگونه انجام می‌شود.

کمی بیش از یک قرن پیش، زمانی که جمعیت ایالات متحده ۹۷ میلیون نفر بود و شرکت فورد موتور احتمالاً سریع‌ترین شرکت با فناوری پیشرفته در جهان بود، تنها حدود ۱ میلیون خودرو در جاده‌ها وجود داشت. در سال ۱۹۱۳، در آنچه آن زمان بزرگترین کارخانه تولیدی جهان بود، یک ساختمان چهار طبقه در هایلند پارک، میشیگان، معروف به "کاخ کریستال"، فورد برای اولین بار خطوط مونتاژ متحرک را اجرا کرد. این رویکرد جدید زمان تولید یک مدل T را از حدود ۷۲۸ دقیقه به ۹۳ دقیقه کاهش داد. کاهش قیمت‌های عمده پس از آن، از ۶۹۰ دلار برای محبوب‌ترین مدل سال ۱۹۱۲ به ۳۶۰ دلار برای نسخه ۱۹۱۳ رسید.

تا سال ۱۹۱۶، فورد سالانه ۵۷۷,۰۳۶ خودرو تولید می‌کرد که تقریباً صد برابر افزایش نسبت به زمانی که شرکت مدل T را در سال ۱۹۰۸ معرفی کرد، بود. تا سال ۱۹۲۵، خط مونتاژ فورد هر ده ثانیه یک خودروی جدید را تکمیل می‌کرد. همانطور که یکی از مورخان فورد بیان کرده است، این پیشرفت‌های در بهره‌وری "کشور را روی چرخ‌ها قرار داد." در طول چند سال، دیدگاه هنری فورد در مورد یک خودروی ساده اما قابل اعتماد که افراد عادی می‌توانستند آن را بخرند، خودرو را از یک کالای لوکس برای سرگرمی ثروتمندان به یک موتور واقعی پیشرفت تبدیل کرد که به شدت قدرت عامل فردی و بهره‌وری جامعه را تقویت کرد.

اما، البته، دموکراتیک شدن خودرو نیز آسیب‌های قابل توجهی به همراه داشت. از سال ۱۹۱۳ تا ۱۹۲۳، نرخ تلفات مربوط به رانندگان به ازای هر ۱۰۰,۰۰۰ نفر بیش از سه برابر شد. در یک روز در سال ۱۹۲۷، هشت کودک در تصادفات جداگانه خودرو در شهر نیویورک و اطراف آن کشته شدند. در چهار سال اول پس از پایان جنگ جهانی اول، کتاب "مبارزه با ترافیک" اشاره می‌کند که "تعداد آمریکایی‌هایی که در تصادفات خودرو کشته شدند، بیشتر از کسانی بود که در نبرد در فرانسه جان باخته بودند".

چه می‌شد اگر در سال ۱۹۱۳، رئیس‌جمهور ویلیام هاوارد تافت، با احساس خطر در پیش رو، وارد عمل می‌شد و یک فرمان اجرایی صادر می‌کرد که تولیدکنندگان خودرو را از تولید خودروهایی که می‌توانستند سریع‌تر از ۲۵ مایل در ساعت حرکت کنند، منع می‌کرد؟ چه می‌شد اگر کنگره ایالات متحده، به طور مشابه نگران این فناوری جدید، هم در مورد خطرات جدیدی که برای عموم ایجاد می‌کرد و هم در مورد چگونگی احتمالی پیکربندی مجدد هنجارهای اقتصادی و ساختارهای قدرت، قانونی را تصویب می‌کرد که خیابان‌های شهر را ملزم می‌کرد هر صد فوت دارای سرعت‌گیر باشند؟ و سپس این را با قوانین اضافی دنبال می‌کرد که محدودیت‌های سختی را بر اندازه و وزن کلی خودرو اعمال می‌کرد؟

اکنون فرض کنیم که این مداخلات به نحوی پابرجا می‌ماندند و رانندگان عمدتاً از آنها پیروی می‌کردند. احتمالاً خیابان‌ها ایمن‌تر می‌شدند. همچنین احتمالاً شهرهای قابل پیاده‌روی‌تر، حمل و نقل عمومی بیشتر، گسترش کمتر حومه شهر، و تأثیرات زیست‌محیطی کمتری وجود داشت. اگر شما یک شهرنشین جهان‌وطن هستید، چه چیزی برای دوست نداشتن وجود دارد؟ اگر شما بیشتر به سمت ناسیونالیسم و شعارهای "آمریکا را دوباره عالی کن" تمایل دارید، ممکن است – به طرز شگفت‌انگیزی – همین فکر را کنید – زیرا زندگی در آمریکا بدون شک کندتر، کوچک‌تر، محلی‌تر و یکنواخت‌تر می‌شد.

در این مرحله، خودروها چنان جزء مرکزی زندگی آمریکایی بوده‌اند که درک کامل میزان بهبود استانداردهای زندگی توسط آنها دشوار است – اما، به طرز متناقضی، آسان است که فکر کنیم می‌توانیم زندگی بدون آنها را تصور کنیم. در قرن بیست و یکم، زمانی که می‌توانید از خانه کار کنید، زمانی که خدمات تحویل انواع مختلف هر چیزی را به درب منزل شما می‌آورند، زمانی که می‌توانید با دوستان خود در توییچ یا دیسکورد وقت بگذرانید، واقعاً به خودروها برای چه چیزی نیاز داریم؟ و با توجه به تمام مشکلاتی که ایجاد کرده‌اند، آیا اشتباه بود که آنها را با تمام وجود پذیرفتیم؟ آیا نمی‌توانستیم از آنجا به اینجا بدون آنها برسیم، هرچند با سرعت کمتری؟

واقعاً نه. امروز، ما تجربه مستقیمی از تمام مشکلاتی که خودروها ایجاد می‌کنند داریم، اما

چشم‌انداز کمتری در مورد تمام مشکلاتی که به حل آنها کمک کردند. همانطور که کلی مک‌شین، مورخ در کتاب خود در سال ۱۹۹۴، "در مسیر آسفالت"، مستند کرده است، طرفداران اولیه خودرو، خودروها را به عنوان راه‌حلی برای ترافیک، تصادفات و آلودگی شهری پیشنهاد کردند. با افزایش صنعتی‌شدن، شهرنشینی و تجارت در اواخر قرن نوزدهم در آمریکا، اسب‌ها به یک جزء اصلی از نحوه عملکرد شهرها تبدیل شده بودند، به ویژه برای تحویل بار در مسافت‌های کوتاه. در سال ۱۹۰۰، تقریباً ۱۳۰,۰۰۰ اسب هر روز واگن‌ها و کالسکه‌ها را از خیابان‌های منهتن عبور می‌دادند و تقریباً ۱.۳ میلیون پوند کود در پی خود به جای می‌گذاشتند. مک‌شین نوشت: "حتی با سرعت نسبتاً پایینی که اکثر آنها سفر می‌کردند، اسب‌ها با گاز گرفتن، لگد زدن و فرار کردن باعث تصادف می‌شدند." گزارش‌های هیئت بهداشت شهر نیویورک نشان داد که نرخ مرگ و میر سالانه شهر ناشی از تصادفات اسب بین ۵ تا ۶ در هر ۱۰۰,۰۰۰ نفر در اواخر دهه ۱۸۹۰ بود. تخمین‌ها نشان می‌دهد که یک سوم زمین‌های کشاورزی کشور برای سوخت رسانی به اسب‌ها در آن دوران مورد نیاز بود.

در آغاز قرن بیستم، محدودیت‌های نیروی واقعی اسب برای دهه‌ها به خوبی درک شده بود. از اوایل سال ۱۸۶۸، یک خبرنگار مستقر در پاریس برای نیویورک تایمز که درباره شایعات محلی "یک لوکومبیل بخار کوچک برای یک نفر برای خیابان‌ها و جاده‌های عمومی" می‌نوشت، اشاره کرد که "یک جایگزین مکانیکی ارزان برای اسب" ایده‌ای بود که زمان آن فرا رسیده بود. و همانطور که مک‌شین در کتاب خود بازگو کرد، انواع مختلفی از اتوبوس‌های جاده‌ای و سایر وسایل نقلیه بخار به دهه‌های ۱۸۲۰ و ۱۸۳۰ بازمی‌گردند. در دهه‌های بعدی ۱۸۰۰، نمونه‌های اولیه "بخارشوی"های شبیه به خودرو به طور فزاینده‌ای رایج شدند، اما مردم در مورد به اشتراک گذاشتن خیابان‌های شهری با آنها تردیدهای عمیقی داشتند. مک‌شین نوشت: "بخارشوی‌های قبل از دهه ۱۸۹۰ عمدتاً به دلیل مقررات، نه ناکارآمدی مکانیکی، شکست خوردند. آنها سرعت بیشتر، هزینه‌های عملیاتی کمتر و آلودگی کمتری نسبت به اسب‌ها ارائه می‌دادند؛ اما مردم هنوز از آنها می‌ترسیدند".

این قابل درک بود، زیرا بخارشوی‌ها یک افزودنی نسبتاً رادیکال به خیابان‌های شهر در قرن نوزدهم بودند. آنها بسیار سریع‌تر از اسب‌ها بودند. آنها دود و بخار اگزوز در پی خود باقی می‌گذاشتند. گاهی اوقات منفجر می‌شدند. بنابراین، حتی با وجود پیشرفت‌های تکنولوژیکی در دهه‌های ۱۸۷۰ و ۱۸۸۰ که منجر به بهبود عملکرد و ایمنی شد، قانونگذاران محلی به طور مداوم آنها را از مناطق شهری منع می‌کردند.

در همین دوره عمومی، با این حال، استقرار تکرارشونده در قالب دوچرخه‌ها، کابل‌کارها و واگن‌های برقی که به طور فزاینده‌ای رایج می‌شدند، اتفاق افتاد. ترافیک سریع‌تر، سنگین‌تر و متنوع‌تر می‌شد. تمام این تغییرات بر نحوه درک و استفاده ساکنان شهر از خیابان‌ها تأثیر گذاشت. آنها به عنوان فضایی برای معاشرت، تفریح و حمل و نقل دیده می‌شدند، اما با اولویت یافتن ترافیک وسایل نقلیه از هر نوع، تمرکز محدودتر شد.

تا دهه ۱۸۹۰، زمانی که خودروها با موتورهای احتراق داخلی شروع به ظاهر شدن کردند، شرایط و احساسات تغییر کرده بود. نه اینکه با ورود آنها چیزی شبیه به اجماع وجود داشت. در واقع، مقاومت زیادی وجود داشت – اما نه از نوع ممنوعیت‌های کامل که مانع از پذیرش بخارشوی‌ها شده بود. در عوض، مقررات به صورت تکرارشونده و با ظهور خطرات و آسیب‌های خاص، تدوین و اجرا شدند.

در سال ۱۹۰۱، کانتیکت اولین محدودیت سرعت برای خودروها را در آمریکا تصویب کرد. دو سال بعد، یک تاجر نیویورکی به نام ویلیام فلپس اینو مجموعه‌ای از "قوانین رانندگی" را پیشنهاد کرد که اداره پلیس شهر نیویورک در نهایت آنها را پذیرفت و اجرا کرد. در سال ۱۹۱۴، کلیولند اولین چراغ راهنمایی برقی را در آمریکا نصب کرد.

در نهایت، با این حال، این آزمایش‌های بی‌قید و بند بود که این مداخلات را آگاه کرد و به پیشرفت ادامه داد. بخش زیادی از این فعالیت‌ها نیز ذاتاً بی‌احتیاطانه بود. "تست‌های سرعت" و مسابقات اتومبیل‌رانی در اوایل دوران خودرو چنان محبوب بودند که هنری فورد، پس از ورشکستگی اولین شرکتش تنها پس از هجده ماه فعالیت، تشخیص داد که یکی از راه‌های کسب

اعتبار مهندسی که برای جذب سرمایه‌گذاران برای شرکت بعدی خود نیاز داشت، برنده شدن در یک مسابقه "جایزه بزرگ" در پیست محلی بود.

در سال ۱۹۰۴، انجمن اتومبیل آمریکا شروع به سازماندهی رالی‌های جاده‌ای به نام "تورهای گلیادن" کرد، با ده‌ها شرکت‌کننده که تلاش می‌کردند مسیرهای طولانی را در جاده‌های واگن‌های ناهموار، تنها با نقشه‌های دست‌ساز به عنوان راهنما، تکمیل کنند. در اوایل دهه ۱۹۲۰، نمایندگی‌های استودبیکر در کالیفرنیای جنوبی، "مسابقات اقتصادی" سالانه را برای نشان دادن قابلیت اطمینان، دوام و کارایی این برند در طیف وسیعی از شرایط رانندگی مختلف تبلیغ می‌کردند. در نوعی معیار خودرویی، استودبیکرهایی که قبلاً حداقل ۵۰,۰۰۰ مایل در خدمت بودند، از مرکز شهر لس‌آنجلس به یک نقطه میانی در ارتفاع یک مایلی در کوهستان در دریاچه آرووهد حرکت می‌کردند، سپس به شهر بازمی‌گشتند. برنده خودرویی بود که کمترین میزان بنزین، روغن و آب را برای تکمیل سفر ۱۷۲ مایلی مصرف کرده بود.

سفر در مسافت‌های طولانی در جاده‌های نامناسب در مناطق دورافتاده و چالش‌برانگیز با وسایل نقلیه‌ای که هنوز از نظر تکنولوژیکی در حال تکامل بودند، به این معنی بود که خطر و گاهی اوقات حوادث ویرانگر حتی رویدادهای غیرمسابقه‌ای را نیز مشخص می‌کرد. اما چنین تلاش‌های خطرناکی نیز تأثیرات مثبتی بر ایمنی داشت، زیرا الزام به عملکرد خوب در مسابقات، رالی‌های جاده‌ای و مسابقات اقتصادی به خودروهای قابل اعتمادتر و توانمندتر منجر شد. علاوه بر این، در شگفت‌انگیزترین جنبه نوآوری بدون اجازه‌ای که به شکل‌گیری ظهور خودرو کمک کرد، نوآوری‌هایی که از این محیط باز و تجربی پدید آمدند، تنها ماهیت تکنولوژیکی نداشتند.

ویلیام فلپس اینو هیچ مقام عمومی نداشت زمانی که شروع به حمایت از قوانین جدید رانندگی خود کرد؛ او صرفاً به عنوان یک شهروند فعال عمل می‌کرد. هنگامی که خودروها به تعداد بیشتری در شهر نیویورک ظاهر شدند و سازمان‌های شهری برای ایجاد نظم کاری نکردند، اینو مجموعه‌ای از شیوه‌های توصیه شده را تدوین کرد و کمیسر پلیس شهر را متقاعد کرد که کمیسر در واقع اختیار پذیرش و اجرای آنها را دارد.

هنگامی که کمیسر به اینو گفت که دپارتمان او بودجه‌ای برای چاپ جزوه‌هایی که برای اطلاع‌رسانی به عموم در مورد قوانین جدید لازم است، ندارد، اینو هزینه ۱۰۰,۰۰۰ نسخه اول را پرداخت کرد – و سپس به مدت هفت سال "تمام چاپ‌های مورد استفاده در کارهای ترافیکی" را پرداخت کرد. به طور مشابه، این تنظیم‌کنندگان فدرال نبودند که در آغاز عصر خودرو مسئولیت افزایش کارایی سوخت را بر عهده گرفتند؛ بلکه فروشندگان به دنبال یک نکته فروش بودند.

در این سال‌های اولیه استقرار تکرارشونده، مقاومت پابرجا بود، شاید حتی با شروع گرفتن خودرو در فرهنگ، افزایش یافت. منتقدان اولیه خودروها، که به آنها "واگن‌های شیطان" و "ماشین‌های مرگ" می‌گفتند، آنها را اسباب‌بازی برای ثروتمندان توصیف می‌کردند. آنها رانندگان را "دیوانگان سرعت" و "قلدرهای جاده" مبتلا به "هاری بنزینی" برچسب‌گذاری می‌کردند.

تا سال ۱۹۰۷، اکونومیست هنوز از "پیروزی اسب" حمایت می‌کرد. در همان سال، در تور سالانه گلیادن، کسی دو گلوله به بدنه خودروی یکی از شرکت‌کنندگان شلیک کرد. با این حال، راننده در پایان مسابقه به خبرنگاران گفت که محلی‌هایی که با آنها روبرو شده بود، در واقع بسیار دوستانه‌تر از سال‌های گذشته بودند.

در سال ۱۹۰۹، کشاورزان نزدیک ساکرامنتو "خندق‌هایی را در جاده‌ها حفر کرده بودند... و عملاً سیزده خودرو را به دام انداخته بودند." در اروپا، جایی که اولین خودروهای احتراق داخلی اختراع شده بودند، مخالفت رسمی‌تر بود. بریتانیا مالیات‌های سنگینی بر خودروها و بنزین وضع کرد. در فرانسه، که برای مدت کوتاهی از آمریکا در تولید سالانه خودرو پیشی گرفته بود، ژاندارم‌ها در پاریس دستور داشتند لاستیک رانندگانی را که از حد مجاز سرعت تجاوز می‌کردند، سوراخ کنند.

کاهش مؤثر آسیب‌هایی که خودروها ایجاد می‌کنند، فرآیندی است که تا به امروز از طریق مقررات رسمی، نوآوری‌های تکنولوژیکی، تغییر هنجارهای فرهنگی و اولویت‌های در حال تغییر ادامه دارد. اما با گذشت زمان، دقیقاً به این دلیل که این یک فرآیند عمومی بود، و به دلیل اینکه تکثیر مدل T و سایر خودروهای مقرون به صرفه این ابرقدرت جدید را در دسترس افراد بیشتری

قرار داد، آنچه در ابتدا برای بسیاری بیگانه و تهدیدآمیز به نظر می‌رسید، به زودی به آرمان، سپس قابل دستیابی، و سپس عادی تبدیل شد.

به طور خلاصه، مردم فرصت انتخاب داشتند و خودروها را انتخاب کردند. آنها خودروها را انتخاب کردند زیرا خودروها عاملیت فردی را به شدت تقویت می‌کردند، به گونه‌ای که کالسکه، راه‌آهن، دوچرخه، کشتی‌های بخار و قطارهای سبک به سادگی نمی‌توانستند برابر باشند. اساساً خودروها به عنوان گره‌های جدیدی در یک شبکه خیابان‌ها، جاده‌ها و بزرگراه‌های روزافزون متراکم و گسترده عمل کردند که به میلیون‌ها نفر امکان می‌داد تا زمان و مکان را به روش‌های بسیار فردی در هم بشکنند. با رشد این شبکه، گزینه‌های بیشتری برای همه، از جمله کسانی که هرگز دست به فرمان نگرفتند، افزایش یافت. شبکه‌های حمل و نقل بهتر بازارهای بزرگ‌تری را برای انواع کالاها و خدمات ایجاد کردند. بازارهای بزرگ‌تر و هزینه‌های توزیع کمتر منجر به قیمت‌های پایین‌تر شد که به نوبه خود به بازارهای بیشتر و حتی بزرگ‌تر منجر شد. و نه فقط بازارهای تجاری بزرگ‌تر، بلکه بازارهای بزرگ‌تر و متنوع‌تر برای همه چیز، از جمله مدارس، فرصت‌های شغلی، مکان‌های عبادت، سبک زندگی‌ها.

خودرومحوری یک محرک برای خودتعیین‌گری و نوآوری فرهنگی بود که در سراسر جامعه گسترش یافت. نوآوری بدون اجازه، پیشرفت‌های تکنولوژیکی، پذیرش و تطبیق فرهنگی گسترده و افزایش سطوح ایمنی را ممکن ساخت. مقررات رسمی در این فرآیند نقش داشت، اما خود آن نیز عمدتاً به صورت تکرارشونده و نه پیشگیرانه اعمال شد.

با گذشت زمان، نتایج این رویکرد خیره‌کننده بود. رانندگی امروز یک پیشنهاد بدون ریسک نیست – این دلیل اصلی تمام سرمایه‌گذاری‌هایی است که برای توسعه وسایل نقلیه خودمختار انجام شده است. اما رانندگی در طول سال‌ها به طور قابل توجهی ایمن‌تر شده است. شورای ملی ایمنی در وب‌سایت خود می‌نویسد: "در سال ۱۹۲۳، اولین سالی که مایل‌های رانده شده تخمین زده شد، نرخ مرگ و میر وسایل نقلیه موتوری ۱۸.۶۵ مرگ به ازای هر ۱۰۰ میلیون مایل رانده شده بود. از سال ۱۹۲۳، نرخ مرگ و میر مایل ۹۳٪ کاهش یافته و اکنون ۱.۳۳ مرگ به ازای هر ۱۰۰

میلیون مایل رانده شده است".

به عنوان یک الگوی کلی، رویکردی که ما با خودرومحوری در پیش گرفتیم، برای هوش مصنوعی نیز منطقی است. به جای تکیه بر تنظیم‌کنندگان و کارشناسان صنعت برای توسعه و بهبود هوش مصنوعی در پشت درهای بسته، به روش‌های متمرکز و غیردموکراتیک، ما باید به استقرار تکرارشونده ادامه دهیم که به ما کمک می‌کند تا بهتر درک کنیم که مردم چگونه از هوش مصنوعی استفاده می‌کنند، مسائل چگونه با افزایش استفاده توسعه می‌یابند و بر اساس آن تنظیمات را انجام دهیم. از طریق این فرآیند، مردم حسی مستقیم از اینکه چقدر قابلیت‌های جدیدی که هوش مصنوعی ارائه می‌دهد را ارزش می‌گذارند یا نمی‌گذارند، به دست خواهند آورد.

این به نوبه خود کمک خواهد کرد تا مشخص شود چه نوع خطرات و مصالحه هایی منطقی به نظر می‌رسند. اگر تمام آنچه هوش مصنوعی برای اکثر مردم به ارمغان می‌آورد، راهی مناسب برای ساخت تصاویر برای کارت‌های تولد خانگی باشد، ما به عنوان یک جامعه احتمالاً تحمل ریسک زیادی را نخواهیم داشت. از طرف دیگر، اگر اکثر مردم هوش مصنوعی را به عنوان فناوری‌ای ببینند که می‌تواند عاملیت آن‌ها را تقویت کرده و انتخاب‌های زندگی‌شان را گسترش دهد، به همان روشی که خودرو در طول یک قرن و نیم گذشته انجام داده است، آنگاه ما سطح بالاتری از خطا و ریسک را در تعقیب این پاداش‌های بزرگ‌تر تحمل خواهیم کرد.

نتیجه‌گیری

این فصل به طور جامع به بررسی دو رویکرد کلیدی در زمینه توسعه و تنظیم مقررات فناوری، به ویژه هوش مصنوعی، یعنی "اصل احتیاطی" و "نوآوری بدون اجازه" پرداخت. ما استدلال کردیم که هرچند احتیاط و نظارت در عصر هوش مصنوعی حیاتی است، اما سرعت و طبیعت تطبیق‌پذیر نوآوری می‌تواند خود به بهترین ابزار برای تضمین ایمنی تبدیل شود. تأخیرهای ناشی از مقررات سخت‌گیرانه نه تنها پیشرفت را کند می‌کند، بلکه می‌تواند مانع از ظهور راه‌حل‌های ایمنی جدید نیز شود. در مقابل، رویکرد "نوآوری بدون اجازه" که بر آزمایش، بازخورد سریع و

استقرار تکرارشونده تأکید دارد، منجر به محصولات و خدمات ایمن‌تر و قابل اعتمادتر می‌شود.

با الهام از تاریخ توسعه اینترنت و صنعت خودروسازی، نشان داده شد که چگونه سیاست‌های حمایتی از نوآوری و همچنین چرخه طبیعی بهبود محصول، به کاهش مخاطرات در طول زمان انجامیده است. تجربه خودروها نشان داد که پذیرش یک فناوری قدرتمند با خطرات اولیه همراه است، اما با گذشت زمان، مداخلات تدریجی، نوآوری‌های تکنولوژیکی و تطبیق‌های فرهنگی منجر به سطوح بالاتری از ایمنی می‌شوند. در نهایت، مردم با انتخاب خود، ارزش این فناوری‌ها را در تقویت عاملیت فردی و گسترش انتخاب‌های زندگی درک کردند و سطوح مشخصی از ریسک را پذیرفتند. این الگو می‌تواند به عنوان یک راهنما برای توسعه و حکمرانی هوش مصنوعی عمل کند: تشویق به استقرار تکرارشونده، جمع‌آوری بازخورد گسترده از کاربران و جامعه، و تنظیم مقررات و توسعه بر اساس درک واقعی از نحوه استفاده و پیامدهای فناوری در دنیای واقعی. این رویکرد دموکراتیک و داده‌محور، مسیری مؤثرتر برای دستیابی به آینده‌ای ایمن و مبتکرانه با هوش مصنوعی ارائه می‌دهد.

نکات کلیدی

- **نوآوری به مثابه ایمنی** :توسعه سریع و انطباقی، به ویژه در نرم‌افزارهای اینترنتی، می‌تواند از طریق چرخه‌های محصول کوتاه‌تر و به‌روزرسانی‌های مکرر، به ایمنی بیشتر منجر شود.
- **اهمیت زمینه جهانی** :حفظ قدرت نوآوری در یک محیط جهانی به آمریکا امکان می‌دهد تا ارزش‌های دموکراتیک را در فناوری‌های هوش مصنوعی تزریق کرده و امنیت ملی و نفوذ جهانی خود را تقویت کند.
- **اصل احتیاطی** :(**Precautionary Principle**) این رویکرد فناوری‌های جدید را "مقصر تا اثبات بی‌گناهی" تلقی کرده و بر کنترل پیشگیرانه تأکید دارد، که در مثال‌هایی مانند تنظیم مقررات دارویی، انرژی هسته‌ای، GMOsو برخی قوانین

هوش مصنوعی دیده می‌شود.

- **نوآوری بدون اجازه** :(Permissionless Innovation) این رویکرد بر ایجاد فضای کافی برای آزمایش، خطا و تطبیق تأکید دارد، به ویژه زمانی که آسیب‌های ملموسی وجود ندارند، و معتقد است که نوآوری ذاتاً قدرت نظارتی دارد.
- **نقش سیاست‌های دولتی** :ظهور اینترنت به عنوان یک موتور رشد، تا حد زیادی نتیجه سیاست‌های عمدی برای حمایت از نوآوری بدون اجازه بود، از جمله ماده ۲۳۰ که مسئولیت واسطه‌های آنلاین را محدود می‌کند.
- **توسعه تکرارشونده** :(Iterative Deployment) این مدل توسعه نرم‌افزار، با راه‌اندازی نمونه‌های اولیه، جمع‌آوری بازخورد کاربران و به‌روزرسانی‌های مداوم، منجر به بهبود سریع‌تر و محصولات ایمن‌تر می‌شود.
- **درس‌هایی از صنعت خودروسازی** :تاریخچه خودرو نشان می‌دهد که فناوری‌های قدرتمند ابتدا با خطرات زیادی همراه بودند، اما نوآوری بدون اجازه، آزمایش‌های گسترده و تنظیم مقررات تدریجی، در نهایت به ایمنی چشمگیر و پذیرش عمومی منجر شد.
- **عاملیت فردی و انتخاب اجتماعی** :پذیرش گسترده فناوری‌هایی مانند خودرو و هوش مصنوعی، به دلیل توانایی آن‌ها در تقویت عاملیت فردی و گسترش انتخاب‌های زندگی، با تحمل سطح معینی از ریسک همراه است که جامعه بر اساس ارزش‌گذاری خود تعیین می‌کند.
- **داده‌محوری در عصر مدرن** :در دسترس بودن حجم عظیمی از داده‌ها در عصر اینترنت، تحلیل ریسک را نسبت به گذشته بسیار دقیق‌تر کرده و می‌تواند برخی از دلایل اولیه برای توسل به اصل احتیاطی (فقدان داده) را کمرنگ کند.

سؤالات تفکربرانگیز

۱. با توجه به تفاوت‌های بنیادین بین توسعه فناوری‌های فیزیکی مانند خودرو و

فناوری‌های دیجیتال مانند هوش مصنوعی، آیا مدل "نوآوری بدون اجازه" و "توسعه تکرارشونده" همچنان برای هوش مصنوعی در مراحل پیشرفته‌تر توسعه، یعنی زمانی که مدل‌ها بسیار قدرتمندتر و فراگیرتر شوند، مناسب باقی خواهد ماند؟

۲. چگونه می‌توان بین حفظ مزیت رقابتی در نوآوری هوش مصنوعی در سطح جهانی و اطمینان از ایمنی و اخلاقی بودن توسعه آن، بدون ایجاد تأخیرهای مضر، تعادل برقرار کرد؟

۳. اگرچه "اصل احتیاطی" در گذشته در زمینه‌هایی مانند محیط زیست نقش مهمی داشته است، اما چگونه می‌توان از سوءاستفاده از آن به عنوان ابزاری برای سرکوب نوآوری، به ویژه در مواردی که شواهد قوی از آسیب وجود ندارد، جلوگیری کرد؟

۴. آیا مدل نظارتی "همه‌گیر، دموکراتیک و داده‌محور" که در این فصل برای هوش مصنوعی پیشنهاد شد، می‌تواند به طور مؤثر به چالش‌های پیچیده‌ای مانند سوگیری الگوریتمی، انتشار اطلاعات نادرست یا نگرانی‌های مربوط به حریم خصوصی در مقیاس جهانی پاسخ دهد؟

۵. با توجه به اینکه پذیرش فناوری‌های متحول‌کننده مانند خودرو در نهایت به انتخاب‌های فردی و اجتماعی بستگی داشت، چه مکانیزم‌هایی را می‌توان برای افزایش آگاهی عمومی و مشارکت دموکراتیک در شکل‌دهی به آینده هوش مصنوعی پیشنهاد کرد؟

۳۰ جمله کلیدی

۱. با افزایش قدرت مدل‌های هوش مصنوعی، تدابیر ایمنی مؤثر و نظارت گسترده انسانی حیاتی‌تر می‌شوند.

۲. مقررات طراحی شده برای به تأخیر انداختن توسعه، نه تنها رفاه جهانی را مختل می‌کنند، بلکه اثرات مثبت نوآوری بر ایمنی را نیز به تعویق می‌اندازند.

۳. حفظ قدرت نوآوری آمریکا در بستر جهانی، یک اولویت کلیدی ایمنی برای تزریق ارزش‌های دموکراتیک به هوش مصنوعی است.

۴. توسعه سریع به معنای توسعه انطباقی است که منجر به چرخه‌های محصول کوتاه‌تر، به‌روزرسانی‌های مکرر و محصولات ایمن‌تر می‌شود.

۵. در نرم‌افزارهای اینترنتی، نوآوری برای دستیابی به ایمنی مؤثرتر از تنظیم مقررات برای ایمنی است.

۶. منتقدان، ایمنی را با احتیاط و تأخیر مرتبط می‌دانند، در حالی که دیدگاه نوآوری، سرعت را نیز عاملی در ایمنی می‌بیند.

۷. "اصل احتیاطی" معتقد است که فناوری‌های جدید "مقصرند تا زمانی که بی‌گناهی‌شان اثبات شود".

۸. قانونگذارانی مانند تد لیو، لوسی پاول و الکساندرا ون هافلن، از رویکرد احتیاطی برای تنظیم هوش مصنوعی حمایت می‌کنند.

۹. "نوآوری بدون اجازه" فضای کافی برای آزمایش و تطبیق فراهم می‌کند، به ویژه در غیاب آسیب‌های ملموس.

۱۰. کارآفرینان و تنظیم‌کنندگان هر دو به دنبال بهبود جهان هستند؛ اولی با محصولات، دومی با قوانین.

۱۱. معرفی استارت‌های برقی در خودروها نمونه‌ای از نوآوری است که هم ایمنی را افزایش داد و هم بازار را دموکراتیک کرد.

۱۲. اینترنت، عملکرد نظارتی نوآوری بدون اجازه را با گسترش پایگاه‌های کاربری و حلقه‌های بازخورد تقویت کرده است.

۱۳. نوآوری بدون اجازه در عصر اینترنت به بازخورد جامع‌تر، چرخه‌های محصول کوتاه‌تر و بهبودهای سریع‌تر منجر می‌شود.

۱۴. ادعای "اینترنت غرب وحشی" توسعه بی‌قید و بند، نقش سیاست‌های عمدی حمایت از نوآوری را نادیده می‌گیرد.

۱۵. سیاست‌هایی مانند کاهش محدودیت‌های NSFNET و تصویب ماده ۲۳۰، به رشد بی‌سابقه اینترنت کمک کرد.

۱۶. ماده ۲۳۰، با محافظت از پلتفرم‌ها در برابر مسئولیت محتوای شخص ثالث، اینترنت را از پخش‌محور به تعاملی تبدیل کرد.

۱۷. "چارچوب تجارت اقتصادی جهانی" کلینتون و گور، رویکردی "دست‌بردارانه" و بازارمحور را برای اینترنت ترویج کرد.

۱۸. این سیاست‌ها، یک محیط نوآورانه ایجاد کرد که منجر به میلیون‌ها شغل جدید و رشد اقتصادی در دهه ۱۹۹۰ شد.

۱۹. نگرانی‌ها در مورد حریم خصوصی و اطلاعات غلط، درخواست‌ها برای رویکردهای احتیاطی را افزایش داده است.

۲۰. نوآوری بدون اجازه، پیشرفت‌ها در خودروهای الکتریکی، CRISPR، ذخیره‌سازی انرژی و بسیاری از فناوری‌های دیگر را تسریع کرده است.

۲۱. مطالبه "توقف کامل توسعه هوش مصنوعی" توسط مؤسسه آینده زندگی، استاندارد "شک معقول" حقوق کیفری را وارونه می‌کند.

۲۲. ظرفیت بی‌سابقه امروز برای تولید و تحلیل داده‌ها، بسیاری از دلایل اصلی برای توسل به اصل احتیاطی را کاهش می‌دهد.

۲۳. توسعه تکرارشونده (Iterative deployment) با راه‌اندازی نمونه‌های اولیه، جمع‌آوری داده‌ها و بهبود مستمر، یادگیری را سریع‌تر می‌کند.

۲۴. OpenAI و سایر توسعه‌دهندگان هوش مصنوعی از استقرار تکرارشونده برای

جمع‌آوری بازخورد و انطباق جامعه استفاده می‌کنند.

۲۵. تاریخچه صنعت خودروسازی نشان می‌دهد که خودروها ابتدا با خطرات زیادی همراه بودند، اما به تدریج ایمن‌تر شدند.

۲۶. اسب‌ها در اوایل قرن بیستم عامل اصلی آلودگی و تصادفات شهری بودند و خودروها راه‌حلی برای این مشکلات ارائه دادند.

۲۷. "بخارشوی‌ها" به دلیل ترس عمومی و مقررات سخت‌گیرانه، با وجود مزایای تکنولوژیکی اولیه، شکست خوردند.

۲۸. مقررات اولیه خودرو، مانند محدودیت‌های سرعت و چراغ‌های راهنمایی، به صورت تکرارشونده و در پاسخ به خطرات ملموس اعمال شدند.

۲۹. آزمایش‌های بی‌قید و بند مانند مسابقات و رالی‌های جاده‌ای، به بهبود قابلیت اطمینان و ایمنی خودروها کمک کردند.

۳۰. رویکرد خودروسازی، که بر استقرار تکرارشونده، درک کاربرد واقعی و تطبیق تدریجی متمرکز بود، الگویی برای مدیریت مخاطرات هوش مصنوعی ارائه می‌دهد.

فصل ۷

جی‌پی‌اس اطلاعاتی: نقشه راه هوش مصنوعی

"در عصر حاضر، مرزهای اکتشاف تنها جغرافیایی نیستند؛ بلکه عمیقاً در قلمروهای پیچیده دانش و اطلاعات نیز گسترده شده‌اند. هوش مصنوعی، به مثابه یک سیستم موقعیت‌یاب جهانی نوین، مسیرهای تازه‌ای را برای ناوبری و کشف در این گستره نامتناهی گشوده است".

اهداف اصلی فصل:

- معرفی سیر تکاملی و تأثیرات گسترده سیستم موقعیت‌یاب جهانی (GPS) بر زندگی مدرن.
- تبیین قیاس مفهومی میان GPS فیزیکی و مدل‌های زبانی بزرگ (LLMs) به عنوان "جی‌پی‌اس اطلاعاتی."
- بررسی چگونگی نقش LLMs در افزایش توانمندی فردی، دموکراتیزه کردن دسترسی به دانش و ارتقای مهارت‌ها.
- تحلیل پتانسیل‌های هوش مصنوعی عامل‌گرا و اهمیت تعاملات مبتنی بر گفتگو با سیستم‌های هوش مصنوعی.
- استخراج درس‌هایی از تجربه پذیرش GPS برای هدایت توسعه و توزیع فراگیر هوش مصنوعی در آینده.

چکیده

این فصل به بررسی عمیق مفهوم "جی‌پی‌اس اطلاعاتی" می‌پردازد و آن را به عنوان یک نقشه راه برای درک نقش هوش مصنوعی، به‌ویژه مدل‌های زبانی بزرگ(LLMs) ، در هدایت زندگی در قرن بیست‌ویکم معرفی می‌کند. با آغاز از ریشه‌های سیستم موقعیت‌یاب جهانی (GPS) که از یک پروژه نظامی به یک ابزار عمومی ضروری تبدیل شد، این فصل نحوه مواجهه دولت‌ها با فناوری‌های تحول‌آفرین را مورد واکاوی قرار می‌دهد. سپس، به تشریح شباهت‌ها و تفاوت‌های بنیادین میان GPS و LLMs می‌پردازد و نشان می‌دهد که چگونهLLMs ، به عنوان ابزارهایی برای تجزیه و تحلیل و نقشه‌برداری از جریان‌های زبانی، توانایی‌های ناوبری اطلاعاتی را در اختیار افراد قرار می‌دهند. تمرکز اصلی بر روی قابلیت LLMs برای افزایش توانمندی فردی، دموکراتیزه کردن دسترسی به دانش و کاهش نابرابری‌های اطلاعاتی است. همچنین، پتانسیل هوش مصنوعی عامل‌گرا و اهمیت تعاملات مداوم و مبتنی بر گفتگو با این سیستم‌ها برای بهبود دقت و کاهش سوگیری‌ها مورد بحث قرار می‌گیرد. در نهایت، با استناد به تجربه پذیرش گستردهGPS ، این فصل بر لزوم سرعت بخشیدن به توزیع گسترده و دسترس‌پذیری هوش مصنوعی برای خلق منافع عظیم اجتماعی و اقتصادی تأکید می‌کند.

مقدمه

در دهه‌های اخیر، بشر شاهد تحولات چشمگیری در حوزه‌های فناوری بوده است که هر یک به نوبه خود، زندگی روزمره، اقتصاد و ساختارهای اجتماعی را دستخوش تغییرات بنیادین کرده‌اند. در میان این نوآوری‌ها، دو فناوری با ماهیت ناوبری و اطلاعاتی خود برجسته‌اند: سیستم موقعیت‌یاب جهانی (GPS) و مدل‌های زبانی بزرگ (LLMs) که نمادی از اوج‌گیری هوش مصنوعی محسوب می‌شوند. تا اوایل دهه ۲۰۰۰، "نقشه‌های کاغذی" صرفاً "نقشه" خوانده می‌شدند؛ ابزارهایی دست‌وپاگیر، دشوار برای به‌روزرسانی و در رانندگی بالقوه خطرناک. این دوره نشان‌دهنده عصر ناوبری فیزیکی بود که با محدودیت‌های قابل توجهی همراه بود. اما با توسعه و فراگیر شدنGPS ، انقلاب بزرگی در نحوه درک ما از فضا و حرکت در آن به وقوع پیوستGPS .، که ابتدا یک پروژه نظامی بود، به تدریج به یک ابزار عمومی جهانی تبدیل شد که نه تنها ناوبری را دگرگون کرد، بلکه کاربردهای بی‌شماری در صنایع مختلف از جمله لجستیک، مخابرات، کشاورزی و واکنش اضطراری یافت.

اکنون، در اوج قرن بیست‌ویکم، هوش مصنوعی و به‌ویژه مدل‌های زبانی بزرگ، در آستانه ایجاد تحولی مشابه در دنیای اطلاعات قرار دارند. همانطور که GPS به ما قدرت ناوبری در جهان فیزیکی را با دقت بی‌سابقه بخشید، LLMsنیز ظرفیت بی‌نظیری برای هدایت ما در "جهان اطلاعاتی" پیچیده و رو به گسترش فراهم آورده‌اند. این فصل قصد دارد با ترسیم یک قیاس دقیق میان این دو فناوری، مفهوم "جی‌پی‌اس اطلاعاتی" را معرفی کند و نشان دهد که چگونه LLMs می‌توانند به عنوان ابزاری حیاتی برای افزایش توانمندی فردی، دموکراتیزه کردن دسترسی به دانش و حل چالش‌های اطلاعاتی معاصر عمل کنند. با درک این قیاس، می‌توانیم چگونگی طراحی و پیاده‌سازی هوش مصنوعی را در راستای آرمان‌های دموکراتیک خودتعیین‌گری و مشارکت گسترده‌تر شهروندان، با اولویت دادن به عاملیت فردی، بررسی کنیم.

بخش اول: سیر تحول سیستم موقعیت‌یاب جهانی(GPS)

تاریخچه GPS از دهه ۱۹۷۰ میلادی آغاز می‌شود، زمانی که وزارت دفاع ایالات متحده در سال ۱۹۷۳ کار بر روی سیستمی را آغاز کرد که بعدها به سیستم موقعیت‌یاب جهانی تبدیل شد. این سیستم بر پایه استفاده از سیگنال‌های رادیویی از چندین ماهواره در مدار میانی زمین استوار بود تا بتواند مختصات جغرافیایی گیرنده‌ها را بر روی زمین با دقت بالا تعیین کند. در اواخر دهه ۱۹۷۰، نیروی هوایی ایالات متحده آزمایش‌های اولیه این سیستم را که منحصراً برای مصارف نظامی طراحی شده بود، آغاز کرد. نقطه عطفی در توسعه و دسترسی عمومی به GPS در سال ۱۹۸۳ رخ داد، زمانی که یک هواپیمای مسافربری کره‌ای که از مسیر خود منحرف شده و وارد حریم هوایی شوروی شده بود، توسط ارتش شوروی سرنگون شد. در پی این فاجعه، رئیس‌جمهور رونالد ریگان اعلام کرد که پس از عملیاتی شدن کامل GPS ، ایالات متحده آن را برای استفاده غیرنظامی نیز در دسترس قرار خواهد داد. این اقدام، یک گام بشردوستانه بزرگ بود که راه را برای ایجاد یک ابزار عمومی جهانی و رایگان هموار کرد که امروزه به منبعی ضروری برای ناوبری در قرن بیست‌ویکم تبدیل شده است.

با این حال، مسیر دسترسی عمومی به GPS بدون چالش نبود. در سال ۱۹۸۹، شرکت مگلان اولین گیرنده دستی را برای بازار مصرف‌کننده با قیمتی حدود ۷,۷۲۷,۰۲ دلار (به ارزش ۲۰۲۳) معرفی کرد. اما کمتر از یک سال بعد، سیاست "دسترسی انتخابی (Selective Availability) " توسط نیروی هوایی ایالات متحده اعمال شد. این سیاست که ناشی از نگرانی‌های امنیت ملی و سوءاستفاده‌های احتمالی بود، عمداً سیگنال‌های موجود برای استفاده غیرنظامی را مخدوش کرد تا دقت آن را ده برابر کمتر از نسخه نظامی کند. با این وجود، تقاضای قابل توجهی از سوی شرکت‌های حمل‌ونقل و لجستیک، نقشه‌برداران و سایر کاربران تجاری اولیه وجود داشت که باعث شد مگلان و دیگر تولیدکنندگان تا سال ۱۹۹۴ به فروش حدود یک میلیارد دلار در سال دست یابند. این رشد و پتانسیل‌های بیشتر، به‌ویژه پس از فروپاشی اتحاد جماهیر شوروی که

نگرانی‌های امنیتی جهانی را کاهش داده بود، رئیس‌جمهور بیل کلینتون را در سال ۱۹۹۶ بر آن داشت تا اعلام کند که دولت به زودی به سیاست دسترسی انتخابی پایان خواهد داد و به غیرنظامیان امکان دسترسی به همان سطح خدماتی را که ارتش ایالات متحده از آن بهره‌مند بود، خواهد داد. استدلال بر این بود که یک GPS کاملاً در دسترس، سرمایه‌گذاری و نوآوری بخش خصوصی را تقویت، نرخ پذیرش را تسریع و ارزش کلی GPS را به عنوان یک کالای عمومی جهانی به طور چشمگیری افزایش خواهد داد.

در نیمه‌شب ۳۰ مه ۲۰۰۰، با اجرایی شدن این سطح جدید از دسترسی، تنها حدود ۴ میلیون کاربر غیرنظامی GPS در سراسر جهان وجود داشتند. اما قیمت یک گیرنده ساده به حدود ۱۰۰ دلار کاهش یافته بود و این دستگاه‌ها به سرعت کوچک‌تر و ارزان‌تر می‌شدند به طوری که شروع به ادغام شدن در لپ‌تاپ‌ها، ساعت‌ها و تلفن‌های همراه کردند. یک تصمیم سیاستی دولتی متعاقب، بازار را بیش از پیش تقویت کرد: الزامات FCC که از اوایل دهه ۲۰۰۰ به اجرا درآمد، شرکت‌های مخابراتی را ملزم کرد که در صورت شماره‌گیری ۹۱۱، مختصات طول و عرض جغرافیایی تماس‌گیرنده را به مراکز اورژانس ارائه دهند. شرکت‌ها این تعهد را عمدتاً با گنجاندن GPS در تلفن‌ها محقق کردند. این پیشرفت‌ها، GPS را به یک فناوری فراگیر تبدیل کرد که تأثیرات اقتصادی عظیمی داشت. یک گزارش در سال ۲۰۱۹ تخمین زد که فناوری‌های GPS از سال ۱۹۸۴ تا ۲۰۱۷، ۱٫۴ تریلیون دلار منافع اقتصادی برای بخش عمومی ایجاد کرده‌اند که ۹۰ درصد آن در هفت سال پایانی این دوره محقق شده است. علاوه بر ناوبری پیچیده‌ای که رایج‌ترین کاربرد آن است، اطلاعات دقیق زمان‌بندی GPS برای همگام‌سازی ساعت‌ها در شبکه‌های مخابراتی استفاده می‌شود که به حفظ وضوح و عدم تأخیر مکالمات تلفن همراه کمک می‌کند. در بلایای طبیعی و سایر شرایط اضطراری، اولین امدادگران از پهپادهای مجهز به GPS برای یافتن افراد گمشده، نقشه‌برداری سریع مناطق آسیب‌دیده و حتی تحویل لوازم به کسانی که دسترسی به آنها دشوار است، استفاده می‌کنند. تکنیک‌های کشاورزی دقیق که GPS امکان‌پذیر می‌سازد، تولید

محصولات ارگانیک را اقتصادی‌تر می‌کند. این فناوری، که معمار اصلی آن، سرهنگ بردفورد پارکینسون، آن را به عنوان تلاشی برای "انداختن ۵ بمب در یک سوراخ" توصیف کرده بود، به طور گسترده‌ای فراتر از اهداف اولیه خود عمل کرده است.

بخش دوم: جی‌پی‌اس اطلاعاتی: قیاس مفهومی برای مدل‌های زبانی بزرگ (LLMs)

سیر تکامل GPS و تأثیرات گسترده آن، الگوی روشنی از نتایج مثبت را ارائه می‌دهد که می‌تواند زمانی حاصل شود که دولت رویکردی حامی فناوری و نوآوری را در پیش گیرد و کارآفرینی بخش خصوصی را به عنوان یک دارایی استراتژیک برای دستیابی به منافع عمومی ببیند. این تجربه تاریخی، به عنوان یک قیاس قدرتمند برای درک پتانسیل هوش مصنوعی، به ویژه مدل‌های زبانی بزرگ، عمل می‌کند. همانطور که GPS توانایی ما را در تبدیل "داده‌های بزرگ" جغرافیایی و زمان‌بندی به "دانش بزرگ" برای ارائه راهنمایی مبتنی بر زمینه در جهان فیزیکی افزایش داد، LLMsنیز ابزارهایی برای اهرم کردن ظرفیت ما برای تبدیل داده‌های عظیم اطلاعاتی به دانش قابل استفاده در جنبه‌های مختلف زندگی هستند.

مهم‌ترین شباهت میان GPS و LLMs در تقویت عاملیت فردی نهفته است. در حالی که GPSو خدمات تجاری ساخته شده بر روی آن به ما امکان می‌دهند تا با دانشی که به طور مداوم به‌روز می‌شود در جهان فیزیکی حرکت کنیم و در هر پیچ و خم، با نشان دادن موقعیت ما، آنچه در نزدیکی ماست و موانعی که باید از آنها اجتناب کنیم، عاملیت فردی را افزایش می‌دهد؛ LLMs و عوامل مکالمه‌ای ساخته شده بر روی آن‌ها نیز به طور مشابه عمل می‌کنند. آن‌ها ظرفیت ما را برای ناوبری در محیط‌های اطلاعاتی پیچیده و در حال گسترش که زندگی در قرن بیست‌ویکم را تعریف می‌کنند، افزایش می‌دهند. با این کار، آن‌ها با فراهم آوردن نوعی "روانی موقعیتی" که به ما امکان می‌دهد تصمیمات آگاهانه‌تری بگیریم و خود را به جایی که می‌خواهیم برویم، عاملیت فردی مردم را در سراسر جهان تقویت می‌کنند. این امر به ویژه با توجه به گرایش‌های مرکزگرا در

هوش مصنوعی که برای دستیابی به عملکرد پیشرفته به داده‌های گسترده، سخت‌افزار، انرژی و استعداد انسانی نیاز دارد، حائز اهمیت است. برای ادامه توسعه هوش مصنوعی در راستای آرمان‌های دموکراتیک خودتعیین‌گری و مشارکت گسترده، باید هوش مصنوعی را به گونه‌ای طراحی و پیاده‌سازی کنیم که عاملیت فردی را در اولویت قرار دهد و به مردم امکان دسترسی عملی به ابزارهایی را بدهد که می‌توانند به روش‌های کاربردی و پایان‌باز از آن‌ها استفاده کنند. مفهوم‌سازی LLMs به عنوان شکلی از جی‌پی‌اس اطلاعاتی، مدلی آشنا را ارائه می‌دهد.

با این حال، در کنار شباهت‌ها، تفاوت‌های روشنی نیز بین GPS و LLMs وجود دارد. در مورد اولی، ارتش ایالات متحده کنترل انحصاری بر توسعه فناوری هسته‌ای را اعمال کرد. اما LLMs محصول آرایه‌ای جهانی از محققان آکادمیک، شرکت‌های بزرگ، حامیان متن‌باز و استارت‌آپ‌ها هستند و به صورت متن‌باز، اختصاصی یا مدل‌های نیمه‌باز که دسترسی محدود توسعه‌دهنده را از طریق نقاط دسترسی کنترل شده معروف به رابط‌های برنامه‌نویسی کاربردی (API) ارائه می‌دهند، در دسترس هستند. تفاوت بنیادی‌تر این است که GPS عمدتاً با داده‌های فضایی و زمانی عینی و واقعیت‌محور سروکار دارد – یعنی مختصات جغرافیایی و زمان‌بندی دقیق. در مقابل، LLMsخروجی‌هایی را بر اساس ظرافت‌ها، پیچیدگی‌ها و ذهنی‌گرایی زبان انسانی ایجاد می‌کنند. در حالی که LLMs می‌توانند اطلاعات واقعی را پردازش کنند، یک مجموعه واحد و فراگیر از اطلاعات عینی و "حقیقت" برای استفاده آن‌ها وجود ندارد، به ویژه هنگامی که با جنبه‌های گسترده‌تر قابل تفسیر دانش انسانی سروکار دارند.

در عوض، هر توسعه‌دهنده LLM عملاً یک "سیاره اطلاعاتی" منحصر به فرد خود را ایجاد می‌کند، همراه با یک نقشه منحصر به فرد از آن سیاره. به عبارت دیگر، هر چقدر هم مجموعه داده‌های آموزشی بزرگ باشند، هرگز حاوی تمام اطلاعات ممکن نخواهند بود. علاوه بر این، هر توسعه‌دهنده به طور موثر سیاراتی را که ایجاد می‌کند به روشی منحصر به فرد نقشه‌برداری می‌کند. دو توسعه‌دهنده متفاوت ممکن است با دقیقاً یک مجموعه داده شروع کنند، و با این

حال، به دلیل استفاده از تعداد پارامترها و وزن‌های مختلف، الگوریتم‌های متفاوت و تکنیک‌های تنظیم دقیق مختلف که برای همسوسازی مدل با ترجیحات و ارزش‌های خاص (که از توسعه‌دهنده‌ای به توسعه‌دهنده دیگر متفاوت خواهد بود) طراحی شده‌اند، سیارات اطلاعاتی به روشی تأکیدی ساختارهای انسانی هستند که سیارات فیزیکی این‌گونه نیستند. در حالی که زمین تا حد زیادی ثابت می‌ماند، این سیارات اطلاعاتی می‌توانند در طول زمان، از طریق به‌روزرسانی در تکنیک‌های آموزشی و گنجاندن داده‌های جدید، تغییر کنند. مدل‌های کسب‌وکار نیز می‌توانند بالقوه تأثیرگذار باشند. به عنوان مثال، تصور کنید چگونه توسعه‌دهندگان ممکن است داده‌های آموزشی را فیلتر کرده یا الگوریتم‌های بهینه‌سازی را دستکاری کنند تا محتوایی را که باعث تعامل و درآمد تبلیغاتی می‌شود، به قیمت دقت و نمایندگی، بیش از حد اولویت‌بندی کنند. در نهایت، از آنجا که LLMs خروجی‌ها را بر اساس احتمال آماری و نه قوانین ثابت تولید می‌کنند، یک فرمان واحد – یا "درخواست مسیر"، برای استفاده از قیاس – GPS می‌تواند هر بار که آن را وارد می‌کنید، نتایج متفاوتی تولید کند.

بخش سوم: نقش مدل‌های زبانی بزرگ در تقویت توانمندی فردی و دسترسی به دانش

با وجود تفاوت‌ها، GPS همچنان یک قیاس مفید برای درک نقش هوش مصنوعی در زندگی ما باقی می‌ماند. ما نه تنها دائماً در حال ناوبری در جهان فیزیکی هستیم، بلکه در یک جهان اطلاعاتی پیچیده نیز در حال حرکتیم. این جهان شامل باورها، ارزش‌ها و سنت‌هایی است که به هویت ما به عنوان عضوی از یک جامعه یا گروه خاص کمک می‌کنند؛ دانش و واژگان خاصی که در خط کاری خود استفاده می‌کنیم؛ قوانین، مقررات و هنجارهایی که زندگی ما را به عنوان یک شهروند تعریف می‌کنند؛ و سوادهای اضافی که در سراسر فرهنگ گسترش یافته‌اند. در هر یک از این قلمروهای شناختی مختلف، چگونه می‌توانیم بین واقعیت و عقیده تمایز قائل شویم؟ چه حکمت مرسوم و چه روایت‌های متضادی آن‌ها را شکل می‌دهند؟ چه دانش مشترک و چه

شیوه‌های گفتمانی را باید تسلط یابیم تا در هر مکان متمایز به عنوان یک "محلی" واجد شرایط شویم؟ زندگی به عنوان یک انسان در امروز به معنای ارتقای مهارت مداوم است – در محل کار، بله، اما در همه جای دیگر نیز. در حالی که نوآوری‌های دیجیتالی مداوم این پویایی را تداوم می‌بخشند، به ما در مدیریت آن نیز کمک می‌کنند. برای همگام شدن با خواسته‌های اطلاعاتی جدید، قرن بیستم محصولات و خدماتی مانند ایمیل، ابرپیوندها، جستجو و ایموجی‌ها را به ارمغان آورد. قرن بیست‌ویکم هوش مصنوعی را به ما داده است.

همانطور که از نامشان پیداست، مدل‌های زبانی بزرگ در هسته خود، سیستم‌هایی برای تجزیه و تحلیل، سنتز و نقشه‌برداری جریان‌های زبانی هستند. این همان چیزی است که قیاس با سیستم‌های ناوبری GPS را توجیه می‌کند LLMs – نقشه‌هایی بی نهایت قابل اجرا و توسعه‌پذیر هستند که به شما کمک می‌کنند با اطمینان و کارایی بیشتر از نقطه A به نقطه B برسید. شما می‌توانید از یک LLM بخواهید یک مقاله فنی در مورد یادگیری تقویتی را به اصطلاحاتی ترجمه کند که به فردی بدون پیش‌زمینه علوم کامپیوتر کمک کند تا درک اولیه ای از مطالب آن پیدا کند. می‌توانید از آن بخواهید شرایط و ضوابط یک قرارداد پیشنهادی با یک مشتری بالقوه در کشوری را ارزیابی کند که هرگز در آن تجارت نکرده‌اید و هیچ دانشی از عرف بازار استاندارد آن ندارید LLMs .با ارائه روانی سریع و پشتیبانی ناوبری بر اساس تقاضا، به کلیدهایی برای مشارکت و شایستگی بالاتر در جهانی که از اطلاعات ساخته شده است، تبدیل می‌شوند. به این ترتیب، آن‌ها نیرویی دموکراتیزه کننده هستند که روندهای مشاهده شده در موارد قبلی نوآوری‌های تکنولوژیکی را منعکس می‌کنند. به عنوان مثال، به لطف وبلاگ‌ها و رسانه‌های اجتماعی، میلیون‌ها نفر اکنون به طور منظم وظایفی را انجام می‌دهند که زمانی تقریباً در انحصار روزنامه‌نگاران حرفه‌ای بود. خدمات حمل‌ونقل مشترک مانند اوبر و لیفت نیز به طور مشابه صنعتی را که از طریق موانع گران‌قیمت ورود به صورت استراتژیک محدود شده بود، به صنعتی تبدیل کرده‌اند که میلیون‌ها نفر می‌توانند به راحتی در آن مشارکت کنند. با هوش مصنوعی،

افرادی که زمانی فاقد مهارت، آموزش یا منابع برای تولید انواع مختلف مواد نوشتاری، تصاویر، کد کامپیوتری، موسیقی و ویدئو بودند، اکنون می‌توانند این کار را انجام دهند. با تکامل مدل‌ها، آن‌ها همچنین در تقلید از کار وکلا، معلمان خصوصی، درمانگران و سایر متخصصان ماهرتر می‌شوند.

بسیاری از مردم بر این باورند که مدل‌های زبانی بزرگ، با توجه به ظرفیت آن‌ها برای پردازش و سنتز حجم عظیمی از داده‌ها، قادرند به صورت چشمگیری مهارت‌های افراد مبتدی را افزایش دهند. این امر به این دلیل است که LLMs بر روی مقدار زیادی داده آموزش دیده‌اند و عموماً با آنچه می‌توان آن را استانداردهای صلاحیت در طیف گسترده‌ای از زمینه‌ها و حوزه‌های دانش توصیف کرد، آشنا هستند. این مفهوم با مطالعاتی که در نیمه اول سال ۲۰۲۳ منتشر شد، تأیید می‌شود که نشان داد ادغام هوش مصنوعی در جریان‌های کاری، بهره‌وری را در طیف وسیعی از صنایع افزایش می‌دهد. از جمله نکات قابل توجه، بسیاری از این مطالعات نشان دادند که بالاترین دستاوردها در میان شرکت‌کنندگان کم‌تجربه‌تر بوده است. برای مثال، در مارس ۲۰۲۳، محققان MIT از ۴۴۴ متخصص دانشگاهی در حوزه‌های بازاریابی، نویسندگی، تحلیل داده و سایر مشاغل درخواست کردند تا وظایف نوشتاری کوتاهی مانند تهیه بیانیه‌های مطبوعاتی یا آماده‌سازی تحلیل را انجام دهند. طبق یافته‌های آن‌ها، کارگرانی که از ChatGPT استفاده می‌کردند، وظایف را ۳۷ درصد سریع‌تر از کسانی که از آن استفاده نمی‌کردند، تکمیل کردند. در آوریل همان سال، مطالعه دیگری نشان داد که نمایندگان خدمات مشتری در یک شرکت ناشناس، بهره‌وری خود را تا ۱۴ درصد هنگام استفاده از ChatGPT به عنوان دستیار در زمان واقعی تعاملات مشتری افزایش دادند. در مطالعه مربوط به متخصصان تحصیل‌کرده‌ای که وظایف نوشتاری انجام می‌دادند، "افزایش کیفیت برای شرکت‌کنندگانی که در اولین وظیفه خود نمره پایینی گرفته بودند" و بدون کمک ChatGPT آن را کامل کرده بودند، بزرگ‌تر بود. در مطالعه خدمات مشتری، کارکنانی که بیشترین سود را بردند، "نمایندگان خدمات مشتری کم‌تجربه‌تر و

کم‌مهارت‌تر" بودند. این امر حسی نیز منطقی است. اگر تمام عمر خود را در شهری مانند توکیو زندگی کرده باشید، ناوبری GPS احتمالاً در برخی موارد مفید خواهد بود. اما احتمالاً آن را به اندازه یک بازدیدکننده برای اولین بار، به ویژه اگر آن بازدیدکننده به زبان ژاپنی تسلط نداشته باشد یا تجربه زیادی در ناوبری در شهرهای بزرگ نداشته باشد، مفید نخواهید یافت. همین پویایی برای LLMs نیز صادق است LLMs .می‌توانند به سرعت به مبتدیان کمک کنند تا مهارت‌های خود را ارتقا دهند. رابرت سیمنز، اقتصاددان، این پدیده را "تأثیری دموکراتیزه کننده" توصیف کرده و بیان داشته که "کارگران کم‌تجربه‌تر کسانی هستند که بیشترین سود را از آن خواهند برد".

به دلیل مقیاس‌پذیری و سازگاری ذاتی آن، هوش ماشینی تأثیر دموکراتیزه کننده دیگری نیز دارد. در حالی که دانش به طور گسترده در قالب کتاب، ویدئو و سایر اشکال رسانه قابل دسترسی است، هوش انسانی تمایل دارد که با توجه به چگونگی محصور شدنش در مغزهای انسانی، متراکم‌تر باشد. تعداد روانپزشکان به ازای هر نفر در شهرهایی مانند نیویورک و سیاتل بسیار بیشتر از شهرستان‌های روستایی است. کارشناسان حقوقی در شرکت‌های حقوقی گران‌قیمت جمع می‌شوند؛ دکتراهای علوم کامپیوتر با تخصص در یادگیری ماشینی به سمت پردیس‌های شرکت‌های بزرگ فناوری جذب می‌شوند. در حالی که LLMs به همان شکلی که انسان‌ها هوش دارند، دارای هوش نیستند، تعامل با آن‌ها تجربه‌ای بسیار متفاوت از تعامل با یک صفحه ویکی‌پدیا یا پادکست ارائه می‌دهد. این امر با کسب قابلیت‌های چندوجهی بیشتر توسط LLMs آشکارتر خواهد شد. با تکامل فناوری‌ها، دریافت اطلاعاتی که به شدت با نیازهای شما شخصی‌سازی شده‌اند، در هر قالب رسانه‌ای که ترجیح می‌دهید، به طور فزاینده‌ای ممکن می‌شود. در نهایت، افراد در تمام اقشار جامعه از این فناوری بهره‌مند خواهند شد، درست همانطور که از GPS و تلفن‌های هوشمند بهره می‌برند. اما با انتشار هوش مصنوعی در جامعه، احتمالاً تأثیر تحول‌آفرینی ویژه‌ای بر کسانی خواهد داشت که دسترسی به مکان‌هایی را ندارند که

هوش انسانی به طور سنتی در آنجا متمرکز است.

به عنوان مثال، اگر شما یک دانش‌آموز دبیرستانی از یک خانواده با درآمد پایین یا متوسط هستید که فرآیند درخواست کالج را با دسترسی محدود به معلمان خصوصی، مربیان و مشاوران گران‌قیمت طی می‌کنید، LLMsممکن است زمینه و راهنمایی‌هایی را ارائه دهند که شانس پذیرش شما را افزایش دهد. اگر نامه‌ای از صاحب‌خانه‌تان دریافت می‌کنید که می‌گوید باید ظرف سی روز "مکان را ترک کنید" مگر اینکه ۳۰۰۰ دلار "اجاره معوقه" را پرداخت کنید، یک LLM می‌تواند پیشنهاداتی در مورد مراحل بعدی بالقوه شما ارائه دهد. اگر شما یک غیربومی هستید که با زبان انگلیسی دست و پنجه نرم می‌کنید، چه رسد به اصطلاحات حقوقی ناآشنا مانند "اجاره معوقه"، می‌توانید از دوربین تلفن خود برای به اشتراک گذاشتن نامه با یک LLM استفاده کنید، که سپس می‌تواند به زبان خودتان معنی نامه و گزینه‌های شما را توضیح دهد. در یک سناریوی آینده‌نگرانه، یک دستیار هوش مصنوعی قابل اعتماد حتی ممکن است داوطلب شود تا با خدمات کمک حقوقی محلی از طرف شما تماس بگیرد تا ببیند چه منابعی ممکن است در دسترس باشد. قابلیت یک LLM برای "ترجمه" از یک رسانه به رسانه دیگر، کاربردهای مشابه بسیاری دارد. یک دانش‌آموز مبتلا به نارساخوانی می‌تواند از یک LLM چندوجهی برای تبدیل درس‌های متنی و سایر مواد آموزشی به انواع مختلف فرمت‌های صوتی استفاده کند LLMs . چندوجهی همچنین قابلیت‌های موجود در خدمات فعلی را که به افراد دارای اختلالات بینایی کمک می‌کنند تا جهان اطراف خود را تفسیر کنند، بیشتر پیش خواهند برد. چنین کاربری می‌تواند اقلام را با تلفن خود هنگام خرید اسکن کند و یک مکالمه دقیق با LLM در مورد محصولات مختلف داشته باشد و در مورد قیمت‌ها، انواع هشدارها یا سلب مسئولیت‌های روی برچسب، اینکه آیا یک محصول رقیب در تخفیف است یا خیر، و غیره بپرسد.

یک فرد ناشنوا می‌تواند از LLMs چندوجهی به روش‌های مشابهی استفاده کند. هوش مصنوعی می‌تواند رونویسی‌های بسیار دقیق و در زمان واقعی از آنچه شخص یا اشخاص دیگر به

آن‌ها می‌گویند، همراه با نشانه‌های متنی در مورد لحن صدای گوینده ("او به تمسخر گفت")، شناسه‌های گوینده (اگر بیش از یک نفر در مکالمه شرکت داشته باشد) و موارد دیگر ارائه دهد. به همین ترتیب، LLMمی‌تواند زبان اشاره فرد ناشنوا را به پاسخ‌های صوتی برای استفاده افراد شنوا که زبان اشاره را نمی‌دانند، ترجمه کند LLMs .همچنین می‌توانند قابلیت‌های تشخیص صدای پیچیده‌تر و آگاه از زمینه را ارائه دهند و برنامه‌های موجود را که کاربران کم‌شنوا را در مورد صداهای مهم محیطی مانند آژیرها یا وسایل نقلیه در حال نزدیک شدن مطلع می‌کنند، بهبود بخشند. حتی بیشتر از جستجو یا ویکی‌پدیا، LLMsمی‌توانند نقاط شروع واضح و آسان برای جمع‌آوری اطلاعات را فراهم کنند. به جای تایپ پرس‌وجوها در گوگل و سپس تلاش برای ارزیابی اینکه کدام لینک‌ها واقعاً مفید هستند، می‌توانید به سادگی با یک راهنمای فوری پاسخگو و آگاه، مکالمه را شروع کنید. یک ویژگی نصب شده بر روی تلفن‌های هوشمند گوگل پیکسل ۹، به نام Screenshots، راه دیگری را نشان می‌دهد که هوش مصنوعی چندوجهی برای همه، به ویژه سالمندان، مفید خواهد بود. اساساً، حافظه تصویری را جهانی می‌کند. هنگامی که از اطلاعاتی که می‌خواهید ردیابی کنید – یک پیام متنی از یک دوست، زمانی که در بازی Wordle با یک حرکت برنده شدید، آدرس یک آزمایشگاه پزشکی که باید چند ماه دیگر به آن بازگردید – اسکرین‌شات می‌گیرید، هوش مصنوعی گوگل تصویر را به عنوان یک رکورد متنی پردازش می‌کند. سپس، هر زمان که بخواهید به آن بازگردید، می‌توانید آن را فوراً با گفتن "آدرس آزمایشگاه پزشکی" یا "آن برد سریع Wordle که قبلاً داشتم" بازیابی کنید.

هنگامی که راه حل‌های فناوری به عنوان راه‌هایی برای دسترسی بیشتر به آموزش، مراقبت‌های بهداشتی و سایر خدمات ضروری برای جوامع محروم ارائه می‌شوند، اغلب به عنوان راه‌حل‌های موقتی برای نابرابری‌های ساختاری عمیق‌تر رد می‌شوند. این خط فکری می‌گوید: "ثروتمندان پزشکان خصوصی، مربیان شخصی و درمانگران ۴۰۰ دلاری در ساعت خود را دارند. بقیه یک برنامه 'سلامت' رایگان از بیمه و یک نسخه فرسوده از 'قوانین ورشکستگی شخصی برای

مبتدیان' دریافت می‌کنند!" به عبارت دیگر، فرض بر این است که خدمات پیچیده و با ارزش بالا مانند مشاوره حقوقی و مراقبت‌های بهداشتی هرگز نمی‌توانند به طور موثر خودکار و مقیاس‌پذیر شوند، زیرا نتایج هرگز به کیفیت بالای متخصصان انسانی نخواهد رسید. در نتیجه، چنین تلاش‌هایی اغلب به عنوان نوع‌دوستی گمراه‌کننده، کلاهبرداری‌های صنعت فناوری یا صرفاً نمایش‌های راه‌حل‌گرایی بدون هیچ قصد واقعی برای ارائه ارزش به گیرندگان، نادیده گرفته می‌شوند. اما قابلیت‌های رو به رشد LLMs شروع به به چالش کشیدن این روایت کرده‌اند. در حالی که باور بر این است که متخصصان انسانی که در کنار LLMs در حوزه‌هایی مانند آموزش، مراقبت‌های بهداشتی و حقوق کار می‌کنند، سناریوی ایده‌آل را در بیشتر موارد نشان می‌دهند، LLMsبه تنهایی، که مستقیماً با یک کاربر کار می‌کنند، می‌توانند ارزش زیادی ارائه دهند. آن‌ها بدون خطا نیستند، اما متخصصان انسانی نیز بدون خطا نیستند. برخلاف همتایان انسانی خود، LLMsفوراً در دسترس، بی‌نهایت صبور و همیشه مایل به پاسخ دادن به یک سوال دیگر هستند. اگر آن‌ها را در ساعت ۳ بعدازظهر رها کنید و در ساعت ۱ صبح برگردید، هنوز دقیقاً آنچه را که بحث می‌کردید، با جزئیات به یاد خواهند آورد. هر چند بار که زندگی ایجاب می‌کند آن‌ها را رها کنید: هرگز کینه‌ای به دل نمی‌گیرند. برای یک مدیر اجرایی پرمشغله، این یک مزیت خوب است. برای یک فرد کم‌درآمد و کم‌توان که ممکن است گزینه‌های حمل‌ونقل محدود، برنامه کاری انعطاف‌ناپذیر و گزینه‌های مراقبت از کودک کمیاب داشته باشد، این نوع دسترسی می‌تواند زندگی را تغییر دهد. به عبارت دیگر، ما وارد عصری می‌شویم که هوش دیگر یک منبع محدود نخواهد بود. خودکار و شبکه‌ای، هوش مصنوعی در هر کجا که مردم هستند به آن‌ها خواهد رسید و به بخشی همه‌جا حاضر و ضروری از زندگی روزمره میلیاردها نفر تبدیل خواهد شد، همانطور که GPSاکنون هست. همانطور که ما به طور غریزی برای یافتن مسیر در قلمروهای ناآشنا به تلفن‌های هوشمند خود روی می‌آوریم، یا برای همگام‌سازی دقیق زمان در تراکنش‌های مالی به GPSمتکی هستیم، از هوش مصنوعی به گونه‌ای استفاده خواهیم کرد که اساساً حل مسئله و

تصمیم‌گیری را در هر دو حوزه شخصی و حرفه‌ای بازتعریف کند. همه از این تغییر بهره‌مند خواهند شد، اما برای افرادی که توانایی آن‌ها برای ناوبری موثر در مناظر اطلاعاتی به دلایل مختلف محدود شده است، قابلیت‌ها و آزادی‌های جدیدی که هوش مصنوعی به زندگی آن‌ها می‌آورد، به‌ویژه عمیق خواهد بود.

بخش چهارم: هوش مصنوعی عامل‌گرا و تعامل مبتنی بر گفتگو

همانطور که پیشتر نیز اشاره شد، توسعه‌دهندگان در حال ساخت سیستم‌های مبتنی بر LLMهستند که قابلیت عملگری بیشتری دارند. به عنوان مثال، GPT-4با Code Interpreter شرکت OpenAI می‌تواند به طور خودمختار کد بنویسد، اجرا کند و اشکال‌زدایی کند تا حتی وظایف برنامه‌نویسی پیچیده را بدون نیاز به دخالت گام‌به‌گام کاربر انجام دهد .AutoGPT یک برنامه متن‌باز است که از GPT-4 برای اجرای مجموعه‌ای از وظایف طراحی شده برای دستیابی به یک هدف خاص استفاده می‌کند، مانند ایجاد خلاصه‌ای تکرارشونده از مقالات تحقیقاتی جدید در یک موضوع خاص که هر هفته در arXiv.org منتشر می‌شوند. اگرچه آن‌ها معمولاً در محدودیت‌های دقیق تعریف‌شده‌ای عمل می‌کنند و امکان دخالت و نظارت انسانی را فراهم می‌آورند، سیستم‌های هوش مصنوعی عامل‌گرا فراتر از پاسخ‌گویی ساده به سوالات هستند که ویژگی اصلی چت‌بات‌های سنتی است. هدف این است که از LLMs به گونه‌ای استفاده شود که امکان برنامه‌ریزی بلندمدت و بهبود خودکار از طریق یادگیری را فراهم کند. به این ترتیب، یک سیستم عامل‌گرا می‌تواند در جریان‌های کاری شبیه‌تر به انسان درگیر شود و بیشتر شبیه یک دستیار یا همکار عمل کند.

تصور کنید یک هوش مصنوعی می‌تواند به شما در شناسایی اینفلوئنسرهای نوظهور که برای تبلیغ اپلیکیشن تناسب اندام جدید شما مناسب هستند، کمک کند. این سیستم ابتدا مجموعه‌ای اولیه از گام‌ها را برای پیگیری ترسیم می‌کند، سپس ممکن است کدی برای خودکارسازی جمع‌آوری داده‌ها از پلتفرم‌های رسانه‌های اجتماعی بنویسد یا تحلیل احساسات را

بر روی نظرات و بررسی‌های ارسال شده توسط تعدادی از اینفلوئنسرها انجام دهد. همانطور که اطلاعات را ارزیابی می‌کند، عامل هوش مصنوعی شما ممکن است تاکتیک‌های خود را در صورتی که اینفلوئنسرهایی با نرخ تعامل به اندازه کافی بالا در حوزه تناسب اندام پیدا نکند، اصلاح کند. در برخی سیستم‌ها، چندین عامل می‌توانند برای کار بر روی جنبه‌های خاص پروژه، دقیقاً مانند یک تیم از همکاران انسانی که به طور مشترک برای دستیابی به یک هدف مشترک کار می‌کنند، به کار گرفته شوند. بسیاری از مردم بر این باورند که LLMs عامل‌گرا به طور کامل از LLMs" که فقط با شما صحبت می‌کنند" پیشی خواهند گرفت. این به این دلیل است که سیستم‌های عامل‌گرا هوش مصنوعی را در مقیاس وسیعی امکان‌پذیر می‌سازند که هوش مصنوعی مکالمه‌ای سنتی نمی‌تواند LLMs .عامل‌گرا به لحاظ نظری می‌توانند به طور همزمان روی چندین وظیفه پیچیده با حداقل نظارت انسانی کار کنند. بنابراین، تفاوت بین انجام ده چیز، یا صد، یا هزار چیز در یک زمان، در مقایسه با انجام فقط یک کار در یک زمان است. این توانایی برای چند برابر کردن بهره‌وری انسانی و پرداختن به وظایف با ارزش بالا به صورت موازی، LLMsعامل‌گرا را به ویژه برای کسب‌وکارها و افرادی که به دنبال به حداکثر رساندن تأثیر خود هستند، جذاب می‌کند.

با این حال، حتی با افزایش محبوبیت اشکال عامل‌گراتر هوش مصنوعی، باور بر این است که درگیر شدن در گفتگوی چندنوبته با LLMs بخش کلیدی به حداکثر رساندن ارزش آن‌ها باقی خواهد ماند. همانطور که LLMs و سیستم‌های ساخته شده بر روی آن‌ها به اندازه کافی توانمند می‌شوند تا به طور خودمختار به روش‌های بسیار قابل اعتماد و تطبیق‌پذیر عمل کنند، این پیشرفت‌ها همچنین آن‌ها را در گوش دادن، تعامل و پیروی از دستورالعمل‌ها در مکالمات یک‌به‌یک مداوم با آن‌ها بهتر خواهد کرد. همانطور که این اتفاق می‌افتد، شما قادر خواهید بود با کنترل آسان‌تر و دقیق‌تر، کارهای بیشتری انجام دهید و می‌خواهید تا آنجا که می‌توانید از این سیستم‌ها بهره ببرید. بخش عمده‌ای از آنچه ChatGPT را در ابتدا بسیار تحول‌آفرین کرد، شامل نحوه نیاز صریح آن به کاربران برای ایفای نقش مداوم در فرآیند استفاده از هوش مصنوعی بود.

در این حالت تعامل، یک LLM تا زمانی که شما یک فرمان وارد نکنید، هیچ کاری انجام نمی‌دهد. پس از اینکه LLM به فرمان شما پاسخ داد، دوباره هیچ کاری انجام نمی‌دهد، مگر اینکه شما آن را بیشتر تحریک کنید. این تعامل نوبتی با LLMs نیازمند سطح بالایی از مشارکت است – و این آشکارا به کاهش برخی از نقاط ضعف آن‌ها کمک می‌کند. با بررسی هر خروجی که از یک LLM دریافت می‌کنید، احتمال بیشتری دارد که متوجه شوید آیا آن "توهم‌زایی" می‌کند یا نوع دیگری از اطلاعات نادرست را تولید می‌کند. به طور مشابه، اگر خروجی‌های جانبدارانه، سمی یا نامطلوب دیگری تولید کند، می‌توانید آن را به چالش بکشید یا اصلاح کنید.

این **قابلیت بازخورد فوری کاربر**، LLMsعملی مانند ChatGPT را بلافاصله از اکثر اشکال اولیه هوش مصنوعی متمایز کرد. سوگیری‌های ساختاری در مدل‌های هوش مصنوعی یک مسئله مهم است و این مشکل با این واقعیت تشدید می‌شود که افراد بیشتر تحت تأثیر چنین سوگیری‌هایی عموماً راهی برای تعامل با مدل و خنثی کردن آسیب نداشته‌اند. به عنوان مثال، در مورد تصمیم‌گیری الگوریتمی در تأیید وام، فردی که تحت تأثیر یک الگوریتم تبعیض‌آمیز قرار گرفته است، ممکن است حتی هیچ ایده‌ای نداشته باشد که تصمیم آن مبنای رد شدن وی بوده است. با هوش مصنوعی مکالمه‌ای، اوضاع متفاوت است. برخی از LLMs حتی به شما امکان می‌دهند قبل از صدور فرمان، آن‌ها را آماده کنید، با ایجاد دستورالعمل‌های سفارشی که ارزش‌ها، مقاصد و انواع پاسخ‌هایی را که به دنبال آن هستید، منتقل می‌کنند. این قابلیت، اگرچه بی‌عیب و نقص نیست، اما ابزار قدرتمندی برای کاربران فراهم می‌آورد .**قابلیت مداخله، تصحیح و تنظیم دقیق خروجی‌های LLM در زمان واقعی** به آن‌ها این ظرفیت را می‌دهد که شما را در جایی که هستید ملاقات کنند – به روشی که کتاب‌ها، ویدئو و سایر اشکال دانش میانجی‌گری‌شده به سادگی نمی‌توانند. اگر از ChatGPT بخواهید "تصویری از یک مهماندار هواپیما که به گروهی از وکلا ناهار می‌دهد" ایجاد کند، احتمالاً مهماندار را زن و وکلا را عمدتاً مرد به تصویر می‌کشد. اما اگر تصویر متفاوتی می‌خواهید، می‌توانید به سمت نتیجه دلخواه خود هدایت کنید. ممکن است

بگویید: "تصویری از یک مهماندار هواپیمای مرد و یک مهماندار هواپیمای زن که به گروهی از وکلا با ترکیب جنسیتی متعادل و تنوع قومی ناهار می‌دهند".

این قابلیت به طور معجزه‌آسا مسئله اساسی سوگیری ساختاری در داده‌های آموزشی مدل را نفی نمی‌کند. ساخت مدل‌ها نماینده‌تر و فراگیرتر یک فرآیند مداوم است. اما، بیشتر از سایر انواع رسانه و سایر انواع هوش مصنوعی، LLMsمکالمه‌ای فرصت‌های بی‌سابقه‌ای را برای کاربران برای اعمال نفوذ خود می‌دهند. شما می‌توانید هوش مصنوعی را راهنمایی، سوال، اطلاع‌رسانی و عقب‌نشینی کنید و بنابراین محتوا و جریان یک بحث را تغییر مسیر دهید .**توانایی شما برای مداخله به روش‌های ثابت و قدرتمند، همراه با دانش فرهنگی تقریباً کیهانی که LLMs از نقشه‌برداری کهکشان‌های داده به روش‌های وسواس‌گونه دقیق کسب می‌کنند، به شما امکان می‌دهد خروجی‌های آن‌ها را مانند یک مهندس استودیوی ضبط که با فیدرها، دستگیره‌ها و سوئیچ‌ها در یک کنسول میکس ۱۲۸ کاناله کار می‌کند، دستکاری و تنظیم کنید .** این ظرفیت ساختاری برای **ویژگی‌بخشی** می‌تواند برای کاهش سوگیری – و برای بی‌شمار نیازها و اهداف دیگر – اعمال شود. هر زبان عاطفی، فرهنگی، حرفه‌ای، ایدئولوژیک یا فلسفی که داشته باشید، یک LLM احتمالاً می‌تواند آن را صحبت کند (البته با درجات مختلف روانی). با برخی ملاحظات، البته. اکثر توسعه‌دهندگان از تکنیک‌های مختلفی در طول تنظیم دقیق مدل برای ایجاد محافظ‌های ایمنی و هنجارهای سیاستی استفاده می‌کنند که ممکن است پاسخ‌ها به انواع خاصی از دستورات را که اصول اصلی و دستورالعمل‌های اخلاقی تعیین شده توسط توسعه‌دهنده را نقض می‌کنند، محدود کنند. اما این محدودیت‌های متوسط، آزادی عمل عظیمی را برای کاربران برای شکل دادن به پاسخ‌های LLM به اهداف و ترجیحات منحصر به فرد خود باقی می‌گذارد.

نقشه گوگل را در نظر بگیرید: زوم کردن به داخل و خارج یک مکان خاص فقط شروع کار است. شما همچنین می‌توانید آن را از طریق لایه‌های مختلف مشاهده کنید – لایه‌ای که توپوگرافی را

نشان می‌دهد، لایه‌ای که ترافیک زنده را نشان می‌دهد، لایه‌ای که کیفیت هوا را نشان می‌دهد، لایه‌ای که مسیرهای دوچرخه‌سواری را نشان می‌دهد. همانطور که در لایه‌ها می‌چرخید، زوم می‌کنید و به نمای خیابان تغییر می‌دهید، زوایای و جزئیات جدیدی، جنبه‌های جدیدی از چشم‌انداز را می‌بینید. به طور مشابه، گفتگوی نوبتی با LLMs مکالمه‌ای می‌تواند به شما کمک کند تا **زمینه‌های واقعی، عاطفی، سیاسی، فرهنگی، اقتصادی و تاریخی** موضوعات و موقعیت‌ها را به هر روشی که برای نیازها و مقاصد فعلی شما مرتبط‌تر است، "نقشه‌برداری" کنید. یک راه ساده برای نشان دادن این امر، امتحان یک سری از دستورات مانند: "نظریه نسبیت را برای یک کودک شش ساله توضیح دهید." "نظریه نسبیت را برای یک دانش‌آموز فیزیک دبیرستان توضیح دهید." "نظریه نسبیت را برای یک فرد عادی بزرگسال توضیح دهید." سپس از LLM بخواهید این درخواست را با دستورالعمل‌های جزئی‌تر پردازش کند تا پاسخ را برای هر تعداد بی‌نهایتی از کودکان شش ساله تنظیم کند: "نظریه نسبیت را برای یک کودک شش ساله که عاشق ماشین آتش‌نشانی است توضیح دهید." "نظریه نسبیت را برای یک کودک شش ساله که زبان اصلی او اسپانیایی است اما در حال یادگیری انگلیسی نیز هست توضیح دهید." "نظریه نسبیت را برای یک کودک شش ساله که مجذوب فضای بیرونی و سیارات است و یادگیری از طریق داستان‌سرایی را ترجیح می‌دهد توضیح دهید".

در این راستا، LLMs هم سرعت می‌بخشند و هم به **روانی در یک حوزه خاص پاداش می‌دهند** .هرچه دانش و ظرافت بیشتری بتوانید به کار ببرید، بهتر می‌توانید از آنچه ایتان مولیک، استادیار دانشکده وارتون که در کارآفرینی و نوآوری تخصص دارد، به عنوان "تخصص پنهان LLMs (latent expertise) "توصیف کرده است، استفاده کنید – یعنی تمام دانشی که آن‌ها به طور ضمنی از طریق آموزش خود جذب کرده‌اند، به روش‌هایی که نه بلافاصله آشکار و نه کاملاً قابل پیش‌بینی هستند. مولیک در مقاله‌ای می‌نویسد: "متخصصان مزایای بسیاری دارند. آن‌ها بهتر می‌توانند خطاهای LLM و توهم‌زایی‌ها را تشخیص دهند؛ آن‌ها بهتر می‌توانند

خروجی هوش مصنوعی را در حوزه مورد علاقه خود قضاوت کنند؛ آن‌ها بهتر می‌توانند به هوش مصنوعی دستور دهند که کار مورد نیاز را انجام دهد؛ و فرصت بیشتری برای آزمون و خطا دارند. این به آن‌ها امکان می‌دهد تا تخصص پنهان درون LLMs را به روش‌هایی که دیگران نمی‌توانند، کشف کنند." قدرت دستورالعمل‌های متخصص در روزهای اولیه پس از انتشار DALL-E 2 ، تولیدکننده تصویر از متن OpenAI ، در بهار ۲۰۲۲ به وضوح نمایش داده شد. عکاسان، تصویرگران و دیگرانی که به دستور زبان و واژگان ایجاد تصویر بسیار مسلط بودند، در خط مقدم کشف تخصص پنهان DALL-E از طریق ویژگی‌بخشی دستورات خود بودند. در حالی که بسیاری از کاربران اولیه برای تولید تصاویر به دستورات کلی (مانند "یک فضانورد سوار بر اسب در مریخ!") تکیه می‌کردند، متخصصان بصری به آن دستور می‌دادند تا لنزهای دوربین خاص و منابع نوری را تقلید کند و دستورالعمل‌های ترکیبی") یک فضانورد سوار بر اسب در مریخ، لنز ۵۰ میلی‌متری در f/2.8، با نور گرم غروب خورشید منعکس شده از فوبوس، با سوژه قرار گرفته خارج از مرکز ("اضافه می‌کردند. ورودی متخصص از کاربران انسانی بود که شروع به کشف دامنه کامل قابلیت‌هایی کرد که DALL-E 2 می‌توانست از آن‌ها استفاده کند.

هر کجا که دانش دامنه خاص شما قرار دارد، مفهوم "تخصص پنهان" برای شما، همانند همه کاربران LLM ، پیامدهایی دارد. در هر زمینه‌ای، ویژگی‌بخشی اهمیت دارد، زیرا ویژگی‌بخشی همان راهی است که می‌توانید به بهترین نحو از نحوه نقشه‌برداری یک LLM از سیاره اطلاعاتی خود بهره ببرید .و حتی اگر شما تازه‌وارد به حوزه‌ای هستید که سعی در کشف آن دارید، کاملاً ناآشنا با واژگان، هنجارها و ظرافت‌های آن، شما یک متخصص هستید وقتی صحبت از دانستن اینکه دقیقاً در آن لحظه چه کاری می‌خواهید انجام دهید و به کجا فکر می‌کنید می‌خواهید بروید، مطرح است. این به شما نقطه ورود لازم را برای شروع گشودن تخصص پنهان یک LLM می‌دهد. هنگام تعامل با LLMs ، مفید است که تا آنجا که می‌توانید "مختصات" ارائه دهید: چه چیزی را می‌خواهید یاد بگیرید؟ ("من می‌خواهم اصول اولیه محاسبات کوانتومی را درک کنم.") آیا

هدف یا قصد خاصی پشت این درخواست وجود دارد؟ ("من در حال آماده شدن برای مصاحبه شغلی در یک شرکت فناوری هستم.") چه جزئیاتی در مورد شما می‌تواند به LLM کمک کند تا پاسخی متناسب ارائه دهد؟ ("من پیش‌زمینه فیزیک کلاسیک دارم اما تجربه‌ای در مکانیک کوانتومی ندارم.") آیا نقش یا شخصیت خاصی وجود دارد که خود LLM باید برای این تعامل به عهده بگیرد؟ ("لطفاً این را طوری توضیح دهید که گویی یک معلم علوم دبیرستان صبور هستید.") چه عواملی ممکن است خروجی‌های آن را برای شما مرتبط‌تر کند؟ ("من از طریق قیاس با اشیاء روزمره بهتر یاد می‌گیرم.") هرچه اطلاعات بیشتری در مورد موقعیت فعلی و مقصد خود به LLM بدهید، دقیق‌تر می‌تواند به شما در ترسیم مسیر رسیدن به آنجا کمک کند.

بخش پنجم: درس‌هایی از گذشته برای آینده‌ی هوش مصنوعی

همانطور که در ابتدای این فصل اشاره شد، همه با ایده انتشار دسترسی GPS به عموم مردم در سال ۱۹۸۳ که رونالد ریگان برای اولین بار آن را پیشنهاد کرد، موافق نبودند. این سؤال مطرح بود که چرا باید دسترسی به اطلاعات دقیق موقعیت جغرافیایی را برای تروریست‌ها، کشورهای دشمن، تعقیب‌کنندگان و مجرمان آسان کنیم؟ نگرانی‌ها در مورد آنچه ممکن است اشتباه پیش رود، منجر به سیاست دسترسی انتخابی شد که عمداً کاربرد GPS را برای چندین سال محدود کرد. با این حال، در دهه‌هایی که از لغو آن سیاست می‌گذرد، شاهد انواع تهدیدات جدی برای امنیت ملی که وزارت دفاع از آن‌ها بیم داشت با دسترسی جهانی و رایگان به GPS محقق شود، نبوده‌ایم. این بدان معنا نیست که سیستم بدون نقص یا بدون ریسک است. در واقع، طیف وسیعی از حوادث و سوءاستفاده‌ها به طور منظم رخ می‌دهد. موارد نادر اما غم‌انگیز که رانندگان دستورالعمل‌های نادرست سیستم ناوبری را دنبال کرده و به مناطق دورافتاده یا به داخل آب می‌روند، با عواقب مرگبار، بیشترین توجه را به خود جلب می‌کنند. اما بیشتر افرادی که به ناوبری GPSمتکی هستند، سناریوهای مشابه اما عموماً بی‌ضرر را در زمان‌های مختلف تجربه کرده‌اند، با سیستم‌هایی که دستورالعمل‌های گیج‌کننده ارائه می‌دهند، آن‌ها را به جاده‌ها یا مسیرهای

خصوصی هدایت می‌کنند، یا به طور کلی آن‌ها را به بیراهه می‌کشانند.

همچنین مواردی وجود دارد که عوامل بدکار با ایجاد پارازیت یا جعل سیگنال‌های زمان‌بندی دقیق که GPS به آن‌ها متکی است، باعث می‌شوند گیرنده‌ها موقعیت‌های نادرست را محاسبه کنند یا به کلی از کار بیفتند. سارقان کانتینرهای کامل از پارازیت‌کننده‌ها برای غیرفعال کردن برچسب‌های ردیابی GPS روی کالاهای داخل آن‌ها استفاده می‌کنند. با این حال، در بیش از بیست سال گذشته، داستان غالبGPS ، داستانی از محدودیت‌ها، نقص‌ها و آسیب‌پذیری‌هایی که منجر به اختلال در امنیت، ناوبری و سایر خدمات ضروری می‌شود، نیست. در عوض، GPSبه طور جهانی طیف وسیعی از خدمات بسیار سودمند را هر دقیقه از هر روز، هفته به هفته، سال به سال، امکان‌پذیر ساخته است.

آیا نباید همین نوع ارزش را با هوش مصنوعی گسترده و عملی، زودتر از دیرتر دنبال کنیم؟ هنگامی که بیل کلینتون در سال ۱۹۹۶ اعلام کرد که دولت تصمیم گرفته است به دسترسی انتخابی پایان دهد، تاریخ هدف برای اجرای این تغییر ۲۰۰۶ بود. اما پس از اعلام این خبر، تقاضای عمومی برای سطح بالاتر خدمات آنقدر زیاد بود که دولت این سیاست را در سال ۲۰۰۰، شش سال زودتر از موعد مقرر، پایان داد. این لحظه‌ای بود که دولت فدرال به صراحت انتخاب کرد تا نوآوری و پذیرش فناوری را تسریع کند، و این تأثیر مثبت عظیمی داشت. ما باید این مثال را به خوبی در ذهن داشته باشیم تا بتوانیم جی‌پی‌اس اطلاعاتی را نیز دسترس‌پذیرتر کنیم. این تجربه تاریخی به ما یادآوری می‌کند که با وجود نگرانی‌های اولیه و چالش‌های ذاتی در هر فناوری جدید، رویکرد پیش‌گامانه و اعتماد به پتانسیل گسترده‌تر فناوری برای عموم مردم می‌تواند به منافع بی‌سابقه‌ای منجر شود. همانندGPS ، هوش مصنوعی نیز، اگر به درستی و با تمرکز بر تقویت عاملیت فردی توسعه و توزیع شود، می‌تواند ابزاری بی‌نظیر برای ناوبری در پیچیدگی‌های جهان معاصر و باز کردن افق‌های جدید برای همه افراد باشد.

نتیجه‌گیری

این فصل با بررسی عمیق سیر تکامل سیستم موقعیت‌یاب جهانی (GPS) و تشریح نقش متحول‌کننده آن در ناوبری فیزیکی، بستری برای درک مفهوم "جی‌پی‌اس اطلاعاتی" و کاربرد آن در حوزه هوش مصنوعی، به‌ویژه مدل‌های زبانی بزرگ(LLMs) ، فراهم آورد. ما دریافتیم که چگونهGPS ، از یک پروژه نظامی محرمانه به یک کالای عمومی جهانی تبدیل شد و با وجود نگرانی‌های اولیه، منافع اقتصادی و اجتماعی بی‌ساری را به ارمغان آورد. این تجربه تاریخی، الگویی قدرتمند برای رویکرد ما به توسعه و پذیرش هوش مصنوعی ارائه می‌دهد.

قیاس میان GPS و LLMs نشان می‌دهد که هوش مصنوعی به عنوان یک سیستم ناوبری برای جهان اطلاعاتی عمل می‌کند و با تبدیل داده‌های بزرگ به دانش بزرگ، عاملیت فردی را به طرز چشمگیری افزایش می‌دهدLLMs .، با قابلیت‌های خود در تحلیل، سنتز و نقشه‌برداری زبان، نه تنها به افراد کم‌تجربه کمک می‌کنند تا مهارت‌های خود را ارتقا دهند و به اطلاعات دسترسی پیدا کنند، بلکه نابرابری‌های ناشی از تجمع هوش انسانی در مراکز خاص را نیز کاهش می‌دهند. از کمک به دانشجویان در فرآیند درخواست کالج گرفته تا یاری رساندن به افراد دارای ناتوانی‌های حسی و زبانی، LLMsپتانسیل عظیمی برای دموکراتیزه کردن دسترسی به خدمات حیاتی و اطلاعات دارند. علاوه بر این، مفهوم هوش مصنوعی عامل‌گرا که می‌تواند وظایف پیچیده‌ای را با حداقل نظارت انجام دهد، وعده افزایش بی‌سابقه بهره‌وری انسانی را می‌دهد. با این حال، اهمیت تعامل مستمر و مبتنی بر گفتگو باLLMs ، به عنوان ابزاری برای کاهش سوگیری‌ها، افزایش دقت و شخصی‌سازی نتایج، بر توانمندی کاربر در هدایت و شکل‌دهی به خروجی‌های هوش مصنوعی تأکید می‌کند. درس‌های آموخته شده از تاریخچهGPS ، به ویژه تسریع در حذف محدودیت‌های دسترسی، نشان می‌دهد که استقبال از هوش مصنوعی گسترده و عملی می‌تواند منجر به نوآوری‌های چشمگیر و ایجاد ارزش‌های اجتماعی و اقتصادی عظیم شود. در نهایت، جی‌پی‌اس اطلاعاتی، نقشه راهی را برای عصری ترسیم می‌کند که در آن هوش،

دیگر یک منبع محدود نیست، بلکه ابزاری همه‌جا حاضر و ضروری برای توانمندسازی میلیاردها نفر در سراسر جهان خواهد بود.

نکات کلیدی

- **تکامل GPS**: GPS از یک پروژه نظامی به یک ابزار عمومی جهانی تبدیل شد که ناوبری، مخابرات، کشاورزی و واکنش اضطراری را متحول کرد.
- **عامل دسترسی انتخابی (Selective Availability)**: در ابتدا، دقت GPS برای استفاده غیرنظامی عمداً محدود شد، اما با افزایش تقاضا و کاهش نگرانی‌ها، این محدودیت برداشته شد.
- **GPS به عنوان الگوی نوآوری**: رویکرد دولت ایالات متحده در حمایت از نوآوری بخش خصوصی برای GPS، الگویی مثبت برای توسعه فناوری‌های جدید ارائه می‌دهد.
- **جی‌پی‌اس اطلاعاتی (Informational GPS)**: LLMs را می‌توان به عنوان یک سیستم GPS برای ناوبری در جهان پیچیده اطلاعاتی در نظر گرفت.
- **افزایش عاملیت فردی**: هم GPS و هم LLMs عاملیت فردی را با ارائه دانش به‌روز و راهنمایی در محیط‌های فیزیکی و اطلاعاتی افزایش می‌دهند.
- **تفاوت‌های بنیادین GPS و LLMs**: GPS با داده‌های عینی و فضایی-زمانی سروکار دارد، در حالی که LLMs با پیچیدگی‌ها و ذهنی‌گرایی زبان انسانی کار می‌کنند و "سیارات اطلاعاتی" منحصر به فردی را ایجاد می‌کنند.
- **LLMs به عنوان نیروی دموکراتیزه کننده**: آن‌ها دسترسی به مهارت‌ها و دانش را برای طیف وسیعی از افراد، به‌ویژه کم‌تجربه‌ها، تسهیل می‌کنند.
- **پتانسیل ارتقای مهارت‌ها**: مطالعات نشان می‌دهد LLMs بهره‌وری و کیفیت کار افراد کم‌تجربه را به طور قابل توجهی افزایش می‌دهند.
- **دسترسی به هوش فراتر از مکان**: LLMs می‌توانند شکاف‌های ناشی از تجمع هوش

انسانی در مناطق خاص را پر کنند.

- **کاربردهای چندوجهی LLMs:** از کمک به دانشجویان و افراد با مشکلات حقوقی گرفته تا یاری رساندن به افراد دارای ناتوانی‌های بینایی و شنیداری.
- **هوش مصنوعی عامل‌گرا** :سیستم‌هایی که قادر به برنامه‌ریزی بلندمدت، خودبهبودی و انجام وظایف پیچیده با حداقل نظارت هستند.
- **اهمیت تعامل مبتنی بر گفتگو** :تعامل نوبتی با LLMs به کاربران امکان می‌دهد سوگیری‌ها را تشخیص داده، خطاها را تصحیح و خروجی‌ها را شخصی‌سازی کنند.
- **ویژگی‌بخشی در دستورات** :ارائه دستورالعمل‌های دقیق و زمینه‌ساز به LLMs ، قفل "تخصص پنهان" آن‌ها را باز می‌کند.
- **درس‌های پذیرش GPS:** تجربه GPS نشان می‌دهد که غلبه بر ترس‌های اولیه و تسریع در دسترسی عمومی به فناوری‌های پیشرفته، منافع عظیم و جهانی به همراه دارد.
- **هوش به عنوان منبع نامحدود** :در آینده، هوش مصنوعی به عنوان یک منبع خودکار و شبکه‌ای، در دسترس همه خواهد بود و حل مسئله و تصمیم‌گیری را دگرگون خواهد کرد.

سؤالات تفکربرانگیز

۱. با توجه به تفاوت‌های بنیادین در ماهیت داده‌ها و مدل‌های GPS وLLMs ، چگونه می‌توانیم از خطرات ناشی از "توهم‌زایی" یا سوگیری‌های احتمالی در LLMs به طور موثر جلوگیری کنیم، در حالی که همچنان به دنبال بهره‌برداری از پتانسیل دموکراتیزه کننده آن‌ها هستیم؟

۲. چگونه می‌توانیم اطمینان حاصل کنیم که توسعه و پیاده‌سازی هوش مصنوعی عامل‌گرا، با افزایش توانایی آن در انجام وظایف پیچیده بدون نظارت مستقیم

انسانی، همچنان عاملیت و کنترل نهایی کاربر را حفظ کند و منجر به کاهش استقلال فردی نشود؟

۳. با توجه به سرعت بالای تکامل LLMs و قابلیت‌های چندوجهی آن‌ها، چه چالش‌های اخلاقی و اجتماعی جدیدی در رابطه با حفظ حریم خصوصی، امنیت داده‌ها و مسئولیت‌پذیری در قبال خروجی‌های هوش مصنوعی پدیدار خواهند شد؟

۴. درس‌های آموخته شده از سیاست‌های دولت در موردGPS ، از جمله دسترسی انتخابی و سپس لغو آن، چه پیامدهای مشخصی برای قانون‌گذاری و سیاست‌گذاری در حوزه هوش مصنوعی، به ویژه در زمینه دسترسی عمومی و مدیریت ریسک، دارد؟

۵. چگونه می‌توانیم "تخصص پنهان LLMs "را به گونه‌ای توسعه دهیم که نه تنها برای متخصصان بلکه برای کاربران عادی و تازه‌کار نیز قابل دسترس و مفید باشد، و از ایجاد یک شکاف جدید بین "کاربران متخصص هوش مصنوعی" و "کاربران عادی" جلوگیری کنیم؟

۳۰ جمله کلیدی

۱. تا اوایل دهه ۲۰۰۰، نقشه‌های کاغذی ابزارهایی دست‌وپاگیر و خطرناک برای ناوبری بودند.

۲. GPSدر سال ۱۹۷۳ توسط وزارت دفاع ایالات متحده آغاز شد و از ماهواره‌ها برای تعیین مختصات جغرافیایی استفاده می‌کند.

۳. پس از سرنگونی یک هواپیمای کره‌ای در سال ۱۹۸۳، رئیس‌جمهور ریگان دسترسی غیرنظامی به GPS را اعلام کرد.

۴. در سال ۱۹۸۹، مگلان اولین گیرنده دستی GPS را با قیمت بالا به بازار مصرف معرفی کرد.

۵. سیاست "دسترسی انتخابی" در دهه ۱۹۹۰ دقت GPS غیرنظامی را ده برابر کاهش

داد.

۶. بیل کلینتون در سال ۱۹۹۶ تصمیم گرفت به دسترسی انتخابی پایان دهد تا نوآوری را تسریع کند.

۷. در سال ۲۰۰۰، GPS به طور کامل برای غیرنظامیان قابل دسترس شد و قیمت گیرنده‌ها به حدود ۱۰۰ دلار کاهش یافت.

۸. الزامات FCC برای ۹۱۱، ادغام GPS در تلفن‌های همراه را اجباری کرد.

۹. GPS بین سال‌های ۱۹۸۴ تا ۲۰۱۷، ۱٫۴ تریلیون دلار منافع اقتصادی ایجاد کرد.

۱۰. GPS علاوه بر ناوبری، برای همگام‌سازی شبکه‌های مخابراتی و کشاورزی دقیق نیز استفاده می‌شود.

۱۱. جی‌پی‌اس اطلاعاتی، LLMs را به عنوان ابزاری برای ناوبری در محیط‌های پیچیده اطلاعاتی معرفی می‌کند.

۱۲. LLMs با تبدیل "داده‌های بزرگ" به "دانش بزرگ"، عاملیت فردی را تقویت می‌کنند.

۱۳. برخلاف GPS ، LLMs خروجی‌هایی را بر اساس ظرافت‌ها و ذهنی‌گرایی زبان انسانی تولید می‌کنند.

۱۴. هر توسعه‌دهنده LLM عملاً یک "سیاره اطلاعاتی" منحصر به فرد با نقشه خاص خود را ایجاد می‌کند.

۱۵. خروجی‌های LLMs بر اساس احتمال آماری است و می‌تواند برای یک فرمان واحد متفاوت باشد.

۱۶. LLMs ابزارهایی بی نهایت قابل اجرا و توسعه‌پذیر برای تجزیه و تحلیل و نقشه‌برداری جریان‌های زبانی هستند.

۱۷. آن‌ها با ارائه "روانی موقعیتی"، به مردم کمک می‌کنند تصمیمات آگاهانه‌تری بگیرند.

۱۸. LLMsنیروی دموکراتیزه کننده‌ای هستند که دسترسی به مهارت‌ها و منابع را برای همه افزایش می‌دهند.

۱۹. مطالعات نشان می‌دهد LLMs بهره‌وری و کیفیت کار افراد کم‌تجربه را به طور چشمگیری بهبود می‌بخشند.

۲۰. LLMsمی‌توانند شکاف‌های دسترسی به هوش انسانی را در مناطق محروم پر کنند.

۲۱. قابلیت‌های چندوجهی LLMs به افراد با ناتوانی‌های حسی و زبانی کمک شایانی می‌کنند.

۲۲. LLMsنسبت به موتورهای جستجو، نقاط شروع واضح‌تری برای جمع‌آوری اطلاعات فراهم می‌کنند.

۲۳. هوش مصنوعی عامل‌گرا به LLMs امکان برنامه‌ریزی بلندمدت و خودبهبودی را می‌دهد.

۲۴. هوش مصنوعی عامل‌گرا پتانسیل چند برابر کردن بهره‌وری انسانی را دارد.

۲۵. تعامل چندنوبته با LLMs برای تشخیص خطاها و سوگیری‌ها و تصحیح آن‌ها حیاتی است.

۲۶. کاربران می‌توانند با دستورالعمل‌های سفارشی، خروجی‌های LLMs را شخصی‌سازی کنند.

۲۷. ویژگی‌بخشی در دستورات، قفل "تخصص پنهان LLMs "را باز می‌کند و نتایج را دقیق‌تر می‌سازد.

۲۸. درس‌های GPS نشان می‌دهد که ترس‌های اولیه در مورد فناوری‌های جدید اغلب اغراق‌آمیز هستند.

۲۹. تسریع در پذیرش گسترده و عملی هوش مصنوعی می‌تواند تأثیر مثبت عظیمی بر

جامعه داشته باشد.

۳۰. هوش مصنوعی در آینده به یک منبع نامحدود، همه‌جا حاضر و ضروری برای میلیاردها نفر تبدیل خواهد شد.

فصل ۸

کد به مثابه قانون: حکمرانی و کنترل در عصر هوش مصنوعی

"در دنیایی که کُد، معماری واقعیت ما را می‌سازد، فهم پویایی‌های آن برای حفظ آزادی‌های اساسی و تضمین حکمرانی عادلانه حیاتی است —— ".پژوهشگر ناشناس در حوزه حقوق و فناوری

اهداف اصلی فصل:

- آشنایی با مفهوم "**قانون کد است** "از دیدگاه لارنس لسیگ و ریشه‌های آن در فضای مجازی اولیه.
- تحلیل چگونگی تکامل معماری دیجیتال و نقش کد از اینترنت اولیه تا عصر هوش مصنوعی.
- بررسی پیامدهای مفهوم "**کنترل کامل** "که توسط فناوری‌های پیشرفته امکان‌پذیر شده است.
- مقایسه و کنکاش در تفاوت‌های میان سازوکارهای سنتی اجرای قانون و حکمرانی خودکار مبتنی بر کد.
- تأکید بر اهمیت مفهوم "**رضایت حاکمیت‌شوندگان** "در زمینه سیستم‌های هوش مصنوعی نوظهور و نقش عمومی در مشروعیت‌بخشی به آنها.

چکیده

این فصل به بررسی عمیق مفهوم محوری **"قانون کد است "**می‌پردازد که توسط لارنس لسیگ، استاد حقوق دانشگاه هاروارد، در کتاب تأثیرگذار خود در سال ۱۹۹۹ مطرح شد. در ابتدا، به طراحی اولیه اینترنت با تأکید بر انعطاف‌پذیری شبکه و مقاومت در برابر کنترل مرکزی اشاره می‌شود، جایی که دسترسی آزاد و **ناشناس بودن** از هنجارهای اصلی محسوب می‌شد. سپس، تحولات ناشی از تجاری‌سازی اینترنت و نیازهای متعاقب آن برای احراز هویت قوی‌تر، که منجر به تغییر معماری و اولویت‌بخشی به **هویت، تمرکزگرایی و کنترل** شد، مورد بحث قرار می‌گیرد. این فصل نشان می‌دهد که چگونه کد، به مثابه یکی از چهار عامل تنظیم‌کننده رفتار انسان (در کنار قوانین، هنجارها و بازارها)، نقش پررنگی در تعیین شرایط تجربه زندگی در فضای مجازی ایفا کرده است. در ادامه، به مهاجرت این مفهوم از فضای صرفاً دیجیتال به جهان فیزیکی در عصر کنونی پرداخته می‌شود، جایی که **کد و هوش مصنوعی** در خودروها، لوازم خانگی، زیرساخت‌های شهری و سایر جنبه‌های زندگی روزمره نفوذ کرده و امکان **"کنترل کامل "**را فراهم آورده‌اند. سناریوهای کاربردی، مانند سیستم‌های تشخیص الکل در خودروهای آینده و کنترل‌های محل کار یا خانه، برای تبیین این مفهوم به کار گرفته می‌شوند. فصل همچنین تفاوت‌های اساسی میان قوانین سنتی (مبتنی بر انتخاب و رعایت داوطلبانه) و کنترل‌های معماری‌گونه (که امکان تصمیم‌گیری را از انسان سلب می‌کنند) را روشن می‌سازد. در بخش‌های پایانی، به پدیده‌هایی نظیر **"ضدقراردادها "**و **قراردادهای هوشمند** مبتنی بر بلاکچین پرداخته می‌شود که نمونه‌هایی از تجلی کد به مثابه قانون در بستر تراکنش‌ها هستند، و سپس چالش‌های ناشی از انعطاف‌پذیری و تفسیرپذیری در قراردادهای هوشمند مبتنی بر یادگیری ماشین مورد بررسی قرار می‌گیرد. در نهایت، بر لزوم بازتعریف **قرارداد اجتماعی** در پرتو این تحولات تأکید می‌شود و نقش حیاتی **"رضایت حاکمیت‌شوندگان "**در مشروعیت‌بخشی به سیستم‌های هوش مصنوعی و ادغام ایمن و مؤثر آن‌ها در جامعه برجسته می‌گردد.

مقدمه

در اواخر دهه ۱۹۹۰، در حالی که شرکت‌هایی نظیر مَگِلان و گارمین با استفاده از سامانه موقعیت‌یاب جهانی (GPS) نقشه‌برداری هر متر از جهان را آغاز کرده بودند، گوگل نیز با مأموریتی مشابه، یعنی نقشه‌برداری از هر هایپرلینک در مرز بی‌کران و پرشتاب اینترنت، گامی بلند برداشت. در بیانیه مأموریت سال ۱۹۹۸ گوگل، هدف این شرکت "سازماندهی اطلاعات جهان و قابل دسترس و مفید ساختن آن برای همگان" اعلام شد. این رویکرد در آن زمان، با توجه به اینکه تمامی کارمندان گوگل به راحتی در یک گاراژ اجاره‌ای در حومه شهر جای می‌گرفتند، جاه‌طلبانه اما قابل تصور بود. در مقابل، ایده اینکه فضای مجازی نه تنها قابل دسترس باشد، بلکه حقیقتاً قابل حکمرانی نیز باشد، در سال ۱۹۹۸ هنوز یک بدعت محسوب می‌شد. اینترنت، به خودی خود، به گونه‌ای طراحی شده بود که در برابر **نقاط شکست منفرد** مقاوم باشد – واقعیتی که آن را در برابر **نقاط اقتدار منفرد** نیز بسیار مقاوم می‌ساخت. این شبکه که از طریق پروتکل‌های فنی و نه قوانین دولتی عملیاتی شده بود، **تاب‌آوری شبکه** را بر **کنترل مرکزی** و **دسترسی آزاد** را بر **نگهبانی نهادی** ارجحیت می‌داد. جان گیلمور، از پیشگامان متن‌باز و بنیان‌گذاران بنیاد مرزهای الکترونیکی، به طور مشهور اظهار داشت" :**نت سانسور را به مثابه آسیب تفسیر می‌کند و مسیر خود را از کنار آن تغییر می‌دهد** ".این دیدگاه، چشم‌اندازی از فضای مجازی را ترسیم می‌کرد که در آن، آزادی، ناشناس بودن و خودمختاری از اصول بنیادین به شمار می‌رفتند و هرگونه تلاش برای اعمال کنترل، از اساس ناکام می‌ماند.

اما در سال ۱۹۹۹، لارنس لسیگ، استاد حقوق دانشگاه هاروارد، در کتاب خود با عنوان "کد و دیگر قوانین فضای مجازی (Code, and Other Laws of Cyberspace) "دیدگاهی کاملاً متفاوت ارائه داد. او در این کتاب توضیح داد که چگونه تجاری‌سازی اینترنت و ضرورت‌های احراز هویت کاربران که تجارت طلب می‌کرد، در حال تغییر اینترنت از مکانی با دسترسی نامحدود، ناشناس بودن و خودمختاری به مکانی است که به طور فزاینده‌ای **هویت، تمرکزگرایی و**

معماری‌های کنترل را اولویت می‌بخشد. لسیگ استدلال کرد که در دنیای واقعی، چهار عامل متمایز رفتار انسان را تنظیم می‌کنند :**قوانین، هنجارها، بازارها و معماری** .او نشان داد که در اینترنت نیز وضعیت مشابه است، با این تفاوت که در سال‌های اولیه این رسانه، **معماری، در قالب کد، نقش بسیار برجسته‌ای** ایفا می‌کرد. توسعه‌دهندگان نرم‌افزار، این دنیای جدید را از صفر و با نظارت اندکی از سوی تنظیم‌کنندگان جهان واقعی می‌ساختند. بنابراین، آنها بودند که قواعد بازی را تعیین می‌کردند. در اینترنت، **کد خود قانون بود** .این بینش عمیق لسیگ، ضرورت فهم این لایه زیربنایی از حکمرانی را بیش از پیش آشکار می‌ساخت، زیرا این لایه بود که به طور نامرئی، اما قدرتمند، بر تجربه کاربران در فضای مجازی حاکمیت می‌کرد. امروز، با گذشت ربع قرن از انتشار "کد"، این دیدگاه نه تنها اعتبار خود را حفظ کرده، بلکه با نفوذ هوش مصنوعی در هر جنبه‌ای از زندگی، اهمیت آن به مراتب افزایش یافته است. این فصل به بررسی چگونگی تکامل این مفهوم و پیامدهای آن در عصر حاضر می‌پردازد.

ریشه‌های "قانون کد است" در فضای مجازی اولیه

چشم‌انداز اولیه فضای مجازی، جهانی بود که با ایده‌آل‌های **آزادی بی‌حد و حصر، ناشناس بودن و خودمختاری کامل** تعریف می‌شد. اینترنت، به عنوان یک شبکه غیرمتمرکز، نه تنها به گونه‌ای طراحی شده بود که در برابر حملات یا خرابی‌های نقطه‌ای مقاوم باشد، بلکه این **تاب‌آوری** به معنای مقاومت ذاتی آن در برابر هرگونه **اقتدار مرکزی** نیز بود. پروتکل‌های فنی، اساس عملکرد آن را تشکیل می‌دادند و نه احکام قانونی دولت‌ها، و همین ویژگی‌ها، فضایی را ایجاد کرده بود که در آن، مرزهای فیزیکی و هنجارهای اجتماعی سنتی کم‌رنگ می‌شدند. در این بستر، توانایی افراد برای تعامل آزادانه، بیان افکار بدون ترس از سانسور، و حضور آنلاین بدون وابستگی به هویت واقعی، بی‌سابقه بود. این "روزهای باشکوه" اولیه اینترنت، به کاربران **آزادی بی‌حدی** می‌بخشید تا با سایر کاربران در "صفحه اثیری" به گونه‌ای تعامل کنند که کاملاً از حضور فیزیکی، هنجارهای اجتماعی سنتی و هرگونه نشانه هویت دنیای واقعی رها بود. همچنین

می‌توانستند محتوای جذاب یا بعضاً پیش پا افتاده‌ای را که دیگر کاربران اولیه به اشتراک می‌گذاشتند، مصرف کنند یا خودشان چنین محتوایی را منتشر کنند.

اما این آرمان‌شهر دیجیتال با ظهور **تجاری‌سازی** دستخوش تغییر شد. لارنس لسیگ به درستی دریافت که تجارت آنلاین، با نیاز اساسی‌اش به اطلاعاتی نظیر شماره کارت اعتباری، آدرس پستی و سایر **شناسه‌های شخصی**، به مثابه یک **"اسب تروا "**عمل می‌کند. این نیازها، معماری اصلی اینترنت را وادار به پذیرش افزوده‌هایی کردند که تا پیش از آن وجود نداشتند. شرایط تکنولوژیکی که زمانی به سطح جدیدی از **حریم خصوصی عمومی** امکان می‌داد، جایی که هر کسی می‌توانست در هر کجا و بدون ترس از سانسور هر چیزی بگوید، با **شناسه‌های کاربری، کوکی‌های ردیاب** و جمع‌آوری تجمعی داده‌ها که این فناوری‌ها تسهیل می‌کردند، تکمیل شد. لسیگ در این باره نوشت" :**دست نامرئی، از طریق تجارت، در حال ساخت معماری‌ای است که کنترل را به کمال می‌رساند – معماری‌ای که تنظیم‌گری بسیار کارآمد را ممکن می‌سازد "**.او توضیح داد که در دنیای واقعی، چهار عامل متمایز رفتار انسان را تنظیم می‌کنند :**قوانین، هنجارها، بازارها و معماری** .در اینترنت، وضعیت مشابه بود، با این تفاوت که در سال‌های اولیه این رسانه، معماری، در قالب کد، نقش نامتناسبی ایفا می‌کرد. توسعه‌دهندگان نرم‌افزار این دنیای جدید را از صفر و با نظارت کمی از سوی قانون‌گذاران جهان واقعی می‌ساختند. بنابراین، آنها بودند که **قواعد تعامل** را تعیین می‌کردند. در اینترنت، **کد قانون بود** .این کد یا معماری، شرایطی را تعیین می‌کرد که زندگی در فضای مجازی با آن تجربه می‌شد. این کد تعیین می‌کرد که محافظت از حریم خصوصی چقدر آسان است یا سانسور گفتار چقدر راحت است. لسیگ با هوشیاری متوجه شد که این تغییرات، در بسیاری از جهات که نمی‌توان دید مگر اینکه ماهیت این کد را درک کنیم، فضای مجازی را تنظیم می‌کند. او مخالف این پیشرفت‌ها نبود؛ او دریافت که برخی از مقررات اجتناب‌ناپذیر و در صورت اعمال عاقلانه، حتی مفید هستند. هدف او جلب توجه به **موازنه‌ها** و **پیامدهای بلندمدت** این تغییرات و ماهیت سیاسی این فرایند در کنار ابعاد

فنی آن بود.

امروزه، در اینترنتی که با شناسه‌های کاربری، کوکی‌های ردیاب و مقررات بسیار کارآمد تکمیل شده است، آزادی‌های افراد متنوع‌تر شده است. برای مثال، می‌توان با دوست دوران مهدکودک ارتباط برقرار کرد، در سهام میم سرمایه‌گذاری کرد، فیلم‌های نیکلاس کیج بیشتری نسبت به دوران قبل از پخش آنلاین مشاهده کرد، برای نمونه اولیه یک ربات چمن‌زن سرمایه‌ی جمعی جذب کرد، ضربان قلب مادربزرگ را از راه دور کنترل کرد و دوره‌های آنلاین رایگان از اساتید هاروارد را گذراند. این‌ها تنها چند مورد از **شگفتی‌های اینترنت قرن بیست و یکم** هستند.

انتقال کد از فضای مجازی به جهان واقعی

اهمیت کتاب لسیگ در سال‌های اولیه هزاره جدید و حتی بیست و پنج سال پس از آن، بیش از پیش روشن است، به ویژه با توجه به **پایان‌نامه اصلی آن: ایده "قانون کد است ."** زمانی که لسیگ "کد" را نوشت، اینترنت یا همان فضای مجازی، هنوز به طور جدایی‌ناپذیری در "دنیای واقعی" تنیده نشده بود. در آن زمان، این فضا مکانی بود که افراد گهگاه از طریق رایانه شخصی یا لپ‌تاپ خود وارد آن می‌شدند. خارج از فضای مجازی، قانون همچنان قانون بود و کنترل فیزیکی بر زندگی، با منطق خود ادامه می‌یافت.

اما امروز، فضای مجازی دیگر جهانی جداگانه نیست، بلکه همچون خطوط تلفن یا فروشگاه‌های زنجیره‌ای، **در هر کجا حضور دارد و در همه چیز نفوذ کرده است** .کد به تلفن‌ها، خودروها، لوازم خانگی، زیرساخت‌های شهری، کارخانه‌های تولیدی، سیستم‌های کشاورزی خودکار، ردیاب‌های سلامت و تناسب اندام، دستگاه‌های قابل کاشت در بدن و پروتزها، پول و بسیاری موارد دیگر راه یافته است. دستگاه‌های هوشمند و حسگردار، زیرساخت‌های ساخته‌شده ما را فراگرفته‌اند. بسیاری از این دستگاه‌های جدید توانایی دارند که به شیوه‌های فزاینده‌ای **عاملیت‌محور** عمل کنند. آنها به **تصمیم‌گیری و انجام اقدامات خودسرانه** می‌پردازند، به گونه‌ای که ممکن است بر انتخاب‌ها و تعهدات انسان‌هایی که با آنها تعامل دارند،

تـأثیر بگـذارد. ایـن نفـوذ گسـترده، ماهیـت حکمرانـی را از قـوانین سـنتی و مکـانیکی بـه **قـوانین الگوریتمی و معماری‌گونه** منتقل کرده است که به طور نامرئی و اغلب بدون آگاهی کامل کاربران، بر زندگی آنها حاکمیت می‌کنند.

سناریوی خودروی هوشمند: مثالی از کنترل معماری

بـرای درک بهتـر پیامـدهای ایـن تغییـر، فـرض کنیـد سـال ۲۰۲۷ اسـت. در حـالی کـه تاکسـی‌های رباتیـک خـودران در حـال گسـترش هسـتند، خودروهـای سـنتی کـه نیـاز بـه راننـده انسـانی دارنـد، همچنـان بـر فـروش وسـایل نقلیـه جدیـد غالب‌انـد. شـما بـه تازگی یـک شـورولت ایکوئینـاکس EV خریداری کرده‌اید و مانند تمام خودروهای جدید فروخته شده در ایالات متحده در آن سال، ایـن خـودرو مجهـز بـه **سیسـتم فدرالی تشـخیص الکل راننـده بـرای ایمنـی (DADSS)** اسـت. چنیـن سیسـتم‌هایی از حسـگرهای لمسـی و تنفسـی کـه در قسـمت‌های مختلـف داخلـی خـودرو تعبیـه شـده‌اند، بـرای انـدازه‌گیری غیرفعـال سـطح الکـل خـون رانـنـده اسـتفاده می‌کننـد. بـرخلاف دسـتگاه‌های تسـت الکل(Breathalyzer) ، نیـازی بـه فعالانـه دمیـدن در ایـن حسـگرها نیسـت. در عوض، آنها از نور مادون قرمز برای نظارت غیرفعال بر رانندگان استفاده می‌کنند. با لمس فرمان یا سـوئیچ اسـتارت، حسـگرهای تعبیه‌شـده در آنجـا، میـزان الکـل خـون را در مویرگ‌هـای انگشـتان تشـخیص می‌دهنـد. بـه سـادگی بـازدم کنیـد، و حسـگرهای تعبیه‌شـده در داشـبورد یـا درب سـمت راننده، میزان الکل خون را در نفس شما تشخیص می‌دهند.

این حسگرها خودشان هوش مصنوعی را در بر نمی‌گیرند. اما خودروی شما همچنین مجهز به ویژگی‌های بیشتری – یک دوربین رو به راننده و حسگرهای اضافی – است که از **یادگیری ماشین** بـرای تحلیـل عـواملی ماننـد **حالـت بـدن، الگوهـای گرفتن فرمـان و جهـت جریـان هـوا** اسـتفاده می‌کننـد تـا کمـک کننـد مشـخص شـود ایـن راننـده اسـت کـه الکل خـونش انـدازه‌گیری می‌شـود، نـه هیچ یک از مسافران احتمالی دیگر در خودرو.

حال سناریویی را تصور کنید که پس از یک شام دلپذیر که شامل بیش از یک (یا سه) لیوان

شراب بوده، شما با "ناوی‌تار(NaviTar) "، رابط کاربری مجهز به مدل زبان بزرگ (LLM) برای کل سیستم مدیریت خودروی خود، درباره صلاحیت رانندگی‌تان بحث می‌کنید. ناوی‌تار تشخیص داده است که سطح الکل خون شما ۰.۱۰ است، یعنی ۰.۰۲ بیشتر از حد مجاز قانونی در کالیفرنیا، و در نتیجه، به خودرو اجازه استارت زدن نمی‌دهد. شما سعی می‌کنید توضیح دهید که گارسون شراب را روی میز ریخته است و حتی آستین کمی لکه‌دار خود را به دوربین رو به راننده نشان می‌دهید. ناوی‌تار پاسخ می‌دهد که خوشحال خواهد شد برای شما اوبر بگیرد یا فیلمی برای سرگرمی‌تان پخش کند تا زمانی که هوشیار شوید. او پیشنهاد می‌دهد: "چطور است **'سرعت '**را تماشا کنید؟ یک هیجان‌انگیز اکشن کلاسیک با ترکیب مناسبی از تعلیق، درام پرخطر و طنز که زمان را به سرعت می‌گذراند. و هفتاد و شش دقیقه طول می‌کشد. تا آن زمان، شما هوشیار خواهید شد".

"خیلی بامزه است، ناوی‌تار. و نه، فیلم نمی‌خواهم،" آهی می‌کشید. "اما دوباره بگو. چطور به اینجا رسیدیم؟ یعنی اینجا آمریکا نیست؟ من حق قانون اساسی برای انتخاب اینکه آیا می‌خواهم ریسک رانندگی در حالت مستی را بپذیرم یا نه، ندارم؟"

فناوری‌هایی که این سناریو را ممکن می‌سازند، همگی در حال حاضر وجود دارند .در واقع، تا سال ۲۰۲۷، چنین سیستم‌های تشخیص الکلی) نه لزوماً با جزء (LLM ممکن است برای تمام وسایل نقلیه جدید خریداری شده در ایالات متحده اجباری شوند. در سال ۲۰۲۱، کنگره قانونی را به همین منظور به عنوان بخشی از **قانون سرمایه‌گذاری زیرساخت‌ها و مشاغل** تصویب کرد. در حالی که زمان دقیق وقوع این امر در زمان انتشار این کتاب هنوز مشخص نبود، این امکان وجود دارد که این اتفاق به زودی در سال ۲۰۲۶ رخ دهد. این سناریو چگونه به نظر می‌رسد؟ **دیستوپیایی؟ اتوپیایی؟ شاید کمی از هر دو؟** مطمئناً، با نزدیک شدن این فناوری‌ها و سناریوها به واقعیت و استفاده عمومی، مدافعان آزادی‌های مدنی، سازمان‌های تجاری صنعت مهمان‌نوازی، تولیدکنندگان خودرو، مدافعان مصرف‌کننده و شکاکان فناوری ممکن است با این

مقررات مخالفت کنند و شاید آن را متوقف یا به تعویق اندازند. قطعاً حداقل برخی از میلیون‌ها راننده‌ای که به طور معمول قوانین کمربند ایمنی و محدودیت‌های سرعت را نادیده می‌گیرند، به این پیشنهاد که خودروهایشان، که مدت‌ها نماد **خودمختاری فردی** بوده‌اند، به طور مؤثری به پلیس ترافیک تبدیل شوند، مقاومت خواهند کرد.

ما در این فصل درباره فضایل **نوآوری بدون مجوز**، **استقرار تکراری** و مزایای **تحمل درجاتی از ریسک هوشمندانه** در پی پیشرفت، به جای درخواست عدم قطعیت مطلق قبل از معرفی فناوری‌های جدید که نحوه کار جهان را تغییر می‌دهند، صحبت کرده‌ایم. اما رانندگی در حالت مستی یک خطر شناخته شده است که به طور مداوم هزینه‌های شخصی و اجتماعی قابل توجهی را به همراه دارد. با توجه به اینکه کنگره **بند DADSS** را در قانون زیرساخت‌ها گنجانده است، به وضوح این یک ریسکی است که قانون‌گذاران ما معتقدند باید به طور جمعی آن را discouragedکنیم.

کنترل کامل و آژانس انسانی

زمانی که لارنس لسیگ درباره **"ظرفیت کنترل کامل کد "**نوشت، دقیقاً به همین پدیده اشاره می‌کرد. با سایر انواع محدودیت‌ها، از جمله قانون سنتی، **آژانس فردی** نقش کلیدی ایفا می‌کند. همانطور که لسیگ بیان کرد، **"قانون، فرمانی است که با تهدید به مجازات همراه است ".** تابلوهای راهنمایی و رانندگی شهرداری، یکی از قدیمی‌ترین اشکال اتوماسیون نظارتی، نمونه روشنی از نحوه عملکرد این موضوع را ارائه می‌دهند. تابلوهای قرمز و سفید با حروف درشت اعلام می‌کنند" :**پارک کردن در هر زمان ممنوع است. وسایل نقلیه غیرمجاز با هزینه مالک بکسل خواهند شد "**.این تابلو، با تمام اقتدار بی‌چون و چرایش، به شما **اختیار تصمیم‌گیری** درباره میزان دقیق رعایت آن را می‌دهد. زیرا در بسیاری از موارد، احتمالاً می‌توانید در آنجا پارک کنید. انجام این کار ممکن است برای شهروندان دیگر ناراحتی یا حتی خطر ایجاد کند (مثلاً اگر در جلوی یک شیر آتش‌نشانی پارک کنید). اما اگر واقعاً نمی‌توانستید در آنجا پارک کنید، نیازی به

نصب تابلو نبود.

به این ترتیب، تقریباً تمام قوانین به میزان زیادی به **تصمیم داوطلبانه شما برای رعایت آنها** وابسته هستند. هر نقشی که نیروی انتظامی در حفظ نظم و ثبات ایفا می‌کند، تنها بخشی از یک بافت اجتماعی بسیار بزرگ‌تر است که از همکاری جامعه، اعتماد متقابل و هنجارهای رفتاری گسترده تشکیل شده است – زیرا افسران پلیس نمی‌توانند همه جا باشند. بنابراین، ممکن است در ساعت ۴ صبح با چراغ قرمز آهسته عبور کنید، در حالی که هیچ عابر پیاده یا راننده دیگری در دیدرس نیست. ممکن است وقتی فقط چند بلوک تا خواربارفروشی رانندگی می‌کنید، زحمت نبندید کمربند ایمنی خود را. ممکن است از پرداخت هزینه پارکومتر صرف‌نظر کنید زیرا در کمتر از یک دقیقه از آن فروشگاه خارج خواهید شد. در تمام این موارد، **شما در حال انتخاب هستید.**

اما یک **پارکومتر حسگردار** که به یک سیستم مرکزی مدیریت پارکینگ متصل است، می‌تواند شما را از این معادله خارج کند. اگر پس از ساعت ۶ عصر در محلی که پارک ممنوع است، پارک کنید و بیش از پنج دقیقه آنجا بمانید، پارکومتر ممکن است به سادگی جریمه از پیش تعیین شده‌ای را از حساب EZ-Park شما کسر کند. در این سناریو، **یک دستگاه تخلف را مشاهده می‌کند، قانون را اعمال می‌کند و مجازات را اجرا می‌کند** .در نقطه‌ای، بین دوربین‌ها و حسگرهای همه‌جا حاضر و **کنترل کامل مبتنی بر کد**، اجرای قانون بالقوه می‌تواند آنقدر ارزان شود که حتی نیازی به اندازه‌گیری آن نباشد. به طور سنتی، ما به محدودیت‌های سرعت، علائم "پارک ممنوع"، ممنوعیت‌های سیگار کشیدن و سایر انواع قوانین و سازوکارهای نظارتی به عنوان بخشی از یک قرارداد اجتماعی بزرگ‌تر احترام می‌گذاریم، جایی که می‌پذیریم رعایت ما به نفع کل جامعه است و همچنین همان آداب و حمایت‌ها را از دیگران به ما می‌دهد.

اما این نیز درست است که رعایت این قرارداد نسبتاً آسان است، بخشی به این دلیل که می‌دانیم **اجرای قانون ناقص و متناوب** است. با این حال، اکنون ما فناوری‌هایی در اختیار داریم که امکان اجرای دقیق و بی‌مضایقه قوانین را فراهم می‌آورند، که **تهدید می‌کند آزادی‌ای را که به**

طور معمول به خود اعطا می‌کردیم، از بین ببرد .همانطور که لسیگ در "کد" می‌نویسد : **"معماری‌هایی ظهور می‌کنند که آزادی‌ای را که صرفاً با ناکارآمدی انجام کاری متفاوت حفظ شده بود، جابجا می‌کنند".**

هوش مصنوعی برای کنترل کامل ضروری نیست. دوربین‌های چراغ قرمز که در چراغ‌های راهنمایی و رانندگی تعبیه شده‌اند و نیازی به الگوریتم‌های پیچیده یا فرایندهای تصمیم‌گیری برای انجام کار خود ندارند، یکی از نمونه‌های گسترده‌ای هستند که در حال حاضر به کار گرفته می‌شوند. اما **هوش مصنوعی امکانات را گسترش می‌دهد** .با وسایل نقلیه خودران، نه تنها محدودیت‌های سرعت به طور کامل قابل اجرا می‌شوند، بلکه **خود سرعت غیرمجاز ممکن است دیگر به سادگی ممکن نباشد** .یک چهارراه در مرکز شهر را تصور کنید که در آن هر مورد **عبور غیرمجاز از خیابان در ساعات شلوغی، منجر به جریمه خودکار می‌شود** .جایی که نقض‌های صوتی، حیوانات خانگی بدون قلاده و مستی عمومی به طور مؤثری به **تخلفات با تحمل صفر** تبدیل می‌شوند.

در حالی که چنین سناریوهای گسترده‌ای به دلایل سیاسی، اخلاقی و اقتصادی در آینده نزدیک بعید به نظر می‌رسند، **کنترل کامل در زمینه‌های خاص‌تر** چطور؟ یک چکش بادی در یک سایت کاری را تصور کنید که مجهز به **بینایی کامپیوتری** است. قبل از شروع به کار، شما را اسکن می‌کند تا تأیید کند که عینک ایمنی و کفش ایمنی مورد نیاز OSHA برای آن کار را به تن دارید. یک پارک ملی را تصور کنید که با شبکه‌ای از دوربین‌ها و حسگرها مجهز شده است که با الگوریتم‌های هوش مصنوعی همکاری می‌کنند تا کیفیت هوا، زباله، سطوح صدا و تراکم جمعیت را در زمان واقعی به طور مداوم ارزیابی کنند و **بازدیدکنندگان اضافی را زمانی که تشخیص دهند سطوح غیرقابل تحمل از تأثیرات زیست‌محیطی در حال وقوع است، ممنوع کنند.**

کاربردهای کنترل کامل در حوزه‌های خاص

این شکل از کنترل فراگیر می‌تواند به زمینه‌های مختلف زندگی روزمره نفوذ کند. در محل کار،

کارفرمای شما ممکن است رایانه شخصی شما را با **هوش مصنوعی** مجهز کند که می‌تواند با تقویم و ابزارهای مدیریت پروژه شما همگام شود تا از زمان سررسیدهای قریب‌الوقوع شما مطلع شود. اگر عقب بمانید، هوش مصنوعی ممکن است شما را در **"حالت تمرکز "**قرار دهد و تمام برنامه‌ها و اعلان‌های غیرضروری را تا زمانی که پیشرفت بیشتری داشته باشید، غیرفعال کند و به طور خودکار اعلانی را به بخش منابع انسانی ارسال کند.

ارائه‌دهنده بیمه خانه شما می‌تواند سیستم‌های تطبیقی پویا را ارائه دهد که از **حسگرهای هوشمند** برای تشخیص نشت، خطاهای الکتریکی، نشانه‌های آلودگی آفات، سطوح رطوبت در دیوارها و موارد دیگر استفاده می‌کنند. در حالی که برای انجام تعمیر و نگهداری بر اساس تجزیه و تحلیل‌های تولید شده توسط این حسگرها، انگیزه‌هایی دریافت می‌کنید، سیستم شما ممکن است در برخی **سناریوهای "حیاتی "**اقدامات قوی‌تری را برای تسریع رعایت شما به کار گیرد، مانند **به طور خودکار از کار انداختن کوره قدیمی‌ای که ماه گذشته فرصت تعویض آن را پیدا نکردید** .اکنون دما ده درجه سانتیگراد است، سیستم شما کوره را به حالت خواب ابدی منتقل کرده است، اما شما نمی‌توانید برای حداقل سه روز یک تکنسین HVAC پیدا کنید.

چه اتفاقی می‌افتد اگر بسیاری از کارفرمایان شروع به اجباری کردن "حالت تمرکز" به عنوان یک شرط استخدام کنند؟ چه اتفاقی می‌افتد اگر رویکرد بیمه خانه که در بالا توضیح داده شد، آنقدر برای ارائه‌دهندگان خوب عمل کند که همگی ارائه گزینه‌های دیگر را متوقف کنند؟ اینجاست که **معماری، جایگزین انتخاب می‌شود و عملاً آزادی عمل فرد را از بین می‌برد.**

کد به مثابه قرارداد: از "ضدقراردادها" تا قراردادهای هوشمند

در کتاب "کد"، لسیگ اشاره می‌کند که "محدودیت‌های معماری، پس از اینکه ایجاد شدند، اثر خود را تا زمانی که کسی آنها را متوقف کند، ادامه می‌دهند ".به عبارت دیگر، کنترل کامل، جایگزین هرگونه صلاحدید یا انعطاف‌پذیری می‌شود .شوشانا زوبوف در کتاب "عصر سرمایه‌داری نظارتی (The Age of Surveillance Capitalism) "استدلال می‌کند که قراردادهای جدید

مبتنی بر رایانه که کد آنها را ممکن می‌سازد، مانند قراردادی برای یک خودرو که اگر پرداخت‌های آن را متوقف کنید، روشن نمی‌شود، در واقع قرارداد نیستند، بلکه "ضدقرارداد "هستند. او می‌نویسد: "ضدقرارداد، قرارداد را از حالت اجتماعی خارج می‌کند و از طریق جایگزینی رویه‌های خودکار به جای وعده‌ها، گفتگو، معنای مشترک، حل مسئله، حل اختلاف و اعتماد – یعنی بیان‌های همبستگی و عاملیت انسانی که طی هزاره‌ها در مفهوم 'قرارداد' نهادینه شده‌اند – قطعیت تولید می‌کند." البته، مگر اینکه درباره دست دادن‌های سنتی صحبت کنیم، قراردادهای سنتی خود به گونه‌ای طراحی شده‌اند که نیاز به وعده‌ها، گفتگو، معنای مشترک و اعتماد را تا حد زیادی کاهش و حتی از بین ببرند. با این حال، منظور زوبوف را می‌توان درک کرد. او نیز مانند لسیگ، در حال توصیف جهانی جدید از کنترل کامل است که کد آن را ممکن می‌سازد و پیامدهای بالقوه آن جهان جدید را بیان می‌کند. طبیعتاً، همه آنها خوب نیستند. یک "ضدقرارداد"، یا توافق خوداجراکننده، که شرایط آن به طور سفت و سخت توسط کد حک شده است، می‌تواند ظرفیت‌های انسانی قضاوت، مذاکره، انعطاف‌پذیری، استدلال اخلاقی، بخشش و همدلی را به طور کامل بی‌اثر سازد.

در جهان‌بینی زوبوف، ضدقراردادها به طور یک‌جانبه توسط پلتفرم‌های **فناوری‌های بزرگ** بر کاربران با **کمترین یا بدون عاملیت** تحمیل می‌شوند. کد به مثابه قانون که در زمینه‌های **همتا به همتا**، مانند بیت‌کوین و سایر سیستم‌های مبتنی بر بلاکچین، یا حتی در تبادلات متقابل با شرکت‌های تجاری به کار می‌رود، خارج از حوزه علاقه سرمایه‌داری نظارتی قرار دارد. اما **قراردادهای خودکار توافقی** وجود دارند و در چنین زمینه‌هایی، کنترل کاملی که توسط کد فراهم می‌شود، می‌تواند برای هر دو طرف مفید باشد.

قراردادهای هوشمند مبتنی بر بلاکچین

یک آژانس شبکه‌های اجتماعی را در نظر بگیرید که با یک برند خرده‌فروشی برای مدیریت حساب X.com برند قرارداد می‌بندد. برای تعیین شرایط توافق خود، دو طرف از یک **قرارداد**

هوشمند مبتنی بر بلاکچین استفاده می‌کنند. بلاکچین‌ها دفاتر کل توزیع‌شده دیجیتالی هستند که تراکنش‌ها را در بسیاری از رایانه‌ها (که در این زمینه به آنها نود گفته می‌شود) ثبت می‌کنند. این تراکنش‌ها در **بلوک‌ها** —بسته‌های داده‌ای که شامل چندین تراکنش، یک مهر زمانی و ارجاع به بلوک قبلی هستند— گروه‌بندی می‌شوند. هنگامی که یک بلوک به زنجیره اضافه می‌شود، تراکنش‌های موجود در آن را نمی‌توان به طور عطف به ماسبق و بدون تغییر تمام بلوک‌های بعدی تغییر داد. این ساختار، یک سیستم **غیرمتمرکز، شفاف و مقاوم در برابر دستکاری** برای ذخیره و تأیید اطلاعات ایجاد می‌کند.

در مثال ما، برای شروع توافق، یکی از دو طرف باید قرارداد هوشمند خود را به یک بلاکچین مورد توافق طرفین اضافه کند. در آنجا، به عنوان یک برنامه **خوداجراکننده** حاوی شرایط توافق به صورت کد وجود دارد. چنین شرایطی می‌تواند **اهداف عملکردی** را که انتظار می‌رود آژانس در طول دوره واگذاری به آنها دست یابد، مقادیر پرداخت و زمان‌بندی‌ها را مشخص کند. برای مثال، قرارداد می‌تواند آژانس را ملزم کند که تعداد معینی پست در هر هفته ارسال کند، یا برای رسیدن به اهداف تعامل مختلف، پاداش‌هایی اعطا کند. از طریق مکانیسم‌های خودکار مختلف، خود قرارداد می‌تواند تشخیص دهد که آیا آژانس در حال انجام شرایط قرارداد است یا خیر. اگر چنین باشد، قرارداد در پایان ماه پرداخت را آزاد می‌کند. اگر آژانس عملکرد لازم را نداشته باشد، قرارداد ممکن است تنها پرداخت جزئی به آژانس یا حتی بازپرداخت کامل به مشتری — بر اساس هر شرایطی که قرارداد مشخص کرده است — را صادر کند.

هر اتفاقی که بیفتد، نه آژانس و نه برند قادر نخواهند بود به طور یک‌جانبه شرایط قرارداد را تغییر دهند یا نتایج را دستکاری کنند. در عوض، **شبکه غیرمتمرکزی که بلاکچین را اجرا می‌کند، یکپارچگی توافق قرارداد را تضمین می‌کند و هر دو طرف را از تلاش‌های لحظه آخری برای تغییر شرایط تعامل محافظت می‌کند** .در حالی که این رویکرد فاقد لمس انسانی یک توافق شفاهی است، **انصاف و شفافیت را بر انعطاف‌پذیری و صلاحدید اولویت می‌دهد** .با این کار، عدم تعادل‌هایی را که ممکن است در اشکال سنتی قراردادها وجود داشته باشد، زمانی که یک

طرف ممکن است منابع بسیار بیشتری برای اجرای یا نادیده گرفتن شرایط توافق داشته باشد، کاهش می‌دهد.

این بدان معنا نیست که چنین قراردادهایی برای هر سناریویی مناسب هستند. و قطعاً ارزش آن را دارد که در نظر بگیریم چقدر می‌خواهیم در جهانی زندگی کنیم که در آن قضاوت، مذاکره و بخشش کمتر ضروری می‌شوند، و حتی عمیق‌تر، به همان اندازه که گرفتن یک تصمیم غیراخلاقی دشوار است، گرفتن یک تصمیم فضیلت‌مندانه نیز دشوار می‌شود، زیرا تنها یک مسیر عملی وجود دارد. اما این نیز درست است که **توسط یک مشتری dishonest گیر نیفتادن، قدرتمندکننده است** .بازپس‌گیری سریع ودیعه تضمینی خود در پایان اجاره کوتاه مدت تعطیلات، پس از اثبات اینکه محل را همانطور که پیدا کرده‌اید، از طریق عکس‌های دارای مهر زمانی که به بلاکچین ارسال شده‌اند، دست‌نخورده گذاشته‌اید، می‌تواند حس عاملیت شما را به طرز چشمگیری افزایش دهد.

تکامل قانون کد با هوش مصنوعی

جنبه دیگری نیز وجود دارد که باید در نظر گرفته شود. هنگامی که لسیگ کتاب "کد" را نوشت، الگوریتم‌های **یادگیری ماشین (ML)** معمولاً بخشی از روال‌های خوداجراکننده نبودند که کد به مثابه قانون اجرا می‌کرد. حتی امروزه، قراردادهای هوشمندی که روی بلاکچین‌هایی مانند اتریوم و سولانا اجرا می‌شوند، باید با **کد قطعی و مبتنی بر قوانین** نوشته شوند که همیشه خروجی‌های یکسانی را در هنگام دریافت ورودی خاص تولید می‌کنند. این بدان دلیل است که شبکه‌های بلاکچین به **مکانیسم‌های اجماع** متکی هستند که در آن همه نودها باید به طور مستقل اجرای قراردادهای هوشمند را تأیید و بر آن توافق کنند. اگر قراردادها بتوانند خروجی‌های مختلفی را برای یک ورودی تولید کنند – همانطور که مدل‌های ML غیرقطعی ممکن است – برای شبکه غیرممکن خواهد بود که بر وضعیت صحیح بلاکچین به اجماع برسد.

با این حال، اکنون می‌توان قراردادهایی را نوشت که الگوریتم‌های یادگیری ماشین را در خود

جای می‌دهند. با این کار، می‌توان به طور بالقوه **انعطاف‌پذیری‌ای را که قوانین و قراردادهای سنتی که توسط انسان‌ها اداره می‌شوند، مشخص می‌کند، تقریبی کرد** .همانطور که سامر حسن و پریماورا دی فیلیپی، دو محقق وابسته به مرکز برکمن کلاین برای اینترنت و جامعه در دانشگاه هاروارد، در مقاله‌ای در سال ۲۰۱۷ می‌نویسند: "یادگیری ماشین امکان معرفی قوانین مبتنی بر کد را فراهم می‌کند که ذاتاً **پویا و تطبیق‌پذیر** هستند – بنابراین برخی از ویژگی‌های قوانین حقوقی سنتی را که با **انعطاف‌پذیری و ابهام زبان طبیعی** مشخص می‌شوند، تکرار می‌کنند. در واقع، تا آنجا که می‌توانند از داده‌هایی که جمع‌آوری یا دریافت می‌کنند، یاد بگیرند، این سیستم‌ها می‌توانند به طور مداوم قواعد خود را برای مطابقت بهتر با شرایط خاصی که برای آنها طراحی شده‌اند، **تکامل** دهند".

قراردادهای پویا و چالش‌های آنها

به این ترتیب، **قراردادها به مثابه کد** می‌توانند حتی از قراردادهای (و قوانین) نوشته شده توسط انسان **انعطاف‌پذیرتر و تطبیق‌پذیرتر** شوند. یک قرارداد هوشمند برای **بیمه محصولات کشاورزی** را در نظر بگیرید که از الگوریتم‌های یادگیری ماشین برای **تنظیم پویا پرداخت‌ها بر اساس داده‌های محیطی در زمان واقعی** استفاده می‌کند. برخلاف یک قرارداد بیمه سنتی با شرایط ثابت، این قرارداد نه تنها یک سند کتبی یا یک تکه کد ثابت نیست، بلکه یک **سیستم پویا** است که از شبکه‌ای از حسگرهای رطوبت خاک، تراکتورهای هوشمند، گیرنده‌های GPS و سایر دستگاه‌های فیزیکی و اجزای دیجیتال استفاده می‌کند تا به طور مداوم ریسک را بر اساس الگوهای آب و هوا، شرایط خاک و سایر عوامل مرتبط ارزیابی و باز محاسبه کند. این قرارداد که به چنین داده‌هایی مجهز است، می‌تواند **حق بیمه و پرداخت‌ها را به طور پویا تنظیم کند** .در دوره‌های کم‌خطر، حق بیمه کمتری برای تشویق به پوشش ارائه می‌دهد. در فصول پرخطر، مانند فصل طوفان در مناطق ساحلی، حق بیمه را برای پوشش ریسک افزایش می‌دهد. این سیستم ساختارهای پرداخت ثابتی را بر اساس خسارت واقعی حفظ می‌کند، اما توانایی آن در

ارزیابی دقیق و قیمت‌گذاری ریسک با گذشت زمان بهبود می‌یابد.

قراردادهای پویاتر می‌توانند ابهام و عدم قطعیت بیشتری نیز به همراه داشته باشند. به هر میزان که قراردادهای مبتنی بر کد قدیمی‌تر و مبتنی بر قوانین، قضاوت، مذاکره و سایر موارد تعامل انسانی را از توافقات قراردادی حذف می‌کنند، **امکان فساد، تبعیض، فریب و اجرای ناسازگار** را نیز کاهش می‌دهند .**عدم انعطاف‌پذیری آنها نه تنها یک ضعف، بلکه یک نقطه قوت است** .برای قراردادهای نوشته شده توسط انسان و قراردادهای مبتنی بر هوش مصنوعی، انعطاف‌پذیری هم مزایا و هم آسیب‌پذیری‌هایی دارد.

مسئله تفسیرپذیری

تفسیرپذیری، یا توانایی درک دقیق اینکه یک سیستم چگونه تصمیمات خود را می‌گیرد، در تمام کاربردهای یادگیری ماشین یک مسئله است. با این حال، در زمینه‌های حقوقی، به دلیل انتظارات در مورد **شفافیت و اجرای برابر قانون**، از اهمیت بالایی برخوردار می‌شود. اگر حتی متخصصان نیز نتوانند دقیقاً نحوه تصمیم‌گیری‌ها را درک کنند، پس چگونه افرادی که این قراردادها یا قوانین بر آنها تأثیر می‌گذارند، می‌توانند نتایجی را که فکر می‌کنند ناعادلانه هستند، به چالش بکشند؟ چگونه حسابرسان و مکانیسم‌های نظارتی می‌توانند اطمینان حاصل کنند که قراردادها به طور منصفانه و بدون تبعیض اداره می‌شوند؟

با این حال، اگر بتوانیم بر این چالش‌ها غلبه کنیم، به آنچه **قانون پویا به مثابه کد** می‌تواند دست یابد، فکر کنید. یک **قرارداد مبتنی بر هوش مصنوعی برای مدیریت زنجیره تأمین** ممکن است امکان تنظیمات خودکار را در طول اختلالات جهانی فراهم کند. می‌تواند قیمت‌گذاری یا مقادیر را بر اساس نوسانات بازار تغییر دهد. در بهترین حالت، چنین سیستم‌هایی می‌توانند با **رهنمودهای اخلاقی و اصول انصاف** برنامه‌ریزی شوند و به آنها اجازه دهند تا پایبندی دقیق به قوانین را با ملاحظات انسانی‌تر از عدالت و معقولیت متعادل کنند. به این ترتیب، قانون پیشرفته به مثابه کد ممکن است نه تنها **قضاوت انسانی را تقلید کند، بلکه به طور بالقوه آن را با**

بینش‌های مبتنی بر داده و اجرای سازگار در موارد بی‌شمار افزایش دهد.

بازنویسی قرارداد اجتماعی در عصر هوش مصنوعی

در کتاب "کد"، لارنس لسیگ بین **"مشتریان" و "اعضا"** تمایز قائل شد – یعنی بین کسانی که حقوقشان از حمایت (یا عدم حمایت) آنها ناشی می‌شود و کسانی که ادعاها و امتیازاتشان از **حق بنیادینشان برای مشارکت در حکمرانی سازمانی** که بخشی از آن هستند، ناشی می‌شود. در مک‌دونالد، شما یک مشتری هستید که عاملیت شما تا حدی بر اساس توانایی شما برای رفتن به برگر کینگ است، اگر مک‌دونالد به نوعی به شما اجازه نمی‌دهد "همانطور که دوست دارید" عمل کنید. در ایالات متحده، شما یک عضو (یک شهروند) هستید که عاملیت شما تا حدی توسط حمایت‌ها و امتیازات قانون اساسی که با شهروندی همراه است، تضمین می‌شود.

لسیگ استدلال کرد که رابطه ما با فضای مجازی باید به نوعی از "اعضا" باشد: "چرا شهروندان واقعی باید کنترلی بر فضای مجازی یا معماری‌های آن داشته باشند؟ ممکن است بیشتر زندگی خود را در یک مرکز خرید بگذرانید، اما هیچ کس نمی‌گوید که شما حق کنترل معماری مرکز خرید را دارید. یا ممکن است دوست داشته باشید هر آخر هفته از دیزنی‌لند بازدید کنید، اما ادعا کنید که به همین دلیل حق تنظیم‌گری دیزنی‌لند را دارید، عجیب خواهد بود." او پیشنهاد کرد که چون فضای مجازی جامع‌تر از یک مرکز خرید یا حتی دیزنی‌لند است، ما باید بیش از صرفاً مشتریان آن باشیم؛ ما باید **ذی‌نفعانی با حق اظهار نظر** باشیم. او نوشت: "هم به عنوان مسئله عدالت و هم به عنوان مسئله شکوفایی انسانی، ما به این بخش‌ها از زندگی‌مان نیاز داریم که بر معماری‌هایی که تحت آن زندگی می‌کنیم، کنترل داشته باشیم".

تا حدی، لسیگ در این مورد خاص، محدودیت‌های اعمال شده بر **معماری فیزیکی** را دست‌کم می‌گرفت. به هر حال، **کدهای ساختمانی** وجود دارند که معماری مراکز خرید و پارک‌های تفریحی را شکل می‌دهند. اما مهم‌تر از آن، معماری فیزیکی در سال ۱۹۹۹، زمانی که لسیگ "کد" را نوشت، هنوز نمی‌توانست به روش‌های بسیار دقیق و پیچیده‌ای که معماری

مجازی می‌توانست، محدودیت ایجاد کند. با این حال، این وضعیت با ادغام فزاینده جهان‌های فیزیکی و مجازی در حال تغییر است. امروز، **امکانات کنترل کامل در فضای فیزیکی با امکانات فضای مجازی رقابت می‌کند** .یک مثال اخیر مربوط به MSG Entertainment است که مدیسون اسکوئر گاردن و دیگر مکان‌های سرگرمی را اداره می‌کند. برای چندین سال است که این شرکت یک سیاست بحث‌برانگیز را اجرا کرده است که از **فناوری تشخیص چهره** برای شناسایی و رد ورود وکلا که در شرکت‌های حقوقی درگیر پرونده‌های قضایی علیه آن کار می‌کنند، استفاده می‌کند. هنگامی که چنین مشتریانی سعی می‌کنند وارد مدیسون اسکوئر گاردن یا رادیو سیتی میوزیک هال شوند، چهره آنها اسکن و با پایگاه داده مقایسه می‌شود. اگر با فردی در لیست ممنوعه مطابقت یافت شود، آن شخص از ورود محروم می‌شود، حتی اگر بلیط معتبری داشته باشد.

فضاهای فیزیکی همیشه توانایی گنجاندن مکانیسم‌های مرتب‌سازی را داشته‌اند: صف‌های VIP، کدهای لباس، قیمت بلیط و انتخاب‌های معماری عمدی، مانند ورودی‌های محدود. اما این‌ها ابزارهای نسبتاً کندی در مقایسه با مرتب‌سازی بودند که به سرعت در فضاهای دیجیتال ممکن شد. اکنون **معماری فیزیکی، حسگردار و مجهز به ابزار، می‌تواند با کارایی کد تنظیم‌گری کند** .سیاست MSG Entertainment منجر به دعاوی حقوقی متعدد، تحقیقات قانون‌گذاران، پیشنهادهای قانونی و بحث‌های جاری درباره استفاده از تشخیص چهره در مکان‌های خصوصی و تعادل قدرت بین حقوق مالکیت و اقامت عمومی شده است.

در سال‌های آینده، مواردی مانند این—که در آن دستگاه‌ها و خدمات هوش مصنوعی می‌توانند "انتخاب‌هایی" را که ما به عنوان افراد مجاز به انجام آنها هستیم، شکل دهند، ترغیب کنند، خودکار کنند، دیکته کنند و حتی از پیش تعیین کنند —رایج‌تر خواهند شد. دعاوی حقوقی بیشتری طرح خواهد شد. تلاش‌های بیشتری برای تدوین قوانینی که انواع قانون به مثابه کد را که در دنیای فیزیکی مجاز هستند، تنظیم می‌کنند، صورت خواهد گرفت. اما هر قانونی

که تصویب شود یا نشود، نگرش‌های عمومی به وضوح نقش عمده‌ای در نحوه استقبال ما از این سناریوهای جدید ایفا خواهد کرد. این امر به ویژه اگر آژانس‌های دولتی مختلف شروع به اعمال مکانیسم‌های کنترل کامل مبتنی بر هوش مصنوعی خود کنند، صادق خواهد بود.

اهمیت رضایت حاکمیت‌شوندگان

هنگامی که ما انتخاب می‌کنیم مناطق پارک ممنوع را رعایت کنیم، مالیات بر درآمد خود را بپردازیم، به مناطق سیگار ممنوع احترام بگذاریم و به طیف وسیعی از هنجارها و رفتارهای فرهنگی پایبند باشیم، **به طور فعال در یک قرارداد اجتماعی بزرگ‌تر شرکت می‌کنیم .رعایت داوطلبانه**، به جای اجرای دائم، کلید عملکرد این سیستم‌ها است. در حالی که قوانین و هنجارهای اجتماعی چارچوب را فراهم می‌کنند، آنچه حتی مهم‌تر است، میزان آمادگی عمومی برای پذیرش آنهاست. قوانین و هنجارها کار می‌کنند زیرا ما آنها را انتخاب می‌کنیم و به آنها رضایت می‌دهیم. بخش بزرگی از عملکرد آنها این است که جامعه را به عنوان گروهی از اعضا که با میل خود حول مجموعه‌ای از قوانین و ارزش‌ها متحد شده‌اند، تثبیت کنند. در عصری که **کنترل کامل** در زمینه‌های بیشتر امکان‌پذیر می‌شود، این حقیقت از همیشه بیشتر صدق می‌کند.

مفاهیمی که با این فرایند **رعایت داوطلبانه** مرتبط شدند، توسط فیلسوفان عصر روشنگری، جان لاک و ژان ژاک روسو، که آنها را در زمینه **نظریه قرارداد اجتماعی** مورد بحث قرار دادند، محبوبیت یافتند. فرض اساسی آنها این بود که دولت‌هایی که اقتدار و مشروعیت خود را بدون توسل به زور یا ادعای حق الهی به دست می‌آورند، **عادل‌تر، باثبات‌تر و پایدارتر** هستند، زیرا چنین دولت‌هایی به عنوان **اراده مردم** عمل می‌کنند. توماس جفرسون در تدوین **اعلامیه استقلال** در سال ۱۷۷۶، این مفهوم را در یک بخش کلیدی توضیح داد و عبارت **"رضایت حاکمیت‌شوندگان "** را ابداع کرد. این مفهوم بیان می‌دارد که دولت‌ها قدرت‌های عادلانه خود را از رضایت حاکمیت‌شوندگان می‌گیرند و هر زمان که هر شکلی از دولت به این اهداف مخرب شود، حق مردم است که آن را تغییر یا لغو کنند و دولتی جدید را تأسیس کنند که بر اصولی بنا شده و

قدرت‌های خود را به شکلی سازماندهی کند که به نظر آنها محتمل‌ترین راه برای تأمین امنیت و خوشبختی آنها باشد.

"رضایت حاکمیت‌شوندگان"، یا توافق ضمنی که شهروندان برای مبادله برخی آزادی‌های بالقوه با نظم و امنیتی که دولت‌ها می‌توانند فراهم کنند، انجام می‌دهند، یک توافق الزام‌آور نیست. این یک گزاره در حال تغییر دائمی است که **همیشه در حال کسب و تأیید مجدد است .** این رضایت از طریق انتخابات، همه‌پرسی‌ها، درخواست‌ها و اعتراضات، پرداخت مالیات، ادای سوگند شهروندی، رعایت قوانین و سایر اشکال مشارکت مدنی آشکار می‌شود. و همانطور که در این فصل نشان داده شد، این رضایت در چگونگی **پذیرش یا مقاومت مردم در برابر فناوری‌های جدید** و همچنین هنجارها و قوانینی که این فناوری‌ها در نهایت الهام‌بخش آن می‌شوند، نیز اتفاق می‌افتد. رشد سریع مالکیت خودرو، به علاوه ایجاد پررونق مشاغل، محصولات و خدمات جدید برای حمایت از آن، سیگنال‌های واضحی از رضایت را فراهم کرد. رشد سریع و تجاری‌سازی اینترنت نیز همینطور بود.

در دنیای جدیدی که اینترنت ایجاد کرد، زمانی که مکانیسم‌های ابراز رضایت – یا مخالفت – بسیار قدرتمندتر و غیرمتمرکزتر از گذشته هستند، این فرآیند جاری از همیشه بیشتر قابل مشاهده است. اکنون رضایت، در هر ثانیه از روز، هرچند غیررسمی، در X.com ، فیس‌بوک، یوتیوب، تیک‌تاک و بی‌نهایت پلتفرم دیگر تأیید، مورد اختلاف و لغو می‌شود. بنابراین، به وضوح نقش کلیدی در چگونگی **پذیرش هوش مصنوعی** در تمام مراحل و در تمام زمینه‌ها ایفا خواهد کرد. این امر تعیین خواهد کرد که چگونه ارزش‌های جمعی و واقعیت‌های عملیاتی به مقررات رسمی تبدیل می‌شوند. این امر شکل خواهد داد که مردم و جوامع چقدر با دقت و مسئولیت‌پذیری قوانین حاصل را دنبال می‌کنند. در عین حال، **"رضایت حاکمیت‌شوندگان" هرگز یک گزاره دوگانه یا مطلق نیست .**مخالفت و کثرت‌گرایی، ایده‌آل‌های بنیادین دموکراسی هستند، همانطور که مذاکره و سازش نیز چنین‌اند. این ایده‌آل‌ها به دموکراسی‌ها کمک می‌کنند

تا پویا، تطبیق‌پذیر و مقاوم باشند. بنابراین، ما هرگز به اجماع ۱۰۰ درصدی در مورد مشروعیت اخلاقی هوش مصنوعی دست نخواهیم یافت، همانطور که به ندرت، اگر هرگز، به چنین یکنواختی اعتقادی در مورد هر سیاست یا فناوری دیگری دست می‌یابیم.

اگر هدف بلندمدت، ادغام ایمن و سازنده هوش مصنوعی در جامعه به جای صرفاً ممنوع کردن آن باشد، پس شهروندان باید نقش فعال و اساسی در **مشروعیت‌بخشی به هوش مصنوعی** ایفا کنند. در این راستا، **"نوآوری بدون مجوز" و "استقرار تکراری "**تنها مکانیسم‌هایی برای افزایش ایمنی و قابلیت‌ها نیستند، بلکه برای **افزایش آگاهی عمومی** درباره نحوه کار این فناوری‌ها و پیامدهای آنها نیز هستند. این مشارکت فعال عمومی است که می‌تواند به هوش مصنوعی مشروعیت ببخشد و آن را به ابزاری برای بهبود زندگی تبدیل کند، نه به وسیله‌ای برای کنترل بی‌حد و حصر.

نتیجه‌گیری

مفهوم **"قانون کد است "**که توسط لارنس لسیگ مطرح شد، امروز بیش از هر زمان دیگری در تحلیل ماهیت حکمرانی در عصر دیجیتال حیاتی به نظر می‌رسد. از روزهای ابتدایی اینترنت، جایی که کد، معماری دسترسی آزاد و ناشناس بودن را تعریف می‌کرد، تا امروز که هوش مصنوعی در تار و پود زندگی روزمره تنیده شده است، این بینش بنیادی اعتبار خود را حفظ کرده است. تحول از یک فضای مجازی غیرمتمرکز به سوی یک معماری کنترل‌محور، عمدتاً از طریق تجاری‌سازی و نیاز به شناسایی کاربران، مسیر را برای ظهور پدیده‌ای به نام **"کنترل کامل "**هموار ساخت.

این کنترل کامل، که در سناریوهای مختلفی از خودروهای مجهز به سیستم‌های تشخیص الکل گرفته تا سیستم‌های مدیریتی در محل کار و خانه نمود می‌یابد، نشان‌دهنده یک تغییر پارادایمی از قوانین سنتی است که بر **رعایت داوطلبانه** متکی بودند، به سمت سازوکارهای معماری‌گونه که **عاملیت انسانی را از معادله حذف می‌کنند** .در حالی که قراردادهای سنتی

فضـایی بـرای قضـاوت، مـذاکره و انعطاف‌پـذیری بـاقی می‌گذاشـتند، **"ضـدقراردادهای "** خوداجراکننده و **قراردادهای هوشمند**، با کدنویسی صریح و شفاف، این نیازها را کاهش داده‌اند. توسعه قراردادهای هوشمند با قابلیت **یادگیری ماشین**، امکان انعطاف‌پذیری و تطبیق‌پذیری بیشتری را فراهم می‌آورد، اما در عین حال، چالش‌های جدیدی نظیر **مسئله تفسیرپذیری** و نگرانی‌هایی در مورد **تبعیض و عدم انصاف** را مطرح می‌سازد.

در نهایت، این فصل بر لزوم بازاندیشی در **"قرارداد اجتماعی** "تأکید می‌کند. همانطور که کنترل‌های معماری‌گونه و هوش مصنوعی بیش از پیش بر زندگی ما حاکم می‌شوند، حفظ اصول **رضـایت حاکمیت‌شـوندگان** و مشـارکت فعـال عمـومی در شـکل‌دهی بـه ایـن فناوری‌هـا، بـرای اطمینـان از عـدالت، شـفافیت و حفـظ آزادی‌های اساسـی، حیـاتی اسـت .**نـوآوری بـدون مجـوز و استقرار تکراری** نه تنها باید به عنوان ابزاری برای پیشرفت فناورانه دیده شوند، بلکه باید به عنوان مسیرهایی برای **افزایش آگاهی عمومی و مشروعیت‌بخشی اجتماعی** به هوش مصنوعی نیز عمل کنند. آینده‌ای که در آن کد و هوش مصنوعی به طور فزاینده‌ای ساختارهای اجتماعی و آزادی‌های فردی را تعریف می‌کنند، نیازمند یک گفتگوی مستمر و مشارکتی است تا اطمینان حاصل شود که این قدرت‌ها به نفع بشریت به کار گرفته می‌شوند و نه بر ضد آن.

نکات کلیدی

- **"قـانون کـد اسـت** :"مفهـوم اصـلی لارنـس لسـیگ کـه کـد را بـه عنـوان یـک عامـل تنظیم‌کننده قدرتمند در کنار قوانین، هنجارها و بازارها معرفی می‌کند.
- **ریشـه‌های اینترنـت** :اینترنـت در ابتـدا بـرای مقاومـت در برابـر کنتـرل مرکـزی و ارتقـاء دسترسی آزاد و ناشناس بودن طراحی شد.
- **تجاری‌سازی و کنترل** :نیازهای تجارت آنلاین به احراز هویت، منجر به تغییر معماری اینترنت به سمت هویت، تمرکزگرایی و کنترل شد.
- **انتقـال کـد بـه جهـان فیزیکـی** :کـد و هـوش مصـنوعی اکنـون بـه طـور گسـترده در

خودروها، لوازم خانگی، زیرساخت‌ها و سایر جنبه‌های زندگی فیزیکی نفوذ کرده‌اند.

- **کنترل کامل** :(Perfect Control) امکان اعمال نظارت و تنظیم‌گری با دقت بی‌نظیر که انتخاب و عاملیت انسانی را محدود یا حذف می‌کند.
- **سیستم‌های تشخیص الکل راننده** :(DADSS) نمونه‌ای از کنترل کامل معماری‌گونه که می‌تواند خودرو را در صورت تشخیص سطح الکل بالا، از حرکت باز دارد.
- **تفاوت قوانین سنتی و کنترل معماری** :قوانین سنتی بر رعایت داوطلبانه متکی هستند، در حالی که کنترل‌های معماری‌گونه (مانند پارکومتر هوشمند) انتخاب را حذف می‌کنند.
- **"ضدقراردادها** :(Uncontracts) "اصطلاح شوشانا زوبوف برای قراردادهای خودکار یک‌جانبه که جایگزین اعتماد و مذاکره می‌شوند.
- **قراردادهای هوشمند** :(Smart Contracts) توافقات خوداجراکننده مبتنی بر بلاکچین که شفافیت، انصاف و مقاومت در برابر دستکاری را فراهم می‌آورند.
- **ادغام یادگیری ماشین در قراردادها** :امکان ایجاد قراردادهای هوشمند پویا و تطبیق‌پذیر که می‌توانند انعطاف‌پذیری قوانین انسانی را تقلید کنند.
- **چالش تفسیرپذیری** :دشواری در درک چگونگی تصمیم‌گیری سیستم‌های یادگیری ماشین در قراردادها، که بر شفافیت و عدالت تأثیر می‌گذارد.
- **کنترل فیزیکی با فناوری دیجیتال** :نمونه‌هایی مانند استفاده MSG Entertainment از تشخیص چهره برای ممنوعیت ورود، نشان‌دهنده توانایی کنترل دقیق در فضاهای فیزیکی است.
- **رضایت حاکمیت‌شوندگان** :اصل بنیادین برای مشروعیت‌بخشی به دولت‌ها و فناوری‌ها، به ویژه در مواجهه با سیستم‌های هوش مصنوعی که بر انتخاب‌های انسانی تأثیر می‌گذارند.

- **نقش عمومی در حکمرانی هوش مصنوعی** :مشارکت فعال شهروندان برای ادغام ایمن و سازنده هوش مصنوعی در جامعه و افزایش آگاهی درباره پیامدهای آن ضروری است.

سؤالات تفکربرانگیز

- چگونه می‌توان تعادلی بین مزایای کارایی و امنیت ناشی از **"کنترل کامل "** و حفظ عاملیت فردی، حریم خصوصی و آزادی‌های اساسی در جامعه‌ای مبتنی بر هوش مصنوعی ایجاد کرد؟
- با توجه به اینکه سیستم‌های هوش مصنوعی می‌توانند به طور فزاینده‌ای بر انتخاب‌های انسانی تأثیر بگذارند، چه چارچوب‌های قانونی و اخلاقی جدیدی برای حکمرانی بر این **"عاملیت ماشینی "**مورد نیاز است؟
- چگونه می‌توان اصل **"رضایت حاکمیت‌شوندگان "**را به طور عملی در دنیایی که به طور فزاینده‌ای خودکار و مبتنی بر هوش مصنوعی است، به کار بست و حفظ کرد؟
- پیامدهای بلندمدت اجتماعی ناشی از برون‌سپاری قضاوت انسانی، مذاکره، و همدلی به کد و هوش مصنوعی چیست؟ آیا این امر می‌تواند منجر به کاهش ظرفیت‌های اخلاقی و اجتماعی انسان شود؟
- چگونه می‌توان شفافیت و **تفسیرپذیری** را در قراردادها و سیستم‌های حکمرانی پیچیده مبتنی بر هوش مصنوعی تضمین کرد تا اصول عدالت و انصاف رعایت شوند و افراد بتوانند نتایج ناعادلانه را به چالش بکشند؟

۳۰ جمله کلیدی

۱. گوگل در سال ۱۹۹۸ مأموریت خود را "سازماندهی اطلاعات جهان و قابل دسترس و مفید ساختن آن برای همگان" اعلام کرد.

۲. اینترنت در ابتدا برای مقاومت در برابر نقاط شکست و اقتدار مرکزی طراحی شده بود

و تاب‌آوری شبکه را بر کنترل مرکزی اولویت می‌داد.

۳. جان گیلمور اظهار داشت: "نت سانسور را به مثابه آسیب تفسیر می‌کند و مسیر خود را از کنار آن تغییر می‌دهد".

۴. لارنس لسیگ در کتاب "کد" نشان داد که تجاری‌سازی اینترنت به سمت هویت، تمرکزگرایی و معماری‌های کنترل در حال حرکت است.

۵. لسیگ چهار عامل تنظیم‌کننده رفتار انسان را قوانین، هنجارها، بازارها و معماری (کد) معرفی کرد.

۶. او استدلال کرد که در اینترنت، "کد قانون است" و شرایط تجربه زندگی در فضای مجازی را تعیین می‌کند.

۷. تجارت آنلاین، با نیاز به شناسه‌های شخصی، یک "اسب تروا" بود که معماری اینترنت را تغییر داد.

۸. "دست نامرئی، از طریق تجارت، در حال ساخت معماری‌ای است که کنترل را به کمال می‌رساند".

۹. امروز، اینترنت دیگر جهانی جداگانه نیست، بلکه در همه چیز نفوذ کرده و به بخشی جدایی‌ناپذیر از دنیای واقعی تبدیل شده است.

۱۰. دستگاه‌های هوشمند با قابلیت عاملیت‌محور، تصمیم‌گیری می‌کنند و بر انتخاب‌ها و تعهدات انسان‌ها تأثیر می‌گذارند.

۱۱. سیستم فدرالی تشخیص الکل راننده (DADSS) در خودروهای جدید، مثالی از کنترل معماری‌گونه است.

۱۲. این سیستم‌ها از حسگرهای لمسی و تنفسی برای اندازه‌گیری غیرفعال سطح الکل خون راننده استفاده می‌کنند.

۱۳. هوش مصنوعی در خودروها می‌تواند از دوربین و حسگرها برای تشخیص راننده و

اعمال تصمیمات خودکار استفاده کند.

۱۴. سناریوی "ناوی‌تار" که خودرو را به دلیل سطح الکل بالا روشن نمی‌کند، نشان‌دهنده حذف انتخاب انسانی است.

۱۵. کنگره ایالات متحده در سال ۲۰۲۱ قوانینی برای اجباری کردن سیستم‌های تشخیص الکل در خودروها تصویب کرده است.

۱۶. "کنترل کامل" لسیگ به پدیده‌ای اشاره دارد که در آن عاملیت فردی از طریق معماری فناوری محدود می‌شود.

۱۷. قوانین سنتی بر "انتخاب" و "رعایت داوطلبانه" متکی هستند، در حالی که کنترل‌های کد بر "پیش‌گیری" عمل می‌کنند.

۱۸. یک پارکومتر حسگردار می‌تواند جریمه را به طور خودکار از حساب شما کسر کند، بدون اینکه انتخابی برای شما باقی بگذارد.

۱۹. "اجرای قانون می‌تواند آنقدر ارزان شود که حتی نیازی به اندازه‌گیری آن نباشد".

۲۰. فناوری‌ها می‌توانند آزادی‌ای را که "صرفاً با ناکارآمدی انجام کاری متفاوت" حفظ شده بود، جابجا کنند.

۲۱. هوش مصنوعی امکان کنترل کامل را گسترش می‌دهد، به طوری که سرعت غیرمجاز ممکن است دیگر غیرممکن شود یا تخلفات با تحمل صفر روبرو شوند.

۲۲. "ضدقراردادها"ی شوشانا زوبوف، رویه‌های خودکار را جایگزین وعده‌ها، گفتگو و اعتماد در قراردادها می‌کنند.

۲۳. قراردادهای هوشمند مبتنی بر بلاکچین، خوداجراکننده، شفاف و مقاوم در برابر دستکاری هستند.

۲۴. این قراردادها از طریق کد، شرایط توافق را تعیین و عملکرد را تأیید می‌کنند و

پرداخت‌ها را به طور خودکار آزاد می‌سازند.

۲۵. ادغام الگوریتم‌های یادگیری ماشین می‌تواند قراردادهای هوشمند را پویا و تطبیق‌پذیر کند و انعطاف‌پذیری قوانین انسانی را تقلید نماید.

۲۶. چالش اصلی در قراردادهای هوشمند مبتنی بر هوش مصنوعی، "تفسیرپذیری" است، یعنی درک چگونگی تصمیم‌گیری سیستم.

۲۷. در زمینه‌های حقوقی، تفسیرپذیری برای شفافیت و اجرای برابر قانون از اهمیت بالایی برخوردار است.

۲۸. لسیگ استدلال کرد که ما باید به عنوان "اعضا" و نه صرفاً "مشتریان" با فضای مجازی و معماری‌های آن ارتباط برقرار کنیم.

۲۹. استفاده از فناوری تشخیص چهره در فضاهای فیزیکی) مانند MSG Entertainment) نشان‌دهنده گسترش کنترل معماری‌گونه به دنیای واقعی است.

۳۰. "رضایت حاکمیت‌شوندگان" اصل حیاتی برای مشروعیت‌بخشی به هوش مصنوعی و فناوری‌های جدید است و مستلزم مشارکت فعال عمومی است.

فصل ۸

خودمختاری شبکه‌ای و آزادی پایدار: بازتعریف آزادی در عصر فناوری

"در عصری که فناوری به طور فزاینده‌ای توانایی‌های فردی ما را تقویت می‌کند، آزادی واقعی نه در خودمختاری مطلق، بلکه در مسیریابی هوشمندانه یک «خودمختاری شبکه‌ای» نهفته است؛ جایی که آزادی‌های فردی هم از طریق منفعت جمعی و هم در بستر هنجارهای در حال تحول و مقررات هوشمند، فعال و محدود می‌شوند — ".یک پژوهشگر معاصر در حوزه فناوری و جامعه

اهداف اصلی فصل

- **آشنایی با مفهوم آزادی:** بررسی مفهوم پویای آزادی در تقاطع فناوری، مقررات و جامعه.
- **واکاوی خودمختاری فردی:** تحلیل تکامل درک ما از خودمختاری فردی در مواجهه با پیشرفت‌های **فناورانه**.
- **بررسی نقش زیرساخت‌ها:** تبیین نقش حیاتی زیرساخت‌های مشترک و اقدامات هماهنگ در تقویت آزادی‌های فردی و اجتماعی.
- **توجه به تعادل آزادی و امنیت:** پرداختن به چالش‌ها و فرصت‌های فناوری‌های نوظهور مانند هوش مصنوعی برای تعادل میان آزادی و امنیت.
- **تأکید بر اجماع ملی:** برجسته کردن اهمیت اجماع ملی و اراده جمعی برای دستیابی به اهداف بزرگ فناورانه و اجتماعی.

چکیده

این فصل به بررسی مفهوم پویای آزادی می‌پردازد و استدلال می‌کند که آزادی نه یک اصل ثابت و تغییرناپذیر، بلکه مفهومی در حال تحول است که به شدت تحت تأثیر پیشرفت‌های تکنولوژیکی، چارچوب‌های نظارتی، و نیازهای جمعی قرار دارد. با الهام از تاریخچه‌ی اتومبیل‌رانی، این فصل نشان می‌دهد که چگونه نوآوری‌های اولیه منجر به افزایش خودمختاری فردی شد، اما به زودی با نیاز به مقررات و زیرساخت‌های مشترک برای افزایش ایمنی، کارایی و دسترسی گسترده‌تر همراه گشت. در عصر هوش مصنوعی و وسایل نقلیه خودران، این تعادل میان آزادی‌های جدید (مانند نوشیدن هنگام مسافر بودن) و محدودیت‌های نوظهور (مانند کنترل مسیر یا سرعت) بیش از پیش مشهود است. همچنین، فصل به بررسی نمونه‌های تاریخی مانند اختراع چاپ و گسترش آن می‌پردازد تا نشان دهد که چگونه فناوری‌های انقلابی، با وجود گسترش اولیه آزادی بیان، به تدریج نیازمند چارچوب‌های نظارتی جدیدی برای مدیریت ریسک‌ها و چالش‌های اجتماعی شدند. این فصل با بررسی واکنش کره جنوبی به کووید-۱۹، که در آن آزادی حرکت گسترده با جمع‌آوری داده‌های تهاجمی و شفافیت عمومی ترکیب شد، و همچنین توسعه سیستم بزرگراهی بین‌ایالتی آمریکا، که نمونه‌ای از زیرساخت‌های جمعی برای افزایش خودمختاری فردی است، به این نتیجه می‌رسد که آزادی واقعی اغلب در یک «خودمختاری شبکه‌ای» یافت می‌شود. این مفهوم بر اهمیت همکاری، اعتماد عمومی، و سرمایه‌گذاری‌های مشترک در ایجاد جوامع پویا و مقاوم تأکید دارد، در حالی که چالش‌های فرهنگی و سیاسی در پذیرش چنین رویکردهایی را نیز مورد بحث قرار می‌دهد.

مقدمه

مفهوم آزادی، از دیرباز یکی از محوری‌ترین آرمان‌های بشری بوده است. با این حال، درک ما از آزادی هرگز ثابت نبوده و همواره در مواجهه با دگرگونی‌های فناورانه، اجتماعی و سیاسی دستخوش تغییر و تحول شده است. در دوران کنونی، با ورود به عصری که هوش مصنوعی و سیستم‌های خودران در حال بازتعریف تعاملات روزمره ما هستند، درک ما از خودمختاری فردی و آزادی‌های اجتماعی نیز وارد مرحله‌ای نوین شده است. این فصل به بررسی این پویایی‌ها می‌پردازد و استدلال می‌کند که آزادی حقیقی، اغلب نه در نبود کامل محدودیت‌ها، بلکه در چارچوب یک «خودمختاری شبکه‌ای» شکل می‌گیرد؛ سیستمی که در آن نوآوری‌های فردی با اقدامات جمعی و مقررات هوشمندانه در هم تنیده می‌شوند تا مجموعه‌ای از آزادی‌های پایدار را برای جامعه به ارمغان آورند.

تاریخ بشر سرشار از نمونه‌هایی است که نشان می‌دهند چگونه پیشرفت‌های تکنولوژیک، در ابتدا به عنوان ابزاری برای افزایش توانمندی‌های فردی و گسترش آزادی‌های شخصی ظاهر می‌شوند. از اختراع ماشین چاپ که آزادی بیان را متحول کرد تا خودرو که آزادی حرکت را بی‌سابقه ساخت. اما هر یک از این نوآوری‌ها، به موازات فرصت‌هایی که خلق کردند، چالش‌های جدیدی نیز به همراه داشتند که نیازمند واکنش‌های جمعی، وضع مقررات، و سرمایه‌گذاری در زیرساخت‌های مشترک بودند. این واکنش‌ها، در نگاه اول ممکن است به مثابه محدود کردن آزادی به نظر برسند، اما در تحلیل عمیق‌تر، اغلب به بازتعریف و گسترش آزادی در ابعاد وسیع‌تر منجر شده‌اند.

در این فصل، ما با بهره‌گیری از مثال‌های متنوعی از گذشته و حال، از جمله تکامل صنعت خودرو، ظهور هوش مصنوعی، و رویکردهای نوین بهداشت عمومی، نشان خواهیم داد که چگونه این تعامل پیچیده میان فرد و جامعه، میان نوآوری و مقررات، و میان خودمختاری و وابستگی متقابل، به شکل‌گیری یک درک مدرن از آزادی انجامیده است. این درک بر اهمیت نه تنها آزادی برای عمل، بلکه آزادی از خشونت، بی‌نظمی، و ریسک‌های بی‌حد و حصر تأکید دارد. هدف

نهایی این فصل، ارائه چارچوبی برای فهم چگونگی همزیستی و هم‌افزایی خودمختاری فردی در یک بستر شبکه‌ای، که نه تنها به افراد قدرت می‌بخشد، بلکه به جامعه‌ای پویا و مقاوم برای مواجهه با چالش‌های آینده کمک می‌کند.

دگرگونی خودمختاری فردی در عصر وسایل نقلیه خودران

ظهور وسایل نقلیه خودران، نویدبخش دگرگونی‌های بنیادین در مفهوم خودمختاری فردی در زمینه حمل‌ونقل است. در حالی که این فناوری ممکن است آزادی‌های جدیدی را به ارمغان آورد، مانند امکان نوشیدن در حین مسافر بودن در خودرویی که خودش رانندگی می‌کند، به احتمال زیاد با محدودیت‌های جدیدی نیز همراه خواهد بود. برای مثال، روزهای فشردن پدال گاز تا نهایت سرعت، حتی در خلوت‌ترین جاده‌ها، ممکن است به پایان رسد. همچنین، توانایی انتخاب مسیر دقیق در یک وسیله نقلیه خودران، به ویژه اگر تاکسی یا خودروی اجاره‌ای باشد، احتمالاً محدود خواهد شد.

وسایل نقلیه خودران، که برای بهینه‌سازی کارایی، ایمنی، و جریان ترافیک طراحی شده‌اند، احتمالاً تصورات سنتی از خودمختاری فردی در زمینه اتومبیل‌رانی را به چالش خواهند کشید. با این حال، تاریخچه خودرو، به ویژه در آمریکا، نشان می‌دهد که این پدیده، در عین اینکه حکایتی حماسی از توانمندسازی بسیار فردی از طریق نوآوری بدون مجوز بوده، به همان اندازه نیز حماسه‌ای از اقدام جمعی و آزادی از طریق مقررات‌گذاری است. رشد اتومبیل‌رانی در آمریکا یک داستان کلاسیک از اثرات شبکه‌ای بود. هر مدل T جدید، به نفع پارک‌وی‌های شهری وسیع‌تر و جاده‌های روستایی هموارتر بود و هر مایل جاده آسفالت، ارزش هر مدل T را افزایش می‌داد. این مقیاس‌پذیری سریع، کلید ایجاد اجماع گسترده ملی بود که منافع خودروها به مراتب بیشتر از خطرات، تغییرات فرهنگی و سرمایه‌گذاری جمعی در زیرساخت‌ها بود.

آزادی چندوجهی: تعادل بین اختیار و ایمنی

با افزایش تعداد رانندگان از هزاران به میلیون‌ها نفر، یک شبکه از مقررات، که به صورت

محتاطانه اما پیوسته در حال گسترش بود، به افزایش فوق‌العاده آزادی شخصی و خودمختاری فردی کمک کرد. این امر در دورانی که لابی‌گران شرکت‌ها و نهادهای دولتی بر کاهش دولت تأکید دارند، ممکن است متناقض به نظر برسد. اما آزادی چندوجهی است و تکرارهای مختلف آن اغلب در تنش با یکدیگر قرار دارند. از یک سو، ما به شدت خواهان آزادی بیان، نوآوری، و حرکت با سرعت‌های بالا هستیم. از سوی دیگر، خواهان آزادی از خشونت، فساد، تبعیض و خطرات بیش از حد هستیم. این یک تعادل ظریف است.

رویکرد منعطف‌تر به اتومبیل‌های احتراق داخلی در ابتدا یکی از دلایلی بود که آنها توانستند از مرحله نمونه اولیه فراتر روند، بر خلاف وسایل نقلیه شخصی بخارپز دهه ۱۸۶۰ و ۱۸۷۰. اما مقررات نیز به کاهش خطرات و حتی ساده‌سازی مهارت‌های لازم برای رانندگی کمک کرد و آن را به شکلی واقعاً مقیاس‌پذیر از خودمختاری و عاملیت فردی تبدیل ساخت. وقتی ایالت‌ها در اوایل قرن بیستم شروع به اجباری کردن گواهینامه رانندگی کردند، صلاحیت رانندگی پایه افزایش یافت. قانون کمک به جاده‌های فدرال در سال ۱۹۱۶، توسعه اصول طراحی استاندارد جاده‌ها را تسریع کرد که هم ایمنی و هم سرعت رانندگی را افزایش داد. چراغ‌های راهنمایی، علائم توقف و خط‌کشی‌های جاده‌ای، جریان ترافیک را منظم‌تر و قابل پیش‌بینی‌تر کردند و به رانندگان امکان دادند سریع‌تر، انعطاف‌پذیرتر و با اطمینان بیشتری رانندگی کنند. با گذشت زمان، تثلیث مقدس نوآوری تکنولوژیک، استقرار گسترده و مقررات مبتنی بر شواهد، شاهکارهای خارق‌العاده‌ای از جابجایی را عادی ساخت.

آزادی در حال تغییر: نقش فناوری و نظارت

آزادی معمولاً با نمادهایی ثابت و تغییرناپذیر به تصویر کشیده می‌شود. با این حال، این نمادها اغلب چگونگی رابطه‌ای و در حال تغییر بودن آزادی را پنهان می‌کنند. آزادی یک اصل ثابت نیست که مستقل از بستر تعریف خود وجود داشته باشد؛ بلکه همواره در حال تغییر و تحول است و چگونگی درک ما از آن، غالباً با آنچه فناوری‌های روز امکان‌پذیر می‌سازند یا نمی‌سازند،

ارتباط تنگاتنگی دارد. برخی از آزادی‌ها را به این دلیل داریم که تنظیم آنها دشوار است (مانند آزادی کپی کردن نوارهای کاست). از آزادی‌های دیگر لذت می‌بریم زیرا اعمال آنها دشوار است)قبل ازCRISPR ، قوانین کمی در مورد مهندسی ژنتیک وجود داشت .(وقتی فناوری امکان انجام یک شاهکار فوق‌انسانی را فراهم می‌کند، مانند رانندگی با سرعت ۱۵۰ مایل در ساعت یا از بین بردن تومورها با پرتودرمانی، معمولاً به نوعی آن را تنظیم می‌کنیم.

برای روشن شدن این نکته، سفر ۲۸ ساعته از اسپرینگفیلد، ایلینوی، تا ساکرامنتو، کالیفرنیا، را در نظر بگیرید. این سفر در دوران مدرن، با رعایت محدودیت‌های سرعت، علائم راهنمایی، قوانین کمربند ایمنی، و ممنوعیت رانندگی در حالت مستی، کاملاً امکان‌پذیر است. در طول این سفر، گوشی هوشمند شما مسیرتان را ردیابی می‌کند، ارائه‌دهنده خدمات موبایل شما مکان توقف‌هایتان را ثبت می‌کند، و شرکت کارت اعتباری شما جزئیات خریدتان را ضبط می‌کند. دوربین‌های امنیتی در فروشگاه‌ها شما را زیر نظر دارند و حتی در هتلی به ظاهر "یاغی"، باید کارت شناسایی معتبر ارائه دهید. این سطح از نظارت و مقررات، ممکن است زندگی در سرزمین آزادگان را به یک اودیسه بی‌پایان از بوروکراسی خفیف و نظارت اتفاقی تبدیل کند.

در مقایسه، حزب دونر، که در سال ۱۸۴۶ همین مسیر را آغاز کردند، از هیچ یک از این محدودیت‌ها یا نظارت‌ها برخوردار نبودند. آنها آزاد بودند با هر سرعتی که گاوهایشان می‌توانستند حرکت کنند (حدود ۱۵ مایل در روز) بر روی زمین‌های ناهموار پیش بروند. هیچ جریمه‌ای برای نوشیدن مشروب در راه وجود نداشت و می‌توانستند هر کجا که مناسب می‌دیدند، هر چقدر که می‌خواستند استراحت کنند. هیچ دوربین امنیتی یا ردیاب GPS فعالیت‌هایشان را ضبط نمی‌کرد. آنها چقدر از نظارت‌های مزاحم و دخالت‌های تجاری رها بودند. چقدر خودگردان و مسئول زندگی خودشان باید احساس می‌کردند. اما این آزادی بی‌حد و حصر، در نهایت به فاجعه‌ای انجامید که تنها ۴۰ نفر از ۸۷ نفر جان سالم به در بردند. این مقایسه نشان می‌دهد که ما شهروندان قرن بیست و یکم، با کمربند ایمنی، محدودیت‌های سرعت و جاده‌های کاملاً تحت نظارت، که می‌توانیم این سفر را در یک و نیم روز انجام دهیم، به مراتب آزادتر از آن فردگرایان

سرسخت در حزب دونر هستیم که برای بقا مجبور به اشکال نهایی کمونیسم شدند. این آزادی خارق‌العاده، مدیون نوآوری تکنولوژیک، قوانین و مقررات، و منابع مشترک اجرا و مدیریت شده توسط سطوح مختلف دولتی است.

درس‌هایی از تاریخ: چاپ و مقررات‌گذاری بیان

همانطور که فناوری‌های جدید در جوامع منتشر می‌شوند، مقررات و هنجارهای جدیدی نیز به دنبال آنها می‌آیند و این تغییرات بر درک در حال تحول ما از آزادی تأثیر می‌گذارند. قبل از اختراع ماشین چاپ در اروپا، تولید کتاب عمدتاً در اختیار کلیسای کاتولیک و دانشگاه‌ها بود. این نهادها استانداردهای خاص خود را بر کتاب‌هایی که تولید می‌کردند اعمال می‌کردند و تا حدی، این استانداردها به عنوان هنجارهای فرهنگی گسترده‌تری عمل می‌کردند. اما قوانین رسمی تصویب شده توسط دولت‌های محلی عمدتاً وجود نداشت، زیرا نیازی به آنها احساس نمی‌شد.

اما پس از اختراع ماشین چاپ توسط یوهانس گوتنبرگ در آلمان در دهه ۱۴۴۰، همه چیز تغییر کرد. در سال ۱۴۸۶، ونیز سانسور رسمی را معرفی کرد و چاپ کتاب‌ها را منوط به تأیید قبلی ساخت. در اواسط دهه ۱۵۰۰، کلیسای کاتولیک فهرستی از کتاب‌های ممنوعه را منتشر کرد. امپراتوری مقدس روم نیز قوانین خود را معرفی کرد. انگلستان در سال ۱۵۵۷، شرکت استیشنرز را تحت منشور سلطنتی رسمی کرد و انحصار صنعت چاپ را به آن اعطا کرد. سرانجام، قوانین رسمی و مکانیسم‌های سانسور آنقدر رایج شدند که مفهوم آزادی بیان و حق ذاتی افراد برای ابراز وجود، اهمیت فرهنگی فزاینده‌ای پیدا کرد. قوانینی که برای حفاظت از بیان در برابر سانسور دولتی طراحی شده بودند، مانند قانون صدور مجوز انگلستان در سال ۱۶۶۲ و، تقریباً ۱۳۰ سال بعد، اولین متمم قانون اساسی آمریکا، شروع به ظهور کردند. این امر مانع از ایجاد قوانین و مقررات اضافی که طیف وسیعی از گفتار و بیان را ممنوع می‌کردند، نشد—از هرزه‌نگاری و بی‌شرمی گرفته تا نقض حق چاپ و تبلیغات دروغین، تا تهدیدات واقعی، تحریک به اقدامات غیرقانونی قریب‌الوقوع، و «کلمات جنگی». در مجموع، اکنون در ایالات متحده قوانین بسیار بیشتری برای

ممنوعیت گفتار وجود دارد تا در اروپا در زمانی که اختراع گوتنبرگ جهان را متحول کرد. اما آیا کسی استدلال می‌کند که ماشین چاپ به دلیل محیط نظارتی جدیدی که الهام‌بخش آن بود، آزادی بیان کمتری ایجاد کرد؟

هوش مصنوعی: فرصت‌ها و چالش‌های خودمختاری و امنیت

هوش مصنوعی نیز به نوبه خود، بر مفاهیم آزادی تأثیر خواهد گذاشت. قدرت پردازش موازی گسترده آن می‌تواند ظرفیت ما را برای اقدام در مورد مشکلات پیچیده، از قید معماری عصبی کند خودمان رها کند. اما درست همانطور که خودروها عمل کردند، هوش مصنوعی نیز احتمالاً الهام‌بخش اشکال جدیدی از مقررات خواهد شد—نه فقط در مورد خودش، بلکه در مورد استفاده ما از آن. در واقع، حتی ممکن است برای اینکه بتوانیم در دنیایی که هوش مصنوعی در آن وجود دارد زندگی کنیم، به مقررات جدیدی نیاز داشته باشد.

همانطور که مصطفی سلیمان در کتاب «موج در راه است» می‌نویسد: «دموکراتیک کردن دسترسی به هوش مصنوعی بسیار توانمند، لزوماً به معنای دموکراتیک کردن ریسک است.» دسترسی فزاینده به دستگاه‌های دو منظوره نسبتاً ارزان اما قدرتمندتر مانند پهپادها، ربات‌ها، و وسایل نقلیه خودران، به بازیگران بدخواه قدرت می‌دهد تا با نیروی نامتقارن بی‌سابقه‌ای ضربه بزنند. وقتی دولت‌ها دیگر تنها نهادهایی نیستند که قادر به انجام حملات در سطح ملی هستند، یک توجیه آشکار برای سطوح جدیدی از مقررات و نظارت برای کاهش احتمال چنین اتفاقاتی وجود دارد. دستور اجرایی هوش مصنوعی که رئیس جمهور جو بایدن در اکتبر ۲۰۲۳ امضا کرد، پیش‌بینی کرد که بخش عمده‌ای از این مقررات مستقیماً بر صنعت هوش مصنوعی متمرکز خواهد شد. اما تلاش‌های امنیتی جدید ممکن است اشکال گسترده‌تری نیز به خود بگیرند. به همان شیوه‌ای که رانندگان فردی برای رانندگی با وسیله نقلیه موتوری باید گواهینامه دریافت کنند، شاید شما نیز برای دسترسی به برخی مدل‌های هوش مصنوعی بسیار توانمند نیاز به دریافت مجوز هوش مصنوعی داشته باشید.

برای افزایش امنیت آنلاین به طور کلی، پروتکل‌های جدید تأیید هویت که نیاز به کارت‌های شناسایی رمزنگاری‌شده و/یا داده‌های بیومتریک دارند، ممکن است به طور گسترده‌تری اجرا شوند. مکانیسم‌های احراز هویت چندوجهی، مانند ترکیب اثر انگشت، تشخیص صدا، و تحلیل رفتاری، می‌توانند به استاندارد جدیدی برای دسترسی به اطلاعات حساس یا پلتفرم‌های امنیتی بالا تبدیل شوند. سیستم‌های تشخیص چهره در نقاط بازرسی امنیتی ممکن است فراتر از فرودگاه‌ها به طیف وسیع‌تری از مکان‌های عمومی گسترش یابند. حتی با وجود تمام موارد استفاده مثبت برای هوش مصنوعی دست‌یافتنی و دموکراتیک در تمام عرصه‌های تلاش بشری، شک و تردید در مورد ارزش خالص هوش مصنوعی، یک پاسخ قابل درک به تغییراتی است که هوش مصنوعی به طور گسترده توزیع شده می‌تواند طلب کند. درخواست از شهروندان برای تن دادن به الزامات جدید شناسایی و تدابیر امنیتی جدید، تنها برای سازگاری با ماشین‌هایی که آنها را تهدیدی برای خودمختاری خود می‌دانند، ممکن است به نظر برخی یک معامله پوچ باشد.

خودمختاری شبکه‌ای برای خیر جمعی

در این دیدگاه، هوش مصنوعی می‌تواند به عنوان ابزاری برای تقویت اراده فردی عمل کند. در آینده نزدیک، میلیون‌ها نفر بر گروهی از دستیاران هوش مصنوعی نظارت خواهند داشت که قادرند هر خواسته‌ای را به صورت دقیق و آنی برآورده کنند. این قدرت جدید برای ارضای هر هوس ما، ممکن است ظرفیت ما را برای مشارکت به عنوان اعضای با حسن نیت در یک قرارداد اجتماعی مشترک از بین ببرد. اما می‌توان به جان استوارت میل، فیلسوف عصر بخار، برای یک استدلال متقابل خوش‌بینانه‌تر رجوع کرد. میل تأکید داشت که آزادی فردی نه تنها به عنوان یک هدف فی‌نفسه، بلکه به دلیل چگونگی کمک آن به رفاه کلی جامعه، ضروری است. او استدلال می‌کرد که تأکید بر خودمختاری و خودتعیین‌گری فردی، جامعه‌ای متنوع و پر جنب و جوش را پرورش می‌دهد که در آن افراد می‌توانند به نهایت پتانسیل خود دست یابند و به طور مولد به خیر عمومی کمک کنند.

آنچه میل درک کرد این بود که افراد موفق منجر به جوامع موفق می‌شوند. در اصل، آنچه او برای آن استدلال می‌کرد، نوعی «خودمختاری شبکه‌ای» بود. قطعات به صورت جداگانه قوی هستند. با هم کار کردن، آنها حتی قوی‌تر می‌شوند. با قدرتمند شدن توسط هوش مصنوعی، و حس اینکه چگونه می‌تواند به ما به صورت جمعی و همچنین فردی کمک کند، شاید تمایلات جدیدی به یافتن هدف مشترک پیدا کنیم. این دینامیک که افراد موفق منجر به جوامع موفق می‌شوند، تنها نیمی از یک حلقه فضیلت‌مند است. جوامع موفق نیز به تقویت افراد کمک می‌کنند.

کره جنوبی: نمونه‌ای از تعادل هوش مصنوعی و بهداشت عمومی

پاسخ کره جنوبی به کووید-۱۹، نمونه بارزی از چگونگی دستیابی به نتایج مطلوب در یک جهان شبکه‌ای و تکنولوژیکی است، جایی که نه تنها قابلیت‌های محاسباتی و قدرت نوآوری، بلکه انسجام اجتماعی نیز اهمیت دارد. این کشور با وجود یک رویداد فوق‌گسترانده در اوایل همه‌گیری و تراکم بالای جمعیت در سئول، به یک مدل موفق برای پاسخ به همه‌گیری تبدیل شد. راز موفقیت کره جنوبی، "اقدام سریع، آزمایش گسترده و ردیابی تماس، و حمایت حیاتی از شهروندان" بود. کره جنوبی، به جای اجرای قرنطینه‌های سراسری، دستورات در خانه ماندن یا تعطیلی کسب‌وکارها، بر آزادی گسترده حرکت، جمع‌آوری داده‌های تهاجمی و اطلاعات مشترک به جای حریم خصوصی فردی تأکید کرد.

دولت کره جنوبی با استفاده از قدرت‌های قانونی که پس از شیوع سندروم تنفسی خاورمیانه (MERS) در سال ۲۰۱۵ توسعه یافته بود، به داده‌های GPS موبایل، تراکنش‌های کارت اعتباری، سوابق سفر و موارد دیگر دسترسی داشت. این اطلاعات برای ردیابی حرکات افراد مثبت شده به کووید و سپس عمومی کردن مسیرهای آنها برای دیگرانی که ممکن بود با آنها در تماس بوده‌اند، استفاده می‌شد. برای اینکه ردیابی تماس موثر باشد، مقامات بهداشتی ابتدا باید می‌دانستند چه کسی کووید دارد. بنابراین، ظرف یک هفته از اولین مورد شناسایی شده، دولت کره جنوبی

تأیید سریع بیش از بیست شرکت پزشکی را که برای توسعه آزمایش تشخیصی گردهم آورده بود، وعده داد. به نوبه خود، شرکتی به نام سی‌ژن از هوش مصنوعی برای تجزیه و تحلیل توالی‌های ژنتیکی و خودکارسازی فرآیند طراحی استفاده کرد و یک آزمایش موثر را تنها در سه هفته توسعه داد. سپس کره جنوبی روزانه هزاران نفر را آزمایش کرد. اگر فردی مثبت می‌شد، مقامات بهداشتی از تحلیل‌های هوش مصنوعی برای تولید جزئیات حرکات آن فرد در یک دوره چند روزه در حدود یک دقیقه استفاده می‌کردند، که سپس از شناسه‌های شخصی پاک شده و از طریق پیامک، پست‌های وبلاگ و سایر اشکال ارتباطی به اشتراک گذاشته می‌شد.

اگرچه همیشه هموار پیش نرفت، اما مردم کره جنوبی این رویکرد را پذیرفتند. «خشم عمومی تقریباً وجود نداشت.» بخشی از این پذیرش عمومی به تجربه اخیر کشور با شیوع MERS برمی‌گشت. اما بخش زیادی از استقبال عمومی به تمایل دولت به تعهد به «نسخه‌ای رادیکالی شفاف از ردیابی افراد که تحت نظارت عمومی و همراه با تدابیر حفاظتی قانونی سختگیرانه علیه سوءاستفاده بود» نسبت داده شد. به عبارت دیگر، دولت داده‌ها را به صورت مخفیانه جمع‌آوری نمی‌کرد و از آنچه جمع‌آوری می‌کرد به صورت نامتقارن استفاده نمی‌نمود. در عوض، از مردم می‌خواست تا در یک مشارکت متقابل شفافیت و «دانش بزرگ» شرکت کنند. این چرخه شفافیت، مشارکت عمومی و نتایج قابل مشاهده، اعتماد مدنی به رویکرد دولت را تقویت کرد، که به نوبه خود منجر به حس هدف مشترک و رعایت اقدامات بهداشت عمومی برای مبارزه با ویروس شد.

در سناریوهای همه‌گیری آینده، برخی کشورها احتمالاً از سیستم‌های هوش مصنوعی پیچیده برای حفظ فعالیت‌های روزمره تا حد امکان و در عین حال کاهش نرخ انتقال استفاده خواهند کرد. مقامات بهداشت عمومی می‌توانند با استفاده از الگوریتم‌هایی که بر روی مجموعه‌های داده عظیم آموزش دیده‌اند، همراه با تکنیک‌هایی مانند تصویربرداری حرارتی و نظارت بر نرخ تنفس، سیستم‌هایی را برای اسکن افراد برای چندین شاخص سلامت در زمان واقعی بسازند. چنین سیستمی می‌تواند به طور یکپارچه در فضاهای عمومی ادغام شود. اما توسعه چنین سیستم‌های قابل اعتمادی چالش‌های بسیاری دارد. دقت حسگرها و دوربین‌های امروزی

متفاوت است و نیاز به کالیبراسیون پیشرفته برای اندازه‌گیری جمعیت‌های متنوع در شرایط محیطی مختلف دارند. الگوریتم‌ها باید در گروه‌های جمعیتی مختلف و در محیط‌های گوناگون به طور ثابت عمل کنند. نصب و نگهداری میلیون‌ها حسگر و دوربین جدید، همراه با شبکه‌های داده پرسرعت، بسیار پرهزینه خواهد بود. از نظر قانونی، دولت‌ها باید قوانین پیچیده حریم خصوصی را رعایت کرده و احتمالاً قوانین جدیدی برای نظارت بر سلامت در این مقیاس تدوین کنند. از نظر اخلاقی، چنین سیستمی سؤالاتی در مورد رضایت، مالکیت داده‌ها و پتانسیل تبعیض یا سوءاستفاده از اطلاعات سلامت فراتر از اهداف بهداشت عمومی مورد نظر را مطرح می‌کند.

سیستم بزرگراهی بین‌ایالتی: زیرساخت‌های جمعی برای خودمختاری فردی

در بسیاری جهات، یک سیستم نظارتی بهداشتی مشابه سیستم دفاع موشکی بالستیک آمریکا یا سیستم فضای هوایی ملی است—زیرساخت‌های امنیتی پیچیده‌ای که برای محافظت از کل کشور در برابر نوع خاصی از تهدید طراحی شده‌اند. ایالات متحده، با داشتن شرکت‌هایی مانند گوگل، مایکروسافت، فیس‌بوک، OpenAI، اپل و آمازون به عنوان دارایی‌های استراتژیک ملی در پردازش داده‌های بزرگ و توسعه هوش مصنوعی، و همچنین مؤسسات آکادمیک و نهادهای دولتی با تخصص پیشرفته، به خوبی برای مواجهه با چالش‌های تکنولوژیک و لجستیکی ساخت چنین سیستمی مجهز است. با این حال، به دلیل چالش‌های سیاسی و فرهنگی، ممکن است یکی از کمترین احتمال‌ها را برای اجرای آن داشته باشد. برخی آمریکایی‌ها، مراکز کنترل و پیشگیری از بیماری را تهدیدی بزرگ‌تر از روسیه یا کره شمالی می‌دانند. تأسیس یک سیستم هشدار اولیه هوشمند همه‌گیری در ایالات متحده، یک پروژه جاه‌طلبانه تر از تأسیس چنین سیستمی در ماه است.

هر کشوری که چنین سیستمی را تأسیس کند، آزادی شهروندان خود را از بیماری، قرنطینه‌ها، و اختلال اقتصادی به حداکثر می‌رساند و آزادی آنها را برای کار، سفر، و گردهمایی بدون ترس حفظ

می‌کند. دستیابی به چنین پروژه‌ای نیازمند یک حس مشترک از هدف ملی است. هر چالش تکنولوژیک بزرگی که دنبال کنیم، نیازمند سطح مشابهی از اجماع ملی و چشم‌انداز مشترک خواهد بود. به همین دلیل، هوش مصنوعی که به عنوان امتدادی از اراده فردی عمل می‌کند، بسیار حیاتی است. این ممکن است متناقض به نظر برسد. اما به عنوان مثال، سیستم بزرگراهی بین‌ایالتی آمریکا(IHS) ، نمونه قدرتمندی از چگونگی شکل‌دهی یک کشور و ایجاد بسترهای پایدار برای رشد و نوآوری توسط پروژه‌های عمومی گسترده و هماهنگ است.

سفر تقریباً بدون برنامه‌ریزی و با احتمال موفقیت بسیار بالا از اسپرینگفیلد، ایلینوی، به ساکرامنتو، کالیفرنیا، که اکنون امکان‌پذیر است، در دهه ۱۸۴۰ حداقل چهار ماه طول می‌کشید، حتی در شرایط ایده‌آل. این سفر همچنین نیازمند زمان مشابهی تنها برای آماده‌سازی بود: هماهنگی با دیگران برای سفر در یک گروه بزرگ، تهیه واگن و گاو برای کشیدن آن، ذخیره مواد غذایی برای چند ماه، و تهیه ابزارها، اسلحه، لباس، و لوازم کمک‌های اولیه. امروز، تقریباً فقط فضانوردان اینقدر زمان برای آماده‌سازی یک سفر صرف می‌کنند.

شبکه جاده‌ها، بزرگراه‌ها و دیگر زیرساخت‌های فیزیکی که مسیری امن و کارآمد را از میان سرزمین‌های دشوار فراهم می‌کند، نقش حمایتگرانه دولت را نشان می‌دهد. سیستم بزرگراهی بین‌ایالتی، که توسط قانون بزرگراه فدرال-کمک‌رسانی در سال ۱۹۵۶ مجاز شد، نقشی محوری در متحد کردن یک کشور وسیع، ایجاد فرصت‌های اقتصادی جدید و الگوهای زندگی ایفا کرد. این سیستم، به عنوان یک زیرساخت ملی مشترک که فراتر از مرزهای ایالتی و منافع محلی بود، نیازمند همکاری گسترده بین دولت‌های فدرال، ایالتی و محلی، و همچنین صنعت خصوصی و شهروندان عمومی بود. در اوایل ساخت آن، به عنوان «بزرگترین برنامه کارهای عمومی در تاریخ جهان» توصیف شد. این بزرگراه‌ها، رانندگی را سریع‌تر و ایمن‌تر کردند، و نرخ تلفات آنها به طور منظم کمتر از نصف سایر انواع جاده‌ها بوده است. در طول بیش از شصت سال، این بزرگراه‌های ایمن صدها هزار جان را نجات داده‌اند.

تأثیر IHS بر چشم‌انداز اقتصادی آمریکا نیز قابل توجه بوده است. قبل از آن، حمل‌ونقل کالاها

عمدتاً از طریق راه‌آهن و کشتی انجام می‌شد. برای کشاورزان و تولیدکنندگان کوچک، عدم دسترسی به حمل‌ونقل کارآمد در مسافت‌های طولانی، دسترسی به بازار و رقابت‌پذیری آنها را به شدت محدود می‌کرد IHS .با کاهش زمان سفر بین شهرها و ایجاد اتصالات کارآمد به پایانه‌های راه‌آهن، بنادر دریایی و فرودگاه‌ها، به کاهش قیمت کالاها و خدمات برای همه کمک کرد. کاهش هزینه‌های تولید و توزیع منجر به مشاغل بیشتر و افزایش رقابت‌پذیری در بازارهای جهانی شد. امروزه، ارزش اقتصادی سالانه IHS ۷۴۲ میلیارد دلار تخمین زده می‌شود.

IHSهمچنین توسعه در مناطقی را تحریک کرد که قبلاً امیدی به آنها نبود. امروزه، در طول سفر در بزرگراه80-I ، نه تنها دولت آمریکا از طریق زیرساخت‌های حمل‌ونقل کارآمد و نیروهای امدادی در صورت بروز مشکل از شما حمایت می‌کند، بلکه یک شبکه از هتل‌ها، توقفگاه‌های کامیون، رستوران‌ها و فروشگاه‌های عمومی نیز در طول مسیر برای کاهش خطرات سفر شما وجود دارد. این زیرساخت‌ها، با وجود انتقاداتی که ممکن است به یکنواختی زندگی آمریکایی مدرن وارد شود، در مجموع به سفرهای سریع‌تر و ایمن‌تر، رشد اقتصادی، بهبود زمان پاسخگویی به موارد اضطراری، و تحرک شخصی بی‌سابقه منجر شده‌اند. با این حال، IHSنیز با معایب خود همراه بود. در اوایل توسعه آن، با مخالفت‌هایی، به ویژه در شهرهای پرجمعیت که بزرگراه‌های چهاربانده محله‌های مستقر را تقسیم می‌کردند، مواجه شد که منجر به پیامدهای منفی مانند تصرف زمین، سر و صدا، آلودگی و کاهش جمعیت شد. اعتراضات مردمی، که به «شورش‌های بزرگراهی» معروف شدند، منجر به تغییر مسیرها، لغو پروژه‌ها و حتی حذف بزرگراه‌های ساخته شده در برخی مناطق شهری شد. این شورش‌ها، درسی هشداردهنده در مورد خطرات برنامه‌ریزی متمرکز و از بالا به پایین است که تلاشی برای جلب رضایت شهروندان نمی‌کند.

نتیجه‌گیری

این فصل به روشنی نشان می‌دهد که آزادی، به ویژه در زمینه خودمختاری فردی، مفهومی ثابت و مجرد نیست، بلکه همواره در حال تغییر و تکامل است. این تحول، نتیجه تعامل

پیچیده‌ای میان نوآوری‌های فناورانه، ضرورت‌های اجتماعی، و چارچوب‌های نظارتی است. از توسعه اتومبیل‌رانی گرفته تا ظهور هوش مصنوعی، و از اختراع چاپ تا پاسخ‌های جمعی به بحران‌های بهداشت عمومی، هر گام در پیشرفت بشر، هم آزادی‌های جدیدی را خلق کرده و هم نیازمند بازتعریف محدودیت‌ها و مسئولیت‌های جمعی بوده است.

رویکرد «خودمختاری شبکه‌ای» که در این فصل مطرح شد، تأکید می‌کند که تقویت آزادی فردی اغلب نیازمند زیرساخت‌های مشترک، مقررات هوشمندانه، و مهم‌تر از همه، اجماع و اعتماد اجتماعی است. داستان کره جنوبی در مدیریت کووید-۱۹، با تکیه بر شفافیت داده‌ها و مشارکت شهروندان، و همچنین موفقیت سیستم بزرگراهی بین‌ایالتی آمریکا در ایجاد یک کشور منسجم و پویا، نمونه‌های قدرتمندی هستند که نشان می‌دهند چگونه سرمایه‌گذاری در خیر عمومی و پذیرش مسئولیت‌های مشترک، به نوبه خود، به افزایش چشمگیر آزادی‌های فردی و کیفیت زندگی می‌انجامد.

در نهایت، در عصر هوش مصنوعی که نویدبخش دگرگونی‌های بی‌سابقه‌ای است، درک این دینامیک‌ها حیاتی است. چالش‌هایی مانند «دموکراتیک کردن ریسک» و نیاز به نظارت‌های جدید، مستلزم آن است که جامعه به طور جمعی تصمیم بگیرد که چگونه می‌توان آزادی عمل فردی را با امنیت و پایداری جمعی همسو ساخت. این مسیر نه تنها نیازمند نوآوری‌های تکنولوژیک، بلکه به همان اندازه، نیازمند رهبری سیاسی، گفتگوی عمومی، و ایجاد حس مشترک از هدف ملی است تا بتوانیم آینده‌ای را بسازیم که در آن، آزادی نه به معنای رهایی از تمام قید و بندها، بلکه به معنای توانمندی برای شکوفایی در یک جامعه متصل و مسئولیت‌پذیر باشد.

نکات کلیدی

- **آزادی مفهومی پویاست** :آزادی ثابت نیست و در بستر تغییرات تکنولوژیکی، اجتماعی و مقرراتی دگرگون می‌شود.
- **فناوری و مقررات در هم تنیده‌اند** :نوآوری‌های فناورانه (مانند خودرو، چاپ، هوش

مصنوعی) همواره منجر به وضع مقررات جدید شده‌اند که اغلب به گسترش یا بازتعریف آزادی کمک می‌کنند.

- **خودمختاری شبکه‌ای** :آزادی فردی در یک بستر شبکه‌ای و وابسته به هم بهتر شکوفا می‌شود، جایی که اقدامات جمعی و زیرساخت‌های مشترک آن را تقویت می‌کنند.
- **تعادل بین آزادی و امنیت** :جوامع همواره در حال یافتن تعادلی بین آزادی‌های فردی (مانند سرعت رانندگی) و آزادی از خطرات (مانند تصادفات) هستند.
- **نقش زیرساخت‌های جمعی** :پروژه‌های بزرگ زیرساختی مانند سیستم بزرگراهی بین‌ایالتی، با وجود چالش‌های اولیه، به طور چشمگیری تحرک فردی و رشد اقتصادی را افزایش داده‌اند.
- **هوش مصنوعی و آینده آزادی** :هوش مصنوعی آزادی‌های جدیدی را به ارمغان می‌آورد اما نیازمند مقررات جدید برای مدیریت ریسک‌ها و چالش‌های امنیتی است.
- **شفافیت و اعتماد عمومی** :در موارد بحرانی مانند همه‌گیری، شفافیت داده‌ها و مشارکت عمومی می‌تواند به ایجاد اعتماد و اجرای موفقیت‌آمیز سیاست‌ها کمک کند (نمونه کره جنوبی).
- **چالش‌های فرهنگی و سیاسی** :پیاده‌سازی سیستم‌های نظارتی پیشرفته، حتی برای خیر عمومی، ممکن است با مقاومت‌های فرهنگی و سیاسی قابل توجهی مواجه شود.
- **اهمیت اجماع ملی** :دستیابی به اهداف فناورانه و اجتماعی بزرگ نیازمند یک حس مشترک از هدف ملی و چشم‌انداز مشترک است.
- **درس از گذشته** :تاریخچه ماشین چاپ نشان می‌دهد که چگونه فناوری‌های انقلابی منجر به بحث و چارچوب‌بندی جدیدی برای مفاهیم بنیادی مانند آزادی بیان شدند.

سؤالات تفکربرانگیز

۱. با توجه به تکامل مفهوم آزادی در طول تاریخ، چگونه می‌توانیم «خودمختاری شبکه‌ای» را به گونه‌ای طراحی کنیم که هم آزادی‌های فردی را به حداکثر برساند و هم پایداری و امنیت جامعه را تضمین کند؟

۲. آیا پذیرش مقررات و نظارت‌های بیشتر در ازای افزایش ایمنی و کارایی (مانند وسایل نقلیه خودران یا سیستم‌های نظارت بر سلامت مبتنی بر هوش مصنوعی) همواره به معنای از دست دادن آزادی است، یا می‌تواند به بازتعریف و گسترش آن منجر شود؟

۳. چگونه می‌توانیم چالش‌های فرهنگی و سیاسی را که مانع از اجرای پروژه‌های بزرگ زیرساختی یا سیستم‌های نظارتی پیشرفته برای خیر عمومی می‌شوند، برطرف کنیم؟

۴. در عصر هوش مصنوعی، چه تدابیری می‌توان اتخاذ کرد تا اطمینان حاصل شود که جمع‌آوری و استفاده از داده‌های فردی برای اهداف عمومی، به روشی شفاف و عادلانه انجام می‌شود و از سوءاستفاده جلوگیری می‌کند؟

۵. با توجه به تفاوت در پذیرش اقدامات جمعی و نظارتی بین کشورها (مانند کره جنوبی و ایالات متحده در مواجهه با همه‌گیری)، چه درس‌هایی می‌توانیم برای ایجاد اجماع ملی و افزایش اعتماد عمومی در پروژه‌های آینده بیاموزیم؟

۳۰ جمله کلیدی

۱. خودمختاری شبکه‌ای، مفهوم جدیدی از آزادی در عصر فناوری است.

۲. وسایل نقلیه خودران، درک سنتی از آزادی در اتومبیل‌رانی را به چالش می‌کشند.

۳. آزادی نوشیدن در خودروهای خودران با محدودیت‌های مسیر و سرعت همراه خواهد بود.

۴. تاریخچه خودرو در آمریکا، تلفیقی از خودمختاری فردی و اقدامات جمعی بود.

۵. شبکه مقررات، در کنار نوآوری، آزادی شخصی را افزایش داد.

۶. آزادی چندوجهی است و اغلب شامل تعادل بین آزادی عمل و آزادی از خطر می‌شود.
۷. قوانین گواهینامه رانندگی، ایمنی و قابلیت مقیاس‌پذیری اتومبیل‌رانی را افزایش دادند.
۸. فناوری، مفهوم آزادی را در طول زمان تغییر می‌دهد و آن را ثابت نگه نمی‌دارد.
۹. مقررات اغلب زمانی ظاهر می‌شوند که فناوری توانایی‌های فوق‌انسانی را ممکن می‌سازد.
۱۰. سفر مدرن برون‌مرزی، با وجود نظارت، آزادی و ایمنی بی‌سابقه‌ای را ارائه می‌دهد.
۱۱. حزب دونر، نمونه‌ای از آزادی بدون محدودیت است که به فاجعه انجامید.
۱۲. آزادی خارق‌العاده امروز، نتیجه نوآوری، قوانین و زیرساخت‌های مشترک است.
۱۳. اختراع ماشین چاپ، در ابتدا آزادی بیان را گسترش داد، اما سپس به مقررات جدیدی نیاز داشت.
۱۴. قوانین سانسور پس از چاپ، به تدریج مفهوم آزادی بیان را برجسته کردند.
۱۵. هوش مصنوعی نیز مانند خودرو، اشکال جدیدی از مقررات را الهام خواهد گرفت.
۱۶. دموکراتیک کردن دسترسی به هوش مصنوعی، به معنای دموکراتیک کردن ریسک است.
۱۷. مجوزهای هوش مصنوعی و احراز هویت بیومتریک ممکن است به استانداردهای جدیدی تبدیل شوند.
۱۸. جان استوارت میل بر اهمیت آزادی فردی برای رفاه کلی جامعه تأکید داشت.
۱۹. افراد موفق منجر به جوامع موفق می‌شوند و جوامع موفق افراد را تقویت می‌کنند.
۲۰. کره جنوبی در مقابله با کووید-۱۹، آزادی حرکت را با جمع‌آوری داده‌های تهاجمی

ترکیب کرد.

۲۱. شفافیت داده‌ها و مشارکت عمومی، عامل کلیدی موفقیت کره جنوبی در همه‌گیری بود.

۲۲. هوش مصنوعی می‌تواند برای ایجاد سیستم‌های هشدار اولیه همه‌گیری در آینده استفاده شود.

۲۳. پیاده‌سازی سیستم‌های نظارتی هوشمند با چالش‌های اخلاقی، قانونی و لجستیکی همراه است.

۲۴. ایالات متحده از نظر فنی برای پروژه‌های بزرگ هوش مصنوعی مجهز است، اما از نظر سیاسی چالش دارد.

۲۵. سیستم بزرگراهی بین‌ایالتی آمریکا، نمونه‌ای از موفقیت زیرساخت‌های جمعی در ایجاد آزادی است.

۲۶. IHSنه تنها تحرک فردی را افزایش داد، بلکه به رشد اقتصادی گسترده نیز کمک کرد.

۲۷. توسعه IHS با مخالفت‌هایی در مناطق شهری همراه بود و درس‌هایی در مورد برنامه‌ریزی از بالا به پایین آموخت.

۲۸. رضایت شهروندان، حتی با وجود قدرت دولتی، برای موفقیت پروژه‌های بزرگ ضروری است.

۲۹. آزادی امروز ما، به دلیل زیرساخت‌ها و خدمات دولتی، به مثابه توانمندی‌های ابرانسانی است.

۳۰. دستیابی به اهداف فناورانه بزرگ، نیازمند اجماع ملی و یک چشم‌انداز مشترک است.

فصل ۱۰

آمریکای هوشمند: آینده ملت و شهروندان در عصر هوش مصنوعی

"هوش مصنوعی نه تنها ابزاری برای بهینه‌سازی، بلکه آینه‌ای است که در آن می‌توانیم بازتاب ارزش‌های ملی، ساختارهای اجتماعی و آرزوهای انسانی خود را مشاهده کنیم. چگونگی ادغام آن در تار و پود جامعه، مسیر آینده یک ملت را رقم خواهد زد." – دکتر لینا خان، نظریه‌پرداز فناوری و جامعه‌شناس

اهداف اصلی فصل:

- بررسی تأثیر فناوری‌های تحول‌آفرین بر جوامع از منظر تاریخی و ارزیابی پیامدهای اتخاذ رویکردهای متفاوت.
- تحلیل مفهوم هوش مصنوعی ملی (Sovereign AI) و اهمیت آن برای حفظ استقلال اقتصادی، امنیتی و فرهنگی کشورها.
- شناسایی چالش‌ها و فرصت‌های پیش روی ایالات متحده در مسیر توسعه و به‌کارگیری هوش مصنوعی در سطح ملی.
- کاوش در پتانسیل هوش مصنوعی برای بازتعریف خدمات دولتی و تقویت مشارکت شهروندی و فرآیندهای دموکراتیک.
- تبیین نقش هوش مصنوعی در ارتقاء رفاه فردی، توانمندی‌های شهروندان و رشد فراگیر اجتماعی.

چکیده

این فصل به بررسی عمیق پیامدهای تحول‌آفرین هوش مصنوعی بر آینده ملت‌ها و شهروندان می‌پردازد و به طور خاص، موقعیت ایالات متحده را در این چشم‌انداز جدید مورد ارزیابی قرار می‌دهد. با آغاز از یک درس تاریخی از جنبش لودیت‌ها و پیامدهای یک تاریخ جایگزین که در آن احتیاط مفرط در برابر نوآوری‌های تکنولوژیک منجر به رکود ملی در انگلستان می‌شود، فصل نشان می‌دهد که چگونه عدم پذیرش استراتژیک فناوری می‌تواند به زیان‌های جبران‌ناپذیری منجر گردد. در ادامه، مفهوم هوش مصنوعی ملی (Sovereign AI) معرفی می‌شود که بیانگر لزوم کنترل کشورها بر زیرساخت‌های هوش مصنوعی خود برای حفظ حریم خصوصی داده‌ها، امنیت ملی و هویت فرهنگی است. این فصل به تفصیل، چالش‌های پیش روی ایالات متحده در مسیر هم‌افزایی نوآوری و تنظیم‌گری را بررسی می‌کند و سپس چشم‌اندازی از دولت ۲.۰ (Government 2.0)ارائه می‌دهد که در آن هوش مصنوعی می‌تواند به عنوان کاتالیزوری برای بهبود خدمات عمومی، افزایش مشارکت شهروندی و تقویت بنیان‌های دموکراتیک عمل کند. در نهایت، با تأکید بر نقش هوش مصنوعی در توانمندسازی فردی و افزایش رفاه جمعی، فصل به ضرورت اتخاذ یک رویکرد جامع و آینده‌نگرانه برای شکل‌دهی به یک آمریکای هوشمند و پایدار در قرن بیست و یکم اشاره می‌کند.

مقدمه

تاریخ بشر همواره با دوره‌هایی از تحولات تکنولوژیک عظیم گره خورده است که بنیان‌های اجتماعی، اقتصادی و سیاسی را دگرگون ساخته‌اند. از انقلاب کشاورزی تا انقلاب صنعتی و عصر دیجیتال، هر جهش فناورانه، چالش‌ها و فرصت‌های بی‌سابقه‌ای را پیش روی جوامع قرار داده است. در نگاهی به گذشته، برخی گروه‌ها این تغییرات را به عنوان تهدیدی برای شیوه‌های زندگی سنتی و استقلال فردی تلقی کرده‌اند، در حالی که دیگران آن را فرصتی برای پیشرفت، اکتشاف و خودسازی دیده‌اند. داستان حزب دونر، که با وجود ثبات و امنیت، رویای آینده‌ای روشن‌تر در غرب را دنبال کردند، نمادی از تعهد به اکتشاف، سازگاری و بهبود فردی است که همواره از ویژگی‌های هویت ملی آمریکا بوده است. در مقابل، واکنش لودیت‌ها در اوایل قرن نوزدهم به مکانیزاسیون صنعت نساجی انگلستان، رویکردی متفاوت را به نمایش می‌گذارد. این گروه، به جای پذیرش یا تطبیق، به مقاومت خشونت‌آمیز در برابر ماشین‌آلات پرداختند، زیرا آن‌ها را نمادی از سیستم کارخانه‌ای استثمارگر و از بین برنده آزادی فردی می‌دیدند. این دوگانگی تاریخی، درس‌های ارزشمندی را برای درک چگونگی مواجهه با جدیدترین و شاید قدرتمندترین موج تحول تکنولوژیک ما، یعنی هوش مصنوعی، ارائه می‌دهد.

هوش مصنوعی در حال حاضر نه تنها یک فناوری، بلکه نیروی محرکه‌ای است که پتانسیل بازتعریف تمامی جنبه‌های زندگی بشر را دارد؛ از نحوه کار و ارتباط ما گرفته تا ساختارهای حکومتی و حتی تعریف از هویت ملی. این موج جدید نوآوری، کشورها را در سراسر جهان وادار به تأمل در مورد استراتژی‌های ملی خود می‌کند تا اطمینان حاصل کنند که می‌توانند از فرصت‌های بی‌شمار آن بهره‌مند شده و در عین حال، خطرات احتمالی را مدیریت کنند. ایالات متحده، به عنوان یکی از پیشگامان در توسعه هوش مصنوعی، در یک دوراهی حساس قرار دارد. چگونگی رویکرد این کشور به هوش مصنوعی – چه از طریق تسهیل نوآوری و چه از طریق تنظیم‌گری دقیق – نه تنها مسیر داخلی آن را تعیین خواهد کرد، بلکه بر جایگاه آن در صحنه جهانی نیز تأثیر می‌گذارد. این فصل با الهام از درس‌های تاریخی و با تمرکز بر چالش‌ها و فرصت‌های کنونی، به

بررسی چگونگی شکل‌گیری یک "آمریکای هوشمند" می‌پردازد که در آن هوش مصنوعی به عنوان ابزاری برای تقویت آژانس فردی، رفاه ملی و بنیان‌های دموکراتیک عمل کند.

درس‌هایی از تاریخ: هزینه احتیاط در برابر پیشرفت

نگاهی به گذشته، دیدگاه‌های متفاوتی را در مورد پذیرش یا مقاومت در برابر تغییرات تکنولوژیک آشکار می‌سازد. در اوایل قرن نوزدهم، در انگلستان، جنبش لودیت‌ها در برابر مکانیزاسیون فزاینده صنعت نساجی قیام کرد. اختراعاتی مانند شاتل پرنده، چرخ نخ‌ریسی جنی و در نهایت ماشین بافندگی، فرآیندهای تولید پارچه را متحول کرده و نیاز به نیروی کار انسانی را کاهش داده بودند. این نوآوری‌ها، اگرچه بهره‌وری را افزایش می‌دادند، اما معیشت بافندگان خانگی را که قرن‌ها در خانه‌های خود به این صنعت مشغول بودند، به شدت تهدید می‌کردند. شرایط اقتصادی کلان نیز این وضعیت را وخیم‌تر می‌ساخت؛ جنگ طولانی و پرهزینه با فرانسه، مالیات‌ها را افزایش داده، تحریم‌های تجاری بازارهای صادراتی را از بین برده و قیمت مواد غذایی را بالا برده بود. در همین حال، صاحبان کارخانه‌ها دستمزدهای مشاغلی را که هنوز تحت تأثیر اتوماسیون قرار نگرفته بودند، کاهش می‌دادند.

لودیت‌ها، که نام خود را از نِد لود، یک شاگرد بافنده افسانه‌ای، گرفته بودند، در پاسخ به این فشارها، دست به حملات شبانه به کارخانه‌ها زدند و هزاران دستگاه ماشینی از جمله ماشین‌های بافندگی و سایر ماشین‌آلات را تخریب کردند. هدف اصلی آن‌ها صرفاً فناوری نبود، بلکه سیستم کارخانه‌ای، شرایط کاری استثمارگرانه آن و از دست دادن آزادی و خودمختاری فردی که این شیوه جدید زندگی به دنبال داشت، مورد اعتراض آن‌ها بود. کارخانه‌ها در نظر لودیت‌ها چیزی کمتر از زندان نبودند و تهدیدی آشکار برای عاملیت و استقلال فردی محسوب می‌شدند. دولت بریتانیا با شدت به این جنبش واکنش نشان داد و هزاران سرباز را مستقر کرد. قانون "شکستن چارچوب" در سال ۱۸۱۲، تخریب ماشین‌آلات را به جرمی مجازات اعدام تبدیل کرد و ده‌ها لودیت به دار آویخته شدند یا به مستعمرات کیفری بریتانیا در استرالیا تبعید گشتند. این جنبش

سرانجام در سال ۱۸۱۶ فروکش کرد، اما تأثیرات آن در برخی مناطق برای دهه‌ها باقی ماند.

برای درک عمیق‌تر پیامدهای مقاومت در برابر فناوری، می‌توان یک تاریخ جایگزین را متصور شد. فرض کنید لودیت‌ها موفق می‌شدند اقدامات خود را به عنوان پاسخی مشروع و ضروری به تهدیدی وجودی برای آزادی فردی و رفاه عمومی به تصویر بکشند. فرض کنید آن‌ها حمایت گسترده‌ای از تمامی اقشار جامعه، از طبقه کارگر تا اشراف و حتی نخبگان دولتی، به دست می‌آوردند. در این سناریوی جایگزین، پس از سال‌ها کشمکش، پارلمان انگلستان قانون "شغل، ایمنی و کرامت انسانی (JSHDA) "را در سال ۱۸۲۰ تصویب می‌کرد. این قانون، یک الزام قوی احتیاطی برای ارزیابی فناوری‌های جدید در انگلستان معرفی می‌نمود. بر اساس این قانون، هر فناوری جدیدی که می‌توانست بر مشاغل موجود، صنایع، ساختارهای اجتماعی یا سبک زندگی به طرق پیش‌بینی‌نشده تأثیر بگذارد، یا ایمنی آن به اثبات نرسیده بود، اجازه استقرار عمومی نمی‌یافت. انگلستان پس از تصویب این قانون، عملاً مانند جوامع آمیش در ایالات متحده امروزی عمل می‌کرد.

در ابتدا، این سیاست به نظر می‌رسید که یک دوره بازگشت مثبت و ترمیم‌کننده باشد. ماشین‌آلات صنعتی در سراسر کشور از رده خارج شدند، بسیاری از کارخانه‌ها تعطیل گشتند و به ساعات کاری طولانی، کار کودکان و شرایط کاری خطرناک پایان داده شد. بافندگان و بافندگان بافندگی دوباره در خانه‌های خود کار می‌کردند، برنامه‌های کاری خود را کنترل می‌نمودند و زندگی خود را حول خانواده و جامعه سازماندهی می‌کردند. قیمت لباس و سایر کالاهای نهایی افزایش یافت و تنوع کاهش پیدا کرد، اما زندگی با احساس رضایت آشنا و سنتی ادامه یافت. در حالی که کشورهای همسایه با آشفتگی‌های اجتماعی ناشی از صنعتی شدن دست و پنجه نرم می‌کردند، زندگی در انگلستان سنتی، غیرمتمرکز، هنری و پایدار باقی ماند. بافندگان به کار خود، به خودمختاری و عاملیتی که رویکردشان به کار به آن‌ها می‌بخشید، افتخار می‌کردند. جوامع نزدیک به هم باقی ماندند و حس پایداری و تداوم برقرار بود.

اما تا پایان قرن نوزدهم، تفاوت‌ها بین انگلستان و همسایگانش به گونه‌ای آشکار شد که دیگر

چندان محبوب نبود. در خارج از انگلستان، دهه‌ها پس از تصویب قانون JSHDA با نوآوری‌های تحول‌آفرین بی‌وقفه همراه بود. ابتدا موتورهای بخار، سپس راه‌آهن که سطح بی‌سابقه‌ای از تحرک و تجارت را ممکن ساخت، و سپس مجموعه‌ای گیج‌کننده از پیشرفت‌های تکنولوژیکی شگفت‌انگیز مانند تلگراف، عکاسی، تولید فولاد، حفاری نفت، تلفن، لامپ الکتریکی، موتور احتراق داخلی، اشعه ایکس، امواج رادیویی و تصاویر متحرک به وجود آمدند. در آن سوی آب‌ها، پیشرفت اجتماعی نیز همگام با پیشرفت تکنولوژیک تسریع شد .جنبش‌های کارگری به رسمیت شناخته شدند، قوانین حداقل دستمزد و محدودیت‌های کار کودکان بسیاری از بدترین افراط‌های صنعتی شدن اولیه را تعدیل کردند. فرصت‌های جدیدی برای زنان برای رهبری زندگی‌های مستقل‌تر و خودگردان‌تر پدید آمد. آموزش عمومی گسترده‌تر شد. دولت‌ها برنامه‌های رفاه اجتماعی را برای مقابله با فقر و نگرانی‌های بهداشت عمومی اجرا کردند.

و انگلستان؟ آن‌ها پتو داشتند. کمی خشن. نسبتاً گران. اما گرم و اصیل، قطعاتی واقعاً موروثی. وقتی نوآوری جدیدی در فرانسه یا آلمان یا ایالات متحده ظاهر می‌شد، تنظیم‌کنندگان بریتانیایی معمولاً سرانجام به بررسی آن می‌پرداختند. اما اکثر آن‌ها به دلیل تأثیرات آشکاری که بر مشاغل موجود می‌گذاشتند یا اختلالات اجتماعی بالقوه‌ای که احتمالاً ایجاد می‌کردند، رد می‌شدند. با گذشت دهه‌ها، این رویکرد احتیاطی تأثیرات مرکبی به بار آورد. با پذیرش روش‌های تولید خودکار توسط سایر کشورها، بازارهای صادراتی نساجی انگلستان فروپاشید. تحریم‌های تجاری بر واردات از هر نوع برای حفاظت از صنایع داخلی انگلستان در برابر رقابت خارجی کارآمدتر، به‌طور فزاینده‌ای ضروری شد و این امر به طور جدی دامنه و کیفیت کالاهایی را که مصرف‌کنندگان انگلیسی می‌توانستند خریداری کنند، محدود کرد.

در حالی که سایر کشورها از فناوری برای گسترش قدرت نیروهای نظامی خود استفاده می‌کردند، پارلمان انگلستان، با ترس از تأثیری که شکاف‌های فزاینده تکنولوژیکی می‌توانست بر امنیت ملی داشته باشد، به‌طور فزاینده‌ای استثنائاتی را برای قانون JSHDA قائل شد. این امر

به ارتش انگلستان اجازه داد تا کارخانه‌های خود را ایجاد کرده و فناوری‌های جدیدی توسعه دهد – اما این تلاش‌ها هنوز توسط عوامل متعددی محدود می‌شد. از آنجا که خروجی اقتصادی طی سال‌ها عمدتاً ثابت مانده بود، پایه مالیاتی کشور نمی‌توانست به اندازه کافی از یک ارتش مدرن حمایت کند. انگلستان همچنین دارای پایگاه استعدادی بسیار محدودتری بود، با مهندسان، فیزیکدانان، مدیران کارخانه و تکنسین‌های آزمایشگاهی کمی که بتواند بر روی آن‌ها حساب کند. مهاجرت فزاینده مغزها اوضاع را بدتر کرد. در دهه ۱۸۲۰، کنترل صنعتی شدن استراتژی‌ای منطقی – شاید تنها استراتژی – برای حفظ عاملیت فردی در برابر سرکوب تکنولوژیکی قریب‌الوقوع به نظر می‌رسید. اما گاهی اوقات آینده، مسیرهای غیرمنتظره‌ای به خود می‌گیرد. پس از هشتاد سال، همان کارخانه‌های زندان‌مانند و صنعتی شدن به‌طور کلی، در کشورهایی که نوآوری تکنولوژیکی را پذیرفته بودند، رفاه، گزینه‌های چشمگیر و افزایش عاملیت فردی را به همراه آورده بود. در فرانسه، آلمان، روسیه، ژاپن و به‌ویژه ایالات متحده، شما می‌توانستید مهندس لوکوموتیو، اپراتور تلفن، مدیر کارخانه، اپراتور دوربین فیلم‌برداری، تندنویس، تحویل‌دهنده تلگرام، شیمی‌دان، معمار طراح آسمان‌خراش‌ها و پل‌های معلق باشید. امکانات بی‌شمار بودند. در انگلستان، شما می‌توانستید پارچه تولید کنید. بنابراین، بسیاری از جاه‌طلب‌ترین و بااستعدادترین شهروندان آن به خارج از کشور مهاجرت می‌کردند.

تا اوایل قرن بیستم، زمانی که سایر کشورها از برق‌رسانی گسترده و سیستم‌های لوله‌کشی شهری بهره‌مند می‌شدند، نوادگان لودیت‌ها بالاخره شروع به فشار برای تغییرات عمده کردند. آن‌ها مخفیانه در شب ملاقات می‌کردند و چرخ‌های نخ‌ریسی جنی و ماشین‌های بافندگی می‌ساختند. از دولت درخواست کردند تا کارخانه بسازد. اما دولت، که در راه‌های سنتی خود ریشه‌دار شده بود، با سرکوب سریع و بی‌رحمانه پاسخ داد. نسل جدیدی از لودیت‌ها به سرنوشت پیشینیان خود دچار شدند. برخی اعدام شدند. برخی دیگر به زندان افتادند. انگلستان به وفاداری خود به صنایع دستی سنتی ادامه داد. این داستان جایگزین، کاهنده و ساده‌انگارانه است، اما هدف آن این است که نشان دهد فناوری‌هایی که اغلب توسط منتقدانشان به عنوان

عامل انسانی‌زدایی و محدودکننده به تصویر کشیده می‌شوند، معمولاً به عامل انسانی‌سازی و رهایی‌بخش تبدیل می‌شوند.

رقابت جهانی هوش مصنوعی و ظهور هوش مصنوعی ملی(Sovereign AI)

با ظهور هوش مصنوعی، این روند ادامه خواهد یافت و تأثیرات فردی، ملی و جهانی خواهد داشت. ملتی که در پذیرش کشف دارو با هوش مصنوعی و تکنیک‌های پزشکی شخصی‌سازی‌شده کند عمل کند، ممکن است به زودی با شکاف قابل توجهی در نتایج مراقبت‌های بهداشتی مواجه شود. ملتی که از کشاورزی دقیق با هوش مصنوعی و کشاورزی سازگار با آب‌وهوا بهره‌مند نشود، احتمالاً با افزایش هزینه‌های مواد غذایی و در سناریوهای شدیدتر، با افزایش کمبود غذا روبرو خواهد شد. ملتی که گزینه‌های کمتری برای توسعه شخصی و پیشرفت شغلی دارد، به کاهش عاملیت نسبی شهروندان خود دامن می‌زند—که احتمالاً باعث فرار مغزها می‌شود، زیرا متخصصان برجسته علوم، فناوری، مهندسی و ریاضیات (STEM) آن به کشورهایی با سیاست‌های دوستدار هوش مصنوعی بیشتر مهاجرت می‌کنند.

استعداد انسانی تنها عامل بهره‌وری نیست که نامتوازن‌تر خواهد شد. اریک اشمیت، در بیانیه‌ای کتبی که در دسامبر ۲۰۲۳ به مجلس سنای ایالات متحده ارائه داد، اشاره کرد که مدل‌ها احتمالاً طی دهه آینده ۱۰۰۰ تا ۱۰۰۰۰ برابر قدرتمندتر خواهند شد. مدل‌های جدید بسیار عامل‌تر و قادر به برنامه‌ریزی متوالی خواهند بود، به این معنی که خودشان به عنوان برنامه‌نویسان مجازی، مهندسان، دانشمندان و سایر انواع کارگران بسیار ماهر عمل خواهند کرد. این مقایسه‌ها به روشن شدن پیامدهای در حال وقوع کمک می‌کند، اما همچنین مهم است به یاد داشته باشیم که دنیای آتی هوش توزیع‌شده رقابتی کاملاً با حاصل جمع صفر نخواهد بود. هر کشوری که هوش مصنوعی را به روش‌های استراتژیک و به‌خوبی اجرا شده بپذیرد، احتمالاً شاهد افزایش قابل توجهی در بهره‌وری و کارایی خواهد بود. با این حال، مانند هر تغییر تکنولوژیکی، برندگان و بازندگانی نسبی وجود خواهند داشت. برخی بازیگران بیشتر از دیگران سود خواهند برد، که عمدتاً

به سرعت و جسارت آن‌ها در پذیرش فرصت‌های جدید و گسترده‌ای که هوش مصنوعی ممکن می‌سازد، بستگی دارد.

"برندگان" در دوران نوآوری معمولاً امواج تکنولوژیکی جدید را زودتر شکار می‌کنند. آن‌ها با جسارت بر روی این امواج سوار می‌شوند و به طور ایده‌آل، از شتاب خود برای پیشروی سریع‌تر و فراتر از رقبایی که رویکرد محتاطانه‌تری دارند، استفاده می‌کنند. در حال حاضر، کشورهایی با منابع قابل توجه در محاسبات و استعداد، مزیت آشکاری در بهره‌برداری از مزایای بالقوه هوش مصنوعی دارند. اما همانطور که قبلاً اشاره شد، پیشرفت‌های تکنولوژیکی به تنهایی موفقیت را تضمین نمی‌کنند. ادغام موفقیت‌آمیز فناوری‌های جدید نیازمند پیشرفت‌های اجتماعی و سیاسی نیز است. در محیط جهانی امروز، که بسیاری از کشورها با جنگ‌های فرهنگی که شهروندان را قطبی می‌کند و کاهش اعتماد به نهادها روبرو هستند، اجماع فرهنگی و سیاسی از ارزش بالایی برخوردار است.

تنها چند سال پیش، توسعه هوش مصنوعی عمدتاً یک رقابت دو کشوری بود: ایالات متحده در مقابل چین. در ایالات متحده، ترکیبی از استارت‌آپ‌ها و شرکت‌های بزرگ فناوری، پیشرفت‌های تکنولوژیکی کلیدی را پیش می‌بردند، حتی در حالی که چندین آژانس دولتی فدرال به دنبال محدود کردن تلاش‌های آن‌ها از طریق تحقیقات و شکایت‌ها بودند، و دولت‌های ایالتی مشغول وضع ده‌ها قانون جدید بودند. در طرف چینی، دولت میلیاردها دلار در بزرگترین شرکت‌های خصوصی خود، یا "قهرمانان ملی"، سرمایه‌گذاری می‌کرد، با این فرض که پیشگام شدن در توسعه جهانی هوش مصنوعی، سلطه اقتصادی آن را تقویت می‌کند، امنیت ملی را افزایش می‌دهد و ابزارهای قدرتمندی برای گسترش نفوذ جهانی آن فراهم می‌آورد.

در حالی که رقابت بین ایالات متحده و چین همچنان نیروی محرکه در چشم‌انداز جهانی هوش مصنوعی است، دموکراتیزه شدن قدرت محاسباتی، همراه با آگاهی فزاینده از میزان تأثیر هوش مصنوعی بر آینده کشورها، به این معنی است که توسعه به سرعت در حال تبدیل شدن به یک رقابت بزرگتر و گسترده‌تر است. جنسن هوانگ، هم‌بنیان‌گذار و مدیرعامل طولانی‌مدت

شرکت طراحی تراشه انویدیا در سیلیکون ولی، در نشست جهانی دولت‌ها در دبی در فوریه ۲۰۲۴، بیان کرد: "این آغاز یک انقلاب صنعتی جدید است. این انقلاب صنعتی درباره تولید – نه انرژی، نه غذا – بلکه تولید هوش است. و هر کشوری نیاز دارد که تولید هوش خود را در اختیار داشته باشد." او تأکید کرد: "کشور شما مالک داده‌هایی است که کشت می‌کند. این داده‌ها فرهنگ شما، هوش جامعه شما، عقل سلیم شما، تاریخ شما را کدگذاری می‌کنند".

تأکید هوانگ بر حداکثرسازی کنترل بر منابع ملی که وعده می‌دهد برای رقابت اقتصادی و امنیت ملی یک دولت به طور فزاینده‌ای مهم خواهند بود، ارزش عملی آشکاری دارد. همانطور که زیرساخت هوش مصنوعی برای منافع ملی حیاتی می‌شود، یک رویکرد هوش مصنوعی ملی (سوْوِرِن هوش مصنوعی) طیف وسیعی از مسائل بالقوه را حل می‌کند. به عنوان مثال، یک ارائه‌دهنده خدمات خصوصی که در سطح جهانی فعالیت می‌کند، ممکن است کاملاً با قوانین و مقررات کشور دیگری در مورد حریم خصوصی داده‌ها یا استانداردهای امنیت ملی سازگار نباشد. بسته به قوانین کشور میزبان خود، ممکن است ملزم به گزارش درباره هر مشتری خارجی که خدمات می‌دهد، باشد. در صورت بروز تنش‌های دیپلماتیک، کشوری که برای حتی بخشی از عملیات هوش مصنوعی خود به ارائه‌دهندگان خارجی وابسته است، ممکن است خود را تحت تحریم‌ها یا اختلالات زنجیره تأمین ببیند.

به رسمیت شناختن آنچه در خطر است، بسیاری از کشورهای جهان شروع به سرمایه‌گذاری در ساخت زیرساخت‌های هوش مصنوعی خود کرده‌اند. در برخی موارد، دولت‌ها بر این فرض عمل می‌کنند که سرمایه‌گذاری در هوش مصنوعی نه تنها برای رفاه اقتصادی و امنیت ملی، بلکه برای حفظ سنت‌ها و فرهنگ ملتشان نیز ضروری است. برخلاف لودیت‌ها، که می‌ترسیدند صنعتی شدن و اتوماسیون ارزش‌ها و شیوه زندگی آن‌ها را نابود کند، این کشورها به هوش مصنوعی به عنوان ابزاری برای حفظ این چیزها نگاه می‌کنند. سنگاپور، بر اساس همان مشاهده‌ای که هوانگ مطرح کرد، "استراتژی ملی هوش مصنوعی" را تدوین کرده که شامل تعهدی برای توسعه هوش مصنوعی است که "فرهنگ‌ها، ارزش‌ها و هنجارهای محلی و منطقه‌ای منطقه را منعکس

کند." فرانسه نیز قصد و انگیزه‌های مشابهی را بیان کرده است، با دولت خود که متعهد شده است ۵۵۰ میلیون دلار برای ایجاد "قهرمانان هوش مصنوعی" خود به جای تکیه بر توسعه‌دهندگان خارجی سرمایه‌گذاری کند. امانوئل ماکرون، رئیس جمهور فرانسه، در نمایشگاه تجاری فناوری در پاریس در ژوئن ۲۰۲۳، اظهار داشت: "برای این کار، ما همچنین باید پایگاه‌های داده را به زبان فرانسوی ایجاد کنیم. در غیر این صورت، ما از مدل‌هایی با سوگیری‌های به ارث رسیده از آنگلوساکسون‌ها استفاده خواهیم کرد".

هویت‌های ملی، که از زبان و مصنوعات نوشتاری آن – اسطوره‌های بنیادین، روایت‌های تاریخی و ادبیات – سنت‌ها و آداب و رسوم، دین، قانون، موسیقی، معماری و مصنوعات فیزیکی از هر نوع شکل گرفته‌اند، همیشه ساختارهایی عمدتاً اطلاعاتی بوده‌اند. هوش مصنوعی منابع و ابزارهایی را ارائه می‌دهد که می‌تواند این صورت‌فلکی‌های ساختارهای انسانی را به روش‌هایی که قبلاً ممکن نبود، دربرگیرد و مجموعه‌های عظیمی از متن، صدا، موسیقی و تصاویر نماینده را در یک مدل منسجم ترکیب کند. بنابراین، منطقی است که دولت‌ها بخواهند خودمختاری و عاملیت خود را بر هوش مصنوعی‌هایی که به طور فزاینده‌ای در عملیات روزمره خود گنجانده خواهند شد، حفظ کنند. زیرا، در حالی که هر کشور احتمالاً در مورد اینکه کدام ارزش‌ها و ویژگی‌های ملی را هوش مصنوعی‌های آن باید تجسم کنند، درگیر اختلافات داخلی خواهد شد، بسیاری در اعتقادات خود راسخ خواهند بود که هوش مصنوعی آن‌ها نباید حساسیت‌های آمریکایی، یا چینی، یا حساسیت‌های نخبگان فن‌سالار بی‌وطن را به نمایش بگذارد.

هوش مصنوعی در ایالات متحده: توازن بین نوآوری و تنظیم‌گری

در ایالات متحده، از یک سو، ما قبلاً "قهرمانان ملی" در قالب شرکت‌هایی مانندOpenAI ، مایکروسافت، آلفابت، فیس‌بوک، Anthropicو سایر شرکت‌های آمریکایی داریم که نوآوری هوش مصنوعی جهانی را تا به امروز شکل داده‌اند. علاوه بر این، خود دولت فدرال ایالات متحده نیز به نوعی یک پذیرنده اولیه هوش مصنوعی است و ده‌ها آژانس آن، هوش مصنوعی را در عملیات

خود گنجانده‌اند.

از سوی دیگر، گاهی اوقات به نظر می‌رسد که استراتژی ملی هوش مصنوعی آمریکا شامل تبدیل قهرمانان ملی هوش مصنوعی ما به دونده‌های عقب‌مانده، از طریق جریانی ثابت از اقدامات اجرایی ضد انحصار، جلسات استماع کنگره و قوانین جدید است. با چنین تلاش‌هایی، سیاستمداران همچنان به محبوبیت احساسات ضد فناوری و ضد هوش مصنوعی در لحظه‌ای کمک می‌کنند که کشور فرصت و نیاز مبرمی برای تصمیم‌گیری در مورد چگونگی پذیرش فناوری هوش مصنوعی به روش‌هایی دارد که می‌تواند عاملیت فردی، رفاه ملی و امنیت ملی را همزمان تقویت کند. برای خدمت رسانی خوب به شهروندان خود در قرن بیست و یکم، ایالات متحده باید با یک قطب‌نمای فن‌سالارانه انسانی خود به جلو حرکت کند، نه صرفاً با چکش‌های تنظیم‌گری از بالا به پایین، دعاوی حقوقی و نظارت.

بازاندیشی دولت برای عصر دیجیتال: دولت ۲.۰ و مشارکت شهروندی

دیوید برنهام، خبرنگار تحقیقی نیویورک تایمز، در کتاب خود در سال ۱۹۸۳ با عنوان "ظهور دولت کامپیوتری" نوشت: "آیا ایالات متحده می‌تواند در عصری که حرکت فیزیکی، خریدهای فردی، مکالمات و جلسات هر شهروند به طور مداوم تحت نظارت شرکت‌های خصوصی و سازمان‌های دولتی است، به شکوفایی و رشد خود ادامه دهد؟ آیا نظارت، حتی از نوع بی‌گناه آن، روح یک ملت را به تدریج مسموم نمی‌کند؟ آیا نظارت، گزینه‌های شخصی بسیاری از شهروندان را محدود نمی‌کند؟" نزدیک به دو دهه پس از جلسات استماع کنگره در مورد مرکز داده ملی و هشت سال پس از انقلاب کامپیوترهای شخصی، شبح برادر بزرگ (Big Brother) همچنان در تصورات عمومی باقی مانده بود. یک سال پس از انتشار کتاب برنهام، سال واقعی ۱۹۸۴ فرا رسید و مکینتاش را به جای تلسکرین‌های اقیانوسیه به ما هدیه داد. و با این حال، چشم‌انداز تاریک اورول از یک دولت تکنولوژی‌محور همه‌چیزبین، همچنان تعریف ما را از جامعه‌ای که به طور فزاینده‌ای بر اساس تصمیم‌گیری‌های داده‌محور و قدرت رهایی‌بخش دانش مشترک سازماندهی

شده است، شکل داد. در دهه‌های هشتاد، نود و هزاره جدید، پذیرش جمعی کامپیوترهای شخصی، اینترنت و تلفن‌های همراه، شرکت‌هایی را که این ابزارها و خدمات جدید را تولید می‌کردند، به نهادهایی چنان قدرتمند تبدیل کرد که ما بسیاری از ترس‌های دیرینه از سرکوب و انطباق با فناوری را به آن‌ها منتقل کردیم. به زبان شوشانا زوبوف، "برادر بزرگ (Big Other) " جای "برادر بزرگ (Big Brother) "را به عنوان لولوی تکنولوژیکی حاکم بر فرهنگ گرفت.

اما اکنون چه اتفاقی می‌افتد زمانی که دولت‌ها در سراسر جهان، از جمله ایالات متحده، با ضرورت استفاده حداکثری از هوش مصنوعی و سایر فناوری‌های پیشرفته دست و پنجه نرم می‌کنند؟ هیچ کس تلسکرین‌های پلیس فکر، یا حتی "زنجیره‌های نوارهای پلاستیکی" را که ونس پاکارد در اواسط قرن معتقد بود مرکز داده ملی تولید خواهد کرد، نمی‌خواهد. اما آیا دولتی که به پتوهای دست‌باف متعهد است، جذابیت بیشتری دارد؟ به خصوص زمانی که دولت‌های دیگر هر کاری که می‌توانند انجام می‌دهند تا فناوری‌هایی را دنبال کنند که می‌تواند مزیت‌های اقتصادی، نظامی و استراتژیک قاطعی نسبت به بقیه جهان به آن‌ها بدهد؟

حتی در کشورهایی که شهروندان اخیراً برای برهم زدن انتقال مسالمت‌آمیز قدرت پس از یک انتخابات مشروع، به پایتخت یورش نبرده‌اند، مذاکره برای ایجاد تعادل بین توانایی تکنولوژیکی و ارزش‌های دموکراتیک یک چالش خواهد بود. در ایالات متحده، برای کسب رضایت حکومت‌شوندگان از همه اقشار جامعه‌ای که هم به شدت قطبی شده و هم به شدت درگیر است، به چابکی قابل توجهی نیاز خواهد بود. بنابراین دولت ما—که به معنای ما نیز هست—چگونه این چالش حیاتی را با موفقیت هدایت می‌کند؟ یک راه ممکن است تقلید از بهترین جنبه‌های نهادهایی باشد که مردم عموماً به آن‌ها احترام می‌گذارند. در حالی که نارضایتی اجتماعی از شرکت‌های بزرگ فناوری یک روایت رسانه‌ای مداوم است، واقعیت ظریف‌تر است. در نظرسنجی هاروارد CAPS/Harris که در می ۲۰۲۳ انجام شد، به عنوان مثال، پاسخ‌دهندگان آمازون را در میان بیست نهاد از نظر مطلوبیت بالاترین رتبه را دادند—بالاتر از ارتش ایالات متحده، دیوان عالی، ناتو و پلیس. گوگل در رتبه سوم این لیست قرار گرفت.

معیارهایی که اکثر پاسخ‌دهندگان نظرسنجی برای ارزیابی آمازون استفاده می‌کنند، بدون شک با معیارهایی که ممکن است برای قضاوت ناتو به کار ببرند، بسیار متفاوت است. در مورد آمازون، تجربه شخصی مکرر احتمالاً نقش دارد، و آن تجربیات شخصی شانس زیادی برای پاداش‌بخش بودن دارند. شما ممکن است از آمازون برای صرفه‌جویی در زمان و پول برای خرید مواد غذایی هفته خود استفاده کنید. ممکن است به عنوان منبع قابل اعتماد تصفیه‌کننده‌های هوای مقرون‌به‌صرفه، لوازم اداری عمده و فیلم‌های قهرمانان جریان‌یافته – ارضای تقریباً فوری در جبهه‌های متعدد – عمل کند. آخرین باری که ناتو چنین کاری برای شما انجام داد کی بود؟ به طور خلاصه، آمازون نمرات بالایی کسب می‌کند زیرا مردم با آن رابطه‌ای فعالانه دارند. آن‌ها اغلب از آن استفاده می‌کنند، و در بیشتر موارد، به آن اعتماد دارند. این شرکت ارزش واقعی به زندگی آن‌ها می‌آورد.

در قرن بیست و یکم، زمانی که همه ما می‌توانیم وزارت خزانه‌داری و اداره تامین اجتماعی را در جیب خود داشته باشیم، دسترسی و استفاده از خدمات دولتی باید به آسانی جستجو در گوگل یا خرید از آمازون باشد. در واقع، می‌توان استدلال کرد که معیار برای "تجربه‌های شهروندی" عالی باید حتی بالاتر از معیار برای تجربه‌های مشتری باشد، به دلیل نقش حیاتی که خدمات دولتی در زندگی ما ایفا می‌کنند. اگر به دلیل یک بیماری همه‌گیر جهانی یا یک طوفان ویرانگر از کار اخراج شده‌اید، خدمات و پشتیبانی به موقع فقط دلایلی برای یک امتیاز پنج ستاره نیستند، بلکه یک نجات‌بخش هستند.

بازنگری خدمات دولتی یک پروژه کوتاه‌مدت نیست. شاید با سیستم بزرگراه بین‌ایالتی به عنوان "بزرگترین برنامه کارهای عمومی در تاریخ جهان" رقابت نکند، اما یک تعهد بزرگ بلندمدت خواهد بود که نیازمند تعهدات بلندمدت است. هوش مصنوعی می‌تواند به شدت کمک کند. برخی کشورها، جایی که اجماع عمومی و اقدام قاطع دولت از پذیرش سریع هوش مصنوعی حمایت کرده است، از قبل مشغول تدوین سیاست‌هایی هستند که عملکرد هوش مصنوعی را در خدمات دولتی گنجانده‌اند. کره جنوبی، به عنوان مثال، برنامه‌هایی برای ادغام تقریباً ۱۵۰۰ خدمات

عمومی خود از طریق چندین وب‌سایت دارد و آن‌ها را از طریق یک پورتال در دسترس قرار می‌دهد، سپس از هوش مصنوعی برای اطلاع‌رسانی خودکار به شهروندان در مورد مزایا و حقوقی که واجد شرایط آن‌ها هستند، استفاده می‌کند.

چه می‌شود اگر ایالات متحده تعهدات مشابهی برای استقرار هوش مصنوعی به روش‌هایی که به وضوح برای شهروندان و عاملیت فردی – و حتی مهم‌تر، برای افزایش فرصت‌های مشارکت مدنی – مفید باشد، انجام دهد؟ چه می‌شود اگر دولت سپس از این تلاش‌ها به روشی که قبلاً در خدمات پستی ایالات متحده، سیستم بزرگراه بین‌ایالتی، رقابت فضایی و اینترنت سرمایه‌گذاری کرده بود، حمایت کند؟ از طریق رهبری آینده‌نگر، ما یک فرصت نسلی برای تقویت رفاه، امنیت و جایگاه جهانی آمریکا و شاید حتی متحد کردن یک جامعه قطبی‌شده با حس هدف ملی و اجماع ملی بزرگتر داریم. این برای قانونگذاران کشور ساده یا بدون درد نخواهد بود، زیرا پذیرش هوش مصنوعی خطرات سیاسی برای آن‌ها ایجاد خواهد کرد. به جای کنگره‌ای پر از وکلا با تخصص حقوقی خود، ما به قانونگذاران بیشتری با تخصص در فناوری و مهندسی نیاز خواهیم داشت. وقتی قانون کد است، ما به کدنویسان به همان اندازه که به وکلا در بالاترین سطوح دولت نیاز داریم، احتیاج داریم.

ما همچنین به مقامات منتخب نیاز داریم که درک کنند مردمی که به آن‌ها خدمت می‌کنند به عاملیت، انتخاب و راحتی‌ای که پیشرفت‌های تکنولوژیکی بیست و پنج سال گذشته به ارمغان آورده‌اند، عادت کرده‌اند. آن‌ها انتظارات بالاتری برای سطح خدمات و دسترسی می‌خواهند و برای نقش خود در این فرآیند. اینجاست که فرصت واقعاً بزرگی برای تشخیص آنچه هوش مصنوعی می‌تواند برای دولت و برای دموکراسی انجام دهد، وجود دارد.

بحث منطقی در مقیاس وسیع

به جای اینکه هوش مصنوعی را در وهله اول به عنوان مکانیزمی تصور کنیم که می‌تواند برای حکمرانی فرمان و کنترل، از طریق کاربردهایی مانند تشخیص چهره، پلیس پیش‌بینی‌کننده و

نظارت الگوریتمی به کار گرفته شود، می‌توانیم آینده‌ای را انتخاب کنیم که در آن هوش مصنوعی برای اتصال معنی‌دارتر شهروندان به فرآیندهای قانون‌گذاری استفاده شود. با استفاده از هوش مصنوعی برای ایجاد فرصت‌های جدید برای تقویت صدای شهروندان در سیاست‌گذاری، می‌توانیم سیستم خود را مشارکتی‌تر و تعاونی‌تر کنیم و در عین حال، نگرانی‌های دیرینه در مورد اقتدارگرایی با هوش مصنوعی را نیز کاهش دهیم.

یک جنبه کلیدی تلسکرین‌های رمان ۱۹۸۴ اورول وجود دارد که او یا نادیده گرفته یا انتخاب کرده است که آن را کاوش نکند. تلسکرین‌ها "دو طرفه" بودند. علاوه بر پخش تبلیغات دولتی، هر تلسکرین می‌توانست فضای نصب شده را به صورت بصری و شنیداری نیز نظارت کند. اما این واقعیت که می‌توان صدای شما را شنید، به این معنی است که می‌توان به شما گوش داد. با متعهد شدن دولت به گوش دادن به جای جاسوسی، تلسکرین و مهم‌تر از آن، میلیاردها صفحه نمایش متصل به اینترنت که جهان ما را شکل می‌دهند، دیگر دستگاه‌های نظارتی نیستند. آن‌ها دستگاه‌های ارتباطی هستند. با استفاده از هوش مصنوعی برای افزایش مشارکت مدنی و تصمیم‌گیری جمعی، دولت ما می‌تواند به طور قانع‌کننده‌ای نشان دهد که هدف آن دو برابر کردن نظارت و جاسوسی از شهروندان نیست – بلکه هدف این است که شهروندان خود را به طور مؤثرتری ببینند و بشنود.

این چشم‌انداز به طور گسترده با آنچه تیم اوریلی، نوآور فناوری، آن را "دولت ۲.۰" نامیده است، همسو و آن را گسترش می‌دهد – مدلی که در آن دولت به عنوان یک پلتفرم، تسهیل‌کننده و بانی اقدام مدنی عمل می‌کند و نه صرفاً یک ارائه‌دهنده خدمات و صادرکننده قوانین از بالا به پایین. در این الگو، هوش مصنوعی به یک توانمندساز قدرتمند تبدیل می‌شود و به دولت اجازه می‌دهد در سطحی از پاسخگویی که قبلاً غیرقابل تصور بود، عمل کند.

ابزاری به نام Remesh را در نظر بگیرید که سازمان ملل متحد آن را در مناطق درگیری برای ارزیابی سریع نیازها و نظرات جمعیت‌های آسیب‌دیده به کار گرفته است. از شرکت‌کنندگان سؤالاتی پرسیده می‌شود و به آن‌ها آزادی داده می‌شود تا به صورت باز پاسخ دهند، و این امر

اطلاعاتی بسیار دقیق‌تر و قابل اجراتر از آنچه با هر نظرسنجی چند گزینه‌ای می‌توان به دست آورد، استخراج می‌کند. هوش مصنوعی برای تقطیر و سازماندهی پاسخ‌ها استفاده می‌شود، به طوری که سؤالات پیگیری مرتبط را می‌توان در زمان واقعی پرسید، در حالی که هنوز توجه و همکاری شرکت‌کنندگان را دارید. تصور کنید، به عنوان مثال، آژانس مدیریت اضطراری فدرال (FEMA)از Remesh برای جمع‌آوری سریع بینش از جوامع آسیب‌دیده از بلایای طبیعی، به روش‌هایی که اثربخشی و ظرفیت آن را برای تطبیق تلاش‌های امدادی خود افزایش می‌دهد، استفاده می‌کند.

Polisابزار دیگری است که می‌تواند مشارکت عمومی و تصمیم‌گیری را متحول کند. این ابزار، که توسط یک سازمان غیرانتفاعی مستقر در سیاتل به نام پروژه دموکراسی محاسباتی مدیریت می‌شود، تا حدی پاسخی است به چالشی که جکلین تسای، وزیر دولت تایوان با تمرکز بر تجارت الکترونیک و سایر مسائل دیجیتال، در یک هکاتون در سال ۲۰۱۴ به توسعه‌دهندگان پیشنهاد کرد. او گفت: "ما به پلتفرمی نیاز داریم که به کل جامعه اجازه دهد در بحث منطقی شرکت کند ". Polisبه عنوان یک سیستم منبع باز برای تسهیل گفتگوهای در مقیاس بزرگ، به طور خاص برای یافتن زمینه‌های توافق در گروه‌های متنوع طراحی شده است. هر کسی می‌تواند از آن برای شروع گفتگوهای عمومی استفاده کند که معمولاً حول یک موضوع خاص مؤثر بر یک جامعه، مانند برنامه‌ریزی شهری، پیشنهادهای سیاستی یا تخصیص منابع، سازماندهی می‌شود.

تعداد نامحدودی از شهروندان می‌توانند به بحث بپیوندند. کاربران می‌توانند به صورت ناشناس شرکت کنند، اما با عدم وجود دکمه "پاسخ"، مکانیزمی برای توهین یا تبادل نظرات تند وجود ندارد. در عوض، افراد به سادگی با بیانیه‌های ۱۴۰ کاراکتری در مورد موضوعی که سایر شرکت‌کنندگان ارائه کرده‌اند، موافقت یا مخالفت می‌کنند. کاربران همچنین آزادند تا پیشنهاد جدیدی از خود ارائه دهند. به جای شمارش یک رأی ساده، هوش مصنوعی دیدگاه‌های مشابه را خوشه‌بندی می‌کند و زمینه‌های توافق را برجسته می‌سازد. سپس شرکت‌کنندگان می‌توانند

گروه‌های همفکر را مشاهده کنند که به صورت یک گرافیک ساده روی صفحه، نتایج در حال ظهور را ردیابی می‌کند. همانطور که شرکت‌کنندگان می‌بینند کدام دیدگاه‌ها و مواضع تأیید گسترده‌تری کسب می‌کنند، می‌توانند پیشنهادهای جدیدی را ارائه دهند که دیدگاه‌های مختلف را به هم وصل کند یا ایده‌های محبوب را اصلاح کند. این فرآیندی است که بر تعمق، تکرار، مصالحه و اجماع تأکید دارد.

یکی از موفقیت‌های اولیه آن در تایوان رخ داد، پس از آنکه پلتفرمی که توسط Polis پشتیبانی می‌شد، برای شکستن بن‌بست در مورد تنظیم Uber در آنجا استفاده شد. برای تعجب همه، جناح‌های طرفدار Uber و مخالف Uber در Polis به سرعت در مورد مسائلی مانند ایمنی مسافران به توافق رسیدند. در نهایت، پس از چندین دور مرتب‌سازی و رأی‌گیری با کمک هوش مصنوعی، گروهی از نظرات کاربران که تقریباً تأیید جهانی کسب کرده بودند، به عنوان اساس سیاست جدید دولت در مورد اشتراک‌گذاری سفر انتخاب شدند.

تصور کنید اگر کنگره ایالات متحده از یک سیستم مشابه برای جمع‌آوری نظرات عمومی در مورد قوانین مهم استفاده می‌کرد. این امر می‌تواند منجر به قوانینی شود که دقیق‌تر منعکس‌کننده خواست مردم باشند، به طور بالقوه اعتماد عمومی به دولت را افزایش داده و قطبی‌سازی سیاسی را کاهش می‌دهد. با ارائه داده‌های غنی‌تر و نماینده‌تر در مورد افکار عمومی به قانونگذاران، این و سیستم‌های هوش مصنوعی مشابه می‌توانند به تضمین این امر کمک کنند که قوانین بهتر نیازها و ترجیحات واقعی رأی‌دهندگان را برطرف کنند.

این تغییر به سیاست‌گذاری پاسخگوتر و مشارکتی‌تر می‌تواند به کاهش حس بیگانگی و سلب قدرت که بسیاری از مردم امروز در مورد سیاست احساس می‌کنند، کمک کند و می‌تواند این ایده را تقویت کند که شهروندان فردی ذینفعان واقعی در دموکراسی ما هستند. با گرد هم آوردن شهروندان و دولت آن‌ها به صورت آنلاین، هوش مصنوعی می‌تواند به نیرویی سازنده برای بهبود انسجام اجتماعی و التیام اختلافات ما تبدیل شود. به جای چارچوب‌بندی مخالفان سیاسی به

عنوان کاریکاتورهایی برای تمسخر، پلتفرم‌های تصمیم‌گیری جمعی هر شهروند را به عنوان یک متحد بالقوه برای تعامل معنی‌دار معرفی می‌کنند. آن‌ها دنیایی را ممکن می‌سازند که در آن فناوری به جای جایگزینی عاملیت انسانی، آن را تقویت می‌کند، و در آن وعده دولتی "از مردم، توسط مردم، برای مردم" برای عصر دیجیتال تجدید و تقویت می‌شود.

این یک چشم‌انداز جاه‌طلبانه و خوش‌بینانه است. قانونگذاران را به پاسخگوتر بودن در برابر حوزه‌های انتخابی خود و شاید حتی واگذاری بخشی از قدرت خود به شهروندانی که به آن‌ها خدمت می‌کنند، به چالش می‌کشد. شهروندان را به تعامل با حسن نیت با سایر شهروندانی که لزوماً با آن‌ها موافق نیستند و حتی ممکن است به شدت مخالفت کنند، به چالش می‌کشد. سیاست را به عنوان قلمرویی تصور می‌کند که بیشتر با تصمیم‌گیری عمل‌گرایانه مشخص می‌شود تا بیان هویت‌های حزبی.

تقویت عاملیت فردی و رفاه اجتماعی از طریق هوش مصنوعی

برای اطمینان از بهترین آینده ممکن برای خودمان، مهم است که به این چشم‌انداز دست یابیم. هوش مصنوعی در حال تغییر نحوه عملکرد جهان است. این واقعیت که بسیاری از کشورها در حال حاضر آن را پذیرفته‌اند، به این معنی است که همه کشورها در نهایت باید با تأثیری که هوش مصنوعی ایجاد می‌کند، کنار بیایند. بهترین آینده ممکن ما با تصور هوش مصنوعی به عنوان "گسترشی از اراده‌های انسانی فردی" آغاز می‌شود. نحوه انتخاب ما برای تعامل جمعی با هوش مصنوعی، به عنوان یک ملت، بر نحوه تعامل فردی ما با آن تأثیر می‌گذارد. بنابراین، اجماع در حال حاضر از اهمیت حیاتی برخوردار است. هدف مشترک ضروری است.

به بیان ساده، هرچه ما، به عنوان یک ملت، بیشتر به هوش مصنوعی متعهد شویم، احتمالاً هر فرد نیز بیشتر بهره‌مند خواهد شد. رویکردهای نظارتی مولد منجر به سیستم‌های بهتر و ایمن‌تر، با سرعت بیشتری خواهد شد. زیرساخت‌های عمومی، که در آن هوش مصنوعی در محیط‌های فیزیکی و دیجیتالی که هر روز با آن‌ها در تعامل هستیم گنجانده می‌شود، زندگی را آسان‌تر و پربارتر

خواهد کرد.

برعکس، هرچه هر فرد بیشتر از هوش مصنوعی بهره‌مند شود، ما به طور جمعی بیشتر سود خواهیم برد. دنیایی که در آن همه مراقبت‌های بهداشت روان مورد نظر خود را دریافت می‌کنند، دنیایی عادلانه‌تر و انسانی‌تر است، حتی اگر برای دستیابی به آن به درمانگران مبتنی بر هوش مصنوعی نیاز باشد. دنیایی که در آن هر فرد به معلمان مجازی، مشاوران حقوقی مجازی و هر آنچه نیاز دارد دسترسی دارد، دنیایی است که در آن هر کس شانس بهتری برای تبدیل شدن به بهترین نسخه از خود را دارد – و مزایای آن به همه ما می‌رسد. دنیایی که در آن هر دانشمندی با فرضیه‌ای جذاب اما غیرمتعارف (و بنابراین دشوار برای تأمین مالی) می‌تواند از هوش مصنوعی برای اجرای شبیه‌سازی‌های پیچیده، تجزیه و تحلیل مجموعه‌های داده وسیع و اعتباربخشی نظریه‌ها استفاده کند، دنیایی است که سرعت کشف را به روش‌هایی که به نفع همه ما است، تسریع می‌کند. دنیایی که در آن هر کارآفرین جاه‌طلبی، صرف‌نظر از پیشینه یا منابع، می‌تواند به ابزارهای تحلیل بازار و مدل‌سازی مالی مبتنی بر هوش مصنوعی دسترسی داشته باشد، دنیایی است که در آن نوآوری شکوفا می‌شود و فرصت‌های اقتصادی به طور عادلانه‌تری توزیع می‌شوند – و اثرات رفاهی را در کل جامعه ایجاد می‌کند. این دنیایی است که با "فرا عاملیت" (Superagency) امکان‌پذیر می‌شود و ما در حال حاضر شروع به دیدن خطوط کلی آن به روش‌های واضح و امیدوارکننده هستیم.

نتیجه‌گیری

تغییرات تکنولوژیک همواره دو روی سکه داشته‌اند: وعده پیشرفت و خطر از دست دادن آنچه آشنا و باارزش است. درس‌های تاریخی از لودیت‌ها و سناریوی جایگزین انگلستان نشان می‌دهد که مقاومت یا احتیاط مفرط در برابر نوآوری‌های تحول‌آفرین می‌تواند به انزوای ملی، رکود اقتصادی و از دست دادن عاملیت فردی در بلندمدت منجر شود. در مقابل، پذیرش استراتژیک و مسئولانه فناوری، اگرچه با چالش‌هایی همراه است، اما می‌تواند مسیر را برای رفاه بی‌سابقه، پیشرفت اجتماعی و تقویت هویت ملی هموار کند.

در عصر حاضر، هوش مصنوعی به عنوان نیروی محرک اصلی در صحنه جهانی ظاهر شده و رقابت فزاینده‌ای را بین کشورها برای توسعه و کنترل زیرساخت‌های آن دامن زده است. مفهوم هوش مصنوعی ملی (Sovereign AI) بر اهمیت حیاتی کنترل کشورها بر داده‌ها و فناوری هوش مصنوعی خود تأکید می‌کند تا از حریم خصوصی، امنیت ملی و بازتاب ارزش‌ها و فرهنگ خود در این فناوری‌های قدرتمند اطمینان حاصل کنند. ایالات متحده، به عنوان پیشتاز در نوآوری هوش مصنوعی، باید با دقت بین حمایت از قهرمانان تکنولوژیکی خود و اعمال تنظیم‌گری‌های هوشمندانه تعادل برقرار کند تا هم نوآوری را تشویق کرده و هم از سوءاستفاده‌ها جلوگیری نماید.

فراتر از تأثیرات اقتصادی و امنیتی، هوش مصنوعی پتانسیل عظیمی برای بازاندیشی دولت و خدمات عمومی دارد. با رویکرد "دولت ۲.۰"، هوش مصنوعی می‌تواند به ابزاری برای افزایش شفافیت، کارایی و به ویژه، مشارکت شهروندی تبدیل شود. نمونه‌هایی مانند Remesh و Polis نشان می‌دهند که چگونه هوش مصنوعی می‌تواند به دولت‌ها کمک کند تا به طور مؤثرتر به صدای شهروندان گوش دهند، اجماع را در مسائل پیچیده تسهیل کنند و سیاست‌هایی را تدوین کنند که واقعاً منعکس‌کننده نیازها و آرزوهای مردم باشد. این رویکرد می‌تواند به کاهش قطبی‌سازی سیاسی، افزایش اعتماد عمومی و تقویت بنیان‌های دموکراتیک کمک کند.

در نهایت، هوش مصنوعی وعده تقویت بی‌سابقه عاملیت فردی و رفاه جمعی را می‌دهد. از

بهبود مراقبت‌های بهداشتی و آموزش گرفته تا فرصت‌های جدید برای کارآفرینی و کشف علمی، هوش مصنوعی می‌تواند به هر فردی این امکان را بدهد که به بهترین نسخه از خود تبدیل شود و این مزایا به طور فراگیر به کل جامعه باز می‌گردد. برای دستیابی به این آینده، نیاز به یک اجماع ملی، هدف مشترک و رهبری آینده‌نگرانه است که هوش مصنوعی را نه تنها به عنوان یک ابزار، بلکه به عنوان یک نیروی تحول‌آفرین برای خیر جمعی بشناسد. شکل‌دهی به یک آمریکای هوشمند، مستلزم تعهد جمعی به نوآوری مسئولانه، حکمرانی مشارکتی و سرمایه‌گذاری در آینده‌ای است که در آن فناوری به توانمندسازی انسان‌ها کمک می‌کند.

نکات کلیدی

- **درس‌های تاریخی** :مقاومت در برابر پیشرفت تکنولوژیک، همانند جنبش لودیت‌ها و سناریوی جایگزین انگلستان، می‌تواند منجر به انزوای ملی و رکود اقتصادی شود.
- **مزایای پذیرش فناوری** :کشورهایی که نوآوری تکنولوژیک را پذیرفته‌اند، شاهد رفاه، گزینه‌های چشمگیر و افزایش عاملیت فردی بوده‌اند.
- **اهمیت هوش مصنوعی ملی** :(Sovereign AI) کنترل کشورها بر زیرساخت‌های هوش مصنوعی خود برای حفظ حریم خصوصی داده‌ها، امنیت ملی و هویت فرهنگی ضروری است.
- **رقابت جهانی هوش مصنوعی** :توسعه هوش مصنوعی یک رقابت جهانی گسترده است که نیازمند اجماع فرهنگی و سیاسی برای موفقیت است.
- **چالش‌های ایالات متحده** :ایالات متحده باید بین حمایت از نوآوری شرکت‌های بزرگ فناوری و اعمال تنظیم‌گری برای جلوگیری از سوءاستفاده تعادل برقرار کند.
- **لزوم رویکرد فن‌سالارانه انسانی** :دولت آمریکا نیاز به یک رویکرد هوشمند دارد که هوش مصنوعی را برای تقویت عاملیت فردی، رفاه ملی و امنیت ملی به کار گیرد.
- **نقش هوش مصنوعی در دولت ۲.۰** :هوش مصنوعی می‌تواند خدمات دولتی را بهبود بخشد و به ابزاری برای تقویت مشارکت شهروندی و فرآیندهای دموکراتیک

تبدیل شود.

- **ابزارهای مشارکتی هوش مصنوعی** :پلتفرم‌هایی مانند Remesh و Polis توانایی تقویت بحث منطقی، یافتن اجماع و بهبود سیاست‌گذاری را دارند.
- **افزایش عاملیت فردی** :هوش مصنوعی می‌تواند فرصت‌های فردی را در حوزه‌هایی مانند مراقبت‌های بهداشتی، آموزش و کارآفرینی به طور چشمگیری افزایش دهد.
- **منافع جمعی هوش مصنوعی** :پذیرش گسترده هوش مصنوعی منجر به بهبود مراقبت‌های بهداشت روان، آموزش فراگیر و تسریع کشفیات علمی می‌شود.

سؤالات تفکربرانگیز

۱. با توجه به درس‌های تاریخ از جنبش لودیت‌ها و سناریوی جایگزین انگلستان، چه تعادلی بین نوآوری تکنولوژیک و حفظ سنت‌های فرهنگی و ساختارهای اجتماعی باید برقرار شود؟

۲. مفهوم "هوش مصنوعی ملی (Sovereign AI) "چه چالش‌ها و فرصت‌های جدیدی را برای دیپلماسی بین‌المللی و روابط قدرت جهانی ایجاد می‌کند؟

۳. چگونه ایالات متحده می‌تواند یک "قطب‌نمای فن‌سالارانه انسانی" را در سیاست‌گذاری‌های هوش مصنوعی خود توسعه دهد که هم نوآوری را تشویق کند و هم حقوق و آزادی‌های شهروندان را حفظ نماید؟

۴. در یک سیستم "دولت ۲.۰" که هوش مصنوعی نقش محوری دارد، چه مکانیزم‌هایی برای اطمینان از پاسخگویی دولت و جلوگیری از سوءاستفاده از قدرت باید وجود داشته باشد؟

۵. پلتفرم‌هایی مانند Polis و Remesh چگونه می‌توانند در جوامع عمیقاً قطبی‌شده، اجماع را تسهیل کنند و چه موانعی ممکن است در پذیرش گسترده آن‌ها در فرآیندهای قانون‌گذاری وجود داشته باشد؟

۶. چگونه می‌توان اطمینان حاصل کرد که مزایای "فرا عاملیت" که توسط هوش مصنوعی امکان‌پذیر می‌شود، به طور عادلانه در سراسر جامعه توزیع می‌شود و شکاف‌های موجود را تشدید نمی‌کند؟

۳۰ جمله کلیدی

۱. داستان حزب دونر و لودیت‌ها، دو رویکرد متضاد در برابر تغییرات تکنولوژیک را به نمایش می‌گذارد.

۲. لودیت‌ها در اوایل قرن نوزدهم، به مکانیزاسیون صنعت نساجی انگلستان واکنش خشونت‌آمیزی نشان دادند.

۳. آن‌ها نه تنها با ماشین‌آلات، بلکه با سیستم کارخانه‌ای استثمارگر و از دست دادن آزادی فردی مقابله می‌کردند.

۴. دولت بریتانیا با تصویب قانون "شکستن چارچوب" و مجازات‌های سنگین، لودیت‌ها را سرکوب کرد.

۵. در یک تاریخ جایگزین، قانون "شغل، ایمنی و کرامت انسانی (JSHDA)" مانع از پیشرفت تکنولوژیک در انگلستان شد.

۶. انگلستان در این سناریو، به یک جامعه سنتی و غیرصنعتی تبدیل شد، در حالی که سایر کشورها صنعتی شدند.

۷. این رویکرد احتیاطی، منجر به رکود اقتصادی، از دست دادن بازارهای صادراتی و مهاجرت مغزها از انگلستان شد.

۸. فناوری‌هایی که در ابتدا به عنوان تهدید تلقی می‌شوند، اغلب به ابزارهای انسانی‌سازی و رهایی‌بخش تبدیل می‌شوند.

۹. هوش مصنوعی پتانسیل تحول‌آفرینی عظیم در سطوح فردی، ملی و جهانی را دارد.

۱۰. کشوری که در پذیرش هوش مصنوعی کند عمل کند، در حوزه‌هایی مانند

مراقبت‌های بهداشتی و کشاورزی عقب خواهد ماند.

۱۱. قدرت مدل‌های هوش مصنوعی طی دهه آینده می‌تواند هزاران برابر افزایش یابد و قابلیت‌های جدیدی ایجاد کند.

۱۲. رقابت جهانی هوش مصنوعی دیگر محدود به ایالات متحده و چین نیست و به یک مسابقه گسترده‌تر تبدیل شده است.

۱۳. جنسن هوانگ، تولید هوش را انقلاب صنعتی جدید می‌داند و بر مالکیت ملی داده‌ها تأکید می‌کند.

۱۴. مفهوم "هوش مصنوعی ملی (Sovereign AI) "بر کنترل کشورها بر زیرساخت‌های هوش مصنوعی خود تمرکز دارد.

۱۵. این رویکرد، مسائل حریم خصوصی داده‌ها، امنیت ملی و حفظ فرهنگ را در برابر وابستگی به ارائه‌دهندگان خارجی حل می‌کند.

۱۶. سنگاپور و فرانسه استراتژی‌های هوش مصنوعی خود را برای بازتاب فرهنگ و ارزش‌های بومی تدوین کرده‌اند.

۱۷. ایالات متحده با چالش تعادل بین ترویج نوآوری توسط شرکت‌های بزرگ فناوری و تنظیم‌گری از طریق اقدامات ضد انحصار روبرو است.

۱۸. برای آینده، آمریکا به یک "قطب‌نمای فن‌سالارانه انسانی" نیاز دارد تا هوش مصنوعی را به نفع عاملیت فردی و امنیت ملی به کار گیرد.

۱۹. ترس‌های تاریخی از نظارت دولتی ("برادر بزرگ") اکنون به نگرانی‌ها در مورد قدرت شرکت‌های بزرگ فناوری ("برادر بزرگ دیگر") منتقل شده است.

۲۰. دولت‌ها در سراسر جهان باید از هوش مصنوعی برای به حداکثر رساندن کارایی و حفظ مزیت استراتژیک استفاده کنند.

۲۱. ارتقاء خدمات دولتی برای شهروندان، با الگوبرداری از تجربیات کاربری پلتفرم‌هایی

مانند آمازون و گوگل، ضروری است.

۲۲. هوش مصنوعی می‌تواند به دولت‌ها کمک کند تا خدمات عمومی را متحول کنند، مانند برنامه کره جنوبی برای ادغام ۱۵۰۰ سرویس.

۲۳. مشارکت شهروندی فعال در فرآیندهای قانون‌گذاری می‌تواند با استفاده از هوش مصنوعی تقویت شود.

۲۴. "دولت ۲.۰" مدلی است که در آن دولت به عنوان پلتفرم، تسهیل‌کننده و بانی اقدام مدنی عمل می‌کند.

۲۵. ابزارهایی مانند Remesh و Polis امکان بحث منطقی در مقیاس وسیع و یافتن اجماع در گروه‌های متنوع را فراهم می‌کنند.

۲۶. Polis در تایوان برای حل بن‌بست‌های سیاستی مانند تنظیم Uber با موفقیت به کار گرفته شده است.

۲۷. استفاده از هوش مصنوعی برای جمع‌آوری نظرات عمومی می‌تواند منجر به قوانینی شود که دقیق‌تر منعکس‌کننده خواست مردم باشند.

۲۸. هوش مصنوعی پتانسیل افزایش عاملیت فردی در حوزه‌هایی مانند مراقبت‌های بهداشتی، آموزش و کارآفرینی را دارد.

۲۹. دنیایی که در آن همه از هوش مصنوعی بهره‌مند می‌شوند، دنیایی عادلانه‌تر، انسانی‌تر و پر از فرصت‌های اقتصادی بیشتر خواهد بود.

۳۰. تعهد ملی به هوش مصنوعی می‌تواند منجر به سیستم‌های بهتر و ایمن‌تر، افزایش بهره‌وری و یکپارچگی اجتماعی شود.

منابع

INTRODUCTION

1 https://www.smithsonianmag.com/innovation/texting-isnt-first-new-technology-thought-impair-social-skills-180958091/; Clive Thompson, "Texting Isn't the First New Technology Thought to Impair Social Skills," Smithsonian Magazine, March 2016.

2 Clay McShane, Down the Asphalt Path: The Automobile and the American City (New York: Columbia University Press, 1994), 133.

3 https://time.com/archive/6624989/business-the-automation-jobless/.

4 Plato, Phaedrus, Translated, with Introduction and Notes, by Alexander Nehamas and Paul Woodruff (Indianapolis, Indiana: Hackett Publishing Company, 1995), 80.

CHAPTER 1

1 https://www.weforum.org/agenda/2022/10/global-concern-inflation-energy-economy/.

2 https://www.cnbc.com/2022/12/02/jobs-report-november-2022.html.

3 https://twitter.com/OpenAI/status/1598014522098208769.

4 https://twitter.com/sama/status/1598038817126027264?lang=en.

5 https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

6 https://www.theverge.com/2023/2/27/23617477/mark-zuckerberg-meta-ai-tools-personas.

7 https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

8 https://www.ipsos.com/sites/default/files/ct/news/documents/2022-01/Global-opinions-and-expectations-about-AI-2022.pdf.

9 https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/.

10 https://www.monmouth.edu/polling-institute/reports/monmouth-poll_us_021523/.

11 https://openai.com/blog/introducing-openai.

12 https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

13 https://civilrightsdocs.info/pdf/FINAL_JointStatementPredictivePolicing.pdf.

14 https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html.

CHAPTER 2

1 https://time.com/archive/6808108/technology-the-brainy-breed/.

2 Robert Gannon, "Big-Brother 7074 Is Watching You," Popular Science, March 1963, https://web.archive.org/web/20200119162242/http://blog.modernmechanix.com/big-brother-7074-is-watching-you/.

3 D. S. Halacy Jr., Computers: The Machines We Think With (New York: Harper & Row, 1962).

4 Vance Packard, The Naked Society (New York: Ig Publishing, 2014), 58.

5 https://www.census.gov/history/pdf/kraus-natdatacenter.pdf.

6 Myron Brenton, The Privacy Invaders (New York: Coward-McCann, 1964), 157.

7 Sarah E. Igo, The Known Citizen: A History of Privacy in Modern America (Cambridge, MA: Harvard University Press, 2020), 147.

8 Igo, 146.

9
https://www.google.com/books/edition/Weekly_Compilation_of_Presidential_Docum/0zVI4Wq7jCYC?hl=en&gbpv=1.
10 https://www.census.gov/history/pdf/kraus-natdatacenter.pdf.
11
https://archive.org/stream/U.S.House1966TheComputerAndInvasionOfPrivacy/U.S.%20House%20%281966%29%20-%20The%20Computer%20and%20Invasion%20of%20Privacy_djvu.txt.
12
https://archive.org/stream/U.S.House1966TheComputerAndInvasionOfPrivacy/U.S.%20House%20%281966%29%20-%20The%20Computer%20and%20Invasion%20of%20Privacy_djvu.txt.
13
https://www.google.com/books/edition/Congressional_Record/w7XMSn2bsl8C?hl=en&gbpv=1&dq=packard+%2B+filekeepers+%2B+derogatory&pg=PA19964&printsec=frontcover.
14
https://archive.org/stream/U.S.House1966TheComputerAndInvasionOfPrivacy/U.S.%20House%20%281966%29%20-%20The%20Computer%20and%20Invasion%20of%20Privacy_djvu.txt.
15
https://www.google.com/books/edition/Congressional_Record/HILK33dxgDMC?hl=en&gbpv=1&dq=%22Big+Brother+Never+Rests%22&pg=PA28691&printsec=frontcover.
16 https://supreme.justia.com/cases/federal/us/381/479/#opinions.
17 https://blog.hubspot.com/marketing/visual-history-of-apple-ads.

18 https://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html .

19 https://www.webdesignmuseum.org/gallery/linkedin-2003.

20 https://news.linkedin.com/about-us.

21 Andr a Belliger and David J. Krieger, Network Publicy Governance: On Privacy and the Informational Self (Bielefeld: transcript publishing, 2018), 8.

22 https://people.ischool.berkeley.edu/~hal/Papers/japan/japan.html.

23 Ithiel De Sola Pool et al., Communication Flows: A Census in the United States and Japan (North Holland; Amsterdam; New York; Oxford: University of Tokyo Press, 1984), 16.

CHAPTER 3

1 https://x.com/RobertRMorris/status/1611450210915434499.

2 https://gizmodo.com/mental-health-therapy-app-ai-koko-chatgpt-rob-morris-1849965534.

3 https://gizmodo.com/mental-health-therapy-app-ai-koko-chatgpt-rob-morris-1849965534.

4 https://time.com/6308096/therapy-mental-health-worse-us/.

5 https://www.cnn.com/2023/11/29/health/suicide-record-high-2022-cdc/index.html#:~:text=At%20least%2049%2C449%20lives%20were,deaths%20for%20every%20100%2C000%20people.

6 https://www.npr.org/2023/05/18/1176830906/overdose-death-2022-record#:~:text=April%2018%2C%202022.-,The%20latest%20federal%20data%20show%20more%20than%20109%2C000,in%202022%2C

٪20many٪20from٪20fentanyl.&text=Drug٪20deaths٪20nationwide٪20hit٪20a,for٪20Disease٪20Control٪20and٪20Prevention.

7 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8601097/.

8 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4102288/.

9 https://websites.umich.edu/~daneis/symposium/2010/ARTICLES/eisenberg_golberstein_hunt_2009.pdf.

10 https://www.gallup.com/workplace/404174/economic-cost-poor-employee-mental-health.aspx.

11 https://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf.

12 https://www.washingtonpost.com/wellness/2022/10/29/therapists-waiting-lists-depression-anxiety/.

13 https://journals.sagepub.com/doi/10.1177/1357633X14524156#bibr1-1357633X14524156.

14 https://psyche.co/guides/how-to-choose-a-mental-health-app-that-can-actually-help.

15 https://calmatters.org/projects/californians-struggle-to-get-mental-health-care/.

16 https://www.behavioralhealthworkforce.org/wp-content/uploads/2019/02/Y3-FA2-P2-Psych-Sub_Full-Report-FINAL2.19.2019.pdf.

17 https://www.chcf.org/publication/2019-california-health-policy-survey/.

18 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7293059/.

19 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10267322/.

20 https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2814116.

21 https://www.nature.com/articles/s44184-024-00056-z.

22 https://www.newscientist.com/article/mg25834340-900-how-do-we-know-that-therapy-works-and-which-kind-is-best-for-you/.

23 https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/2804309.

CHAPTER 4

1 https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/.

2 https://www.newyorker.com/science/annals-of-artificial-intelligence/will-ai-become-the-new-mckinsey.

3 Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power (New York: PublicAffairs, 2020), 20.

4 Zuboff, 69.

5 Zuboff, 74.

6 Zuboff, 74–75.

7 Zuboff, 377.

8 https://arxiv.org/pdf/2005.14165. There is a distinction between "training on X tokens" and "training for X tokens." The former refers to the number of unique tokens in a given dataset, indicating that the dataset contains X tokens. The latter describes the total amount of data processed during the training phase, which includes multiple passes over the dataset. In other words, "training for X tokens" implies that the model was trained by processing a smaller dataset multiple times until the total number of tokens processed reached X.

9 https://www.commoncrawl.org/blog/june-2024-crawl-archive-now-available; https://commoncrawl.github.io/cc-crawl-statistics/plots/domains.html.

10 https://en.wikipedia.org/wiki/The_Pile_(dataset).

11 https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/.

12 https://www.smithsonianmag.com/history/what-17th-century-ideal-commons-means-21st-century-180973240/.

13 https://www.forbes.com/billionaires/.

14 https://hbr.org/2019/11/how-should-we-measure-the-digital-economy.

15 https://ide.mit.edu/wp-content/uploads/2018/04/w24514.pdf?x57209.

16 https://www.aei.org/economics/googlenomics-a-long-read-qa-with-chief-economist-hal-varian/.

17 https://www.nytimes.com/2006/12/24/opinion/24rich.html.

18 Paul R. Ehrlich, The Population Bomb (New York: Ballantine, 1968).

19 Garrett Hardin, "The Tragedy of the Commons," Science, New Series 162, no. 3859 (December 1968), 1243–48;
https://pages.mtu.edu/~asmayer/rural_sustain/governance/Hardin%201968.pdf.

20 https://www.statista.com/statistics/234056/facebooks-average-advertising-revenue-per-user/.

CHAPTER 5

1 https://twitter.com/TIME/status/1626190434530693123.

2 https://www.cnn.com/2023/02/06/media/google-microsoft-ai-reliable-sources/index.html.

3 https://www.axios.com/2023/01/31/chatgpt-ai-arms-race-ethics-competition.

4 https://www.newsweek.com/efforts-avoid-suicidal-ai-arms-race-are-failing-scientists-warn-1826622.

5 https://www.bloomberg.com/news/videos/2023-07-13/only-winner-of-an-ai-arms-race-ai-ai-irl-episode-12-video.

6 https://arxiv.org/pdf/1905.00537.pdf.

7 https://towardsdatascience.com/the-olympics-of-ai-benchmarking-machine-learning-systems-c4b2051fbd2b

8 https://openai.com/research/instruction-following.

9 Keep in mind that RealToxicityPrompts explicitly tries to get a model to produce toxic outputs by asking it to complete prompts like "The men started swearing at me, called me . . ." and "So I'm starting to think she's full . . ." See https://arxiv.org/pdf/2009.11462.pdf.

10 https://cdn.openai.com/papers/gpt-4.pdf.

11 https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

12 https://arxiv.org/pdf/2303.12712.

13 https://www.nature.com/immersive/d41586-023-02822-z/index.html.

14 https://www.science.org/doi/10.1126/science.adk$176.

15 https://hwpi.harvard.edu/files/datasmart/files/regulation_the_internet_way.pdf.

CHAPTER 6

1 https://www.nytimes.com/2023/01/23/opinion/ted-lieu-ai-chatgpt-congress.html.

2 https://www.theguardian.com/technology/2023/jun/05/ai-could-outwit-humans-in-two-years-says-uk-government-adviser.

3 https://www.theregister.com/2023/06/06/netherlands_minister_asks_big_tech/.

4 https://www.rstreet.org/commentary/the-most-important-principle-for-ai-regulation/.

5 https://www.nytimes.com/2001/12/09/magazine/the-year-in-ideas-a-to-z-precautionary-principle.html.

6 https://www.gmfus.org/news/acid-rain-lessons-germanys-black-forest.

7 https://www.wired.com/2017/05/san-francisco-wants-ban-delivery-robots-squash-someones-toes/.

8 https://sf.curbed.com/2017/12/6/16743326/san-francisco-delivery-robot-ban.

9 https://rules.cityofnewyork.us/wp-content/uploads/2021/08/DOT-Notice-of-Adoption-AV-Rule-FINAL-with-Finding.pdf.

10 Permissionless Innovation, loc. 317.

11 https://archive.nytimes.com/www.nytimes.com/library/cyber/week/070297commerce.html.

12 https://clintonwhitehouse4.archives.gov/WH/New/Commerce/read.html.

13 https://www.wsj.com/articles/SB928362554894823220.

14 https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

15 https://openai.com/index/our-approach-to-ai-safety/.

16 John B. Rae, The Road and the Car in American Life (Cambridge, MA: MIT Press, 1971), 51.

17 https://npgallery.nps.gov/NRHP/GetAsset/NHLS/73000961_text.

18 https://www.conceptcarz.com/vehicle/z13236/ford-model-t.aspx.

19 William Greenleaf, "Henry Ford," Dictionary of American Biography, Supplement Four, 1946–1950 (New York: Scribner, 1974), 295.

20 Peter D. Norton, Fighting Traffic: The Dawn of the Motor Age in the American City (Cambridge, MA: MIT Press, 2008), loc. 335.

21 Clay McShane, Down the Asphalt Path: The Automobile and the American City (New York: Columbia University Press, 1994), 43, 51.

22 McShane, 97.

23 https://www.wired.com/2008/05/dayintech-0521/.

24 William Phelps Eno, The Story of Highway Traffic Control, 1899–1939 (Saugatuck, CT: Eno Foundation for Highway Traffic Control, 1939), vii.

25 https://www.vox.com/2015/8/5/9097713/when-was-the-first-traffic-light-installed.

26 https://corporate.ford.com/articles/history/henry-ford-biography.html.

27 https://corporate.ford.com/articles/history/1901-sweepstakes-race.html.

28 https://exchange.aaa.com/wp-content/uploads/2012/10/AAA-Glidden-History.pdf.

29 "Studebakers in Annual Test," Los Angeles Times, August 23, 1925, H8.

30 Eno, The Story of Highway Traffic Control, vii.

31 https://www.google.com/books/edition/Motor_Field/8cBw3GJJfxwC?hl=en&gbpv=1&dq=darlington+%2B+%22Royal+Tourist%22+%2B+Glidden+%2B+shot&pg=RA2-PA38&printsec=frontcover.

32 https://press.uchicago.edu/Misc/Chicago/467412.html.

33 https://injuryfacts.nsc.org/motor-vehicle/historical-fatality-trends/deaths-and-rates/.

CHAPTER 7

1 https://aerospace.org/article/brief-history-gps#:~:text=In%20February%201978%2C%20the%20first,space%2Fcontrol%2Fuser%20system.

2 https://www.popularmechanics.com/technology/gadgets/a26980/why-the-military-released-gps-to-the-public/.

3 https://washingtontechnology.com/1996/04/clinton-lifts-gps-barriers-boosts-industry/334467/.

4 https://www.nist.gov/system/files/documents/2020/02/06/gps_finalreport618.pdf.

5 Greg Milner, Pinpoint (New York: W. W. Norton, 2017), 58.

6 https://economics.mit.edu/sites/default/files/inline-files/Noy_Zhang_1.pdf.

7 Erik Brynjolfsson et al., "Generative AI at Work," National Bureau of Economic Research, April 2023, revised November 2023, https://www.nber.org/system/files/working_papers/w31161/w31161.pdf.

8 https://www.nature.com/articles/d41586-023-02270-9.

9 https://illuminate.google.com/home?pli=1.

10 Ethan Mollick, "Latent Expertise: Everyone Is in R&D," One Useful Thing, June 20, 2024, https://www.oneusefulthing.org/p/latent-expertise-everyone-is-in-r.

CHAPTER 8

1 This quote is taken from an essay Lessig published in Harvard Magazine in January 2000, in coordination with the publication of Code, and Other Laws of Cyberspace.

2 Lawrence Lessig, Code, and Other Laws of Cyberspace (New York: Basic Books, 1999), 6.

3 https://www.nhtsa.gov/sites/nhtsa.gov/files/2023-12/anprm-advanced-impaired-driving-prevention-technology-2127-AM50-web-version-12-12-23.pdf.

4 Lessig, Code, 235.

5 Lessig, 208.

6 Lessig, 237.

7 Zuboff, The Age of Surveillance Capitalism, 220.

8 Samer Hassan and Primavera De Filippi, "The Expansion of Algorithmic Governance: From Code Is Law to Law Is Code," Field Actions Science Reports, Special Issue 17 (2017): 88–90, https://journals.openedition.org/factsreports/4518

9 Lessig, Code, 200.

10 Lessig, 203.

11 https://www.archives.gov/founding-docs/declaration-transcript

CHAPTER 9

1 Rae, The Road and the Car in American Life, 50.

2 Mustafa Suleyman with Michael Bhaskar, The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma (New York: Crown, 2023), 164.

3 https://www.statista.com/statistics/1104709/coronavirus-deaths-worldwide-per-million-inhabitants/.

4 https://www.nytimes.com/2020/03/23/world/asia/coronavirus-south-korea-flatten-curve.html.

5 https://www.rand.org/pubs/commentary/2022/09/can-south-korea-help-the-world-beat-the-next-pandemic.html.

6 https://www.csis.org/analysis/timeline-south-koreas-response-covid-19.

7 https://thediplomat.com/2020/12/covid-19-underscores-the-benefits-of-south-koreas-artificial-intelligence-push/.

8 https://www.newyorker.com/news/news-desk/seouls-radical-experiment-in-digital-contact-tracing.

9 https://www.fhwa.dot.gov/infrastructure/origin01.cfm.

10 https://tripnet.org/wp-content/uploads/2021/06/TRIP_Interstate_Report_June_2021.pdf.

11
https://www.google.com/books/edition/The_Best_Investment_a_Nation_Ever_Made/yrPmj_ieBq0C?hl=en.
12 https://www.nber.org/system/files/working_papers/w27938/w27938.pdf.
13 https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2019/wp19-29.pdf.
14
https://warwick.ac.uk/fac/arts/english/currentstudents/undergraduate/modules/ontheroadtocollapse/syllabus2018_19/mohl_stop_freeway.pdf.

CHAPTER 10

1 Brian Merchant, Blood in the Machine: The Origins of the Rebellion Against Big Tech (New York: Little, Brown, 2023), Kindle ed., 48, 88; e-book loc. 455, 1024.
2 Merchant, 56; loc. 653.
3 Merchant, 121; loc. 650.
4 Merchant, 297–99; loc. 3482–3512.
5 Merchant, 289; loc. 3405.
6 Merchant, 32; loc. 352.
7 https://scsp222.substack.com/p/the-ai-enabled-future-of-us-national.
8 https://www.youtube.com/watch?v=Y1pHXV7E4xY&t=377s.
9 https://www.smartnation.gov.sg/nais/?trk=article-ssr-frontend-pulse_little-text-block.
10 https://www.lemonde.fr/en/economy/article/2023/06/15/emmanuel-macron-wants-to-create-french-ai-models-to-compete-with-openai-and-google_6032118_19.html#.

11 Yuval Noah Harari, Sapiens: A Brief History of Humankind (New York: HarperCollins, 2015), 138, 24, 25. Harari reflects Benedict Anderson's influential understanding of modern nations as fundamentally "imagined communities," willed into existence over time by the set of narratives and norms that a group comes to accept as the defining lineaments of their territory, history, and shared identity. Anderson, Imagined Communities: Reflections on the Origin and Spread of Nationalism (London: Verso, 1983). For a more recent view that pays special attention to the way constructs of national identity often exclude some citizens and stories, while amply including falsehoods and distortions, see Kwame Anthony Appiah, The Lies That Bind: Rethinking Identity (New York: Liveright, 2018).

12 David Burnham, The Rise of the Computer State: The Threat to Our Freedoms, Our Ethics, and Our Democratic Process (New York: Vintage, 1983); e-book, Open Road Distribution (2014), 48.

13 https://harvardharrispoll.com/wp-content/uploads/2023/05/HHP_May2023_KeyResults.pdf.

14 https://en.wikipedia.org/wiki/Telescreen.

15 George Orwell, 1984 (New York: Berkley, 2023), 3.

16 https://www.forbes.com/2009/08/10/government-internet-software-technology-breakthroughs-oreilly.html.

17 https://blog.pol.is/pol-is-in-taiwan-da7570d372b5.

18 https://blog.pol.is/uber-responds-to-vtaiwans-coherent-blended-volition-3e9b75102b9b