شما در برابر امنیت اطلاعات خود مسئول نیستید!

در عصر هوش مصنوعی، دانستن اینکه اطلاعات شخصی شما کجا مورد استفاده قرار گرفته‌است غیر ممکن است.

علی رادمرد
دانشجو مهندسی کامپیوتر
دانشکدگان فارابی دانشگاه تهران
aliradmard5@gmail.com

در سال ۲۰۱۰، مارک زاکربرگ در مراسم اهدای جوایز TechCrunch به حضار گفت که جوانان، به ویژه کاربران شبکه‌های اجتماعی، دیگر به حریم شخصی اهمیت نمی‌دهند. او گفت: «مردم واقعاً راحت‌تر شده‌اند، نه تنها اطلاعات بیشتری و انواع مختلفی از آن‌ها را به اشتراک می‌گذارند، بلکه با صراحت بیشتر و با افراد بیشتری این کار را می‌کنند. این قاعده اجتماعی فقط چیزی است که با گذر زمان تکامل یافته‌است.» این بیان نمایانگر یک باور متداول است که نقض حریم شخصی وقتی رخ می‌دهد که افراد اطلاعات شخصی خود را آشکار می‌کنند. به عبارت دیگر، وقتی چیزی در ردیت یا تیک‌تاک وایرال می‌شود یا عکس ارسال شده به یک طرفدار فاش می‌شود، در ابتدا، اشتباه از طرف کسی است که آن را منتشر کرده‌است. این مدل از مسئولیت فردی بسیار مقاوم و ماندگار است، اما کاملاً اشتباه است و در عصر هوش مصنوعی بی فایده است.

هوش مصنوعی مفهوم مسئولیت فردی در حفظ حریم شخصی را کاملا نابود می‌کند، زیرا نمی‌توانید دسترسی این الگوریتم‌ها به اطلاعات شما را کنترل کنید یا ببینید که آن‌ها چه کارهایی با آن انجام می‌دهند. ابزارهایی مانند چت‌جی‌پی‌تی، دال-ای و گوگل بارد بر اساس داده‌های جمع‌آوری شده بدون موافقت یا حتی بدون اطلاع از آن‌ها، آموزش داده شده‌اند. در بدترین حالت، ست‌های آموزشی، مقدار زیادی از اطلاعات دیجیتال را جذب و ترکیب کرده و به عنوان مواد اولیه برای هوش مصنوعی مورد استفاده قرار می‌گیرند. در حالی که شرکت‌های فناوری در حال تلاش برای گنجاندن هوش مصنوعی در هر محصول قابل تصوری هستند، از موتورهای جستجو گرفته تا بازی‌ها و ابزارهای نظامی، امکان ندارد بدانیم این خروجی به کجا می‌رود یا چگونه ممکن است تفسیر شود. فروشنده های داده (Data brokers) داده‌های زیادی را از وب استخراج کرده‌اند و پرونده‌های بزرگی برای مردم تشکیل داده‌اند اما خروجی این داده‌ها برای یک فرد عادی، یا موتورهای جستجو به طور رایگان قابل دسترسی نیست. دسترسی گسترده به هوش مصنوعی، احتمال نقض حریم شخصی را افزایش می‌دهد و افراد بیشتری در خطر مواجهه با عواقب مضر این عمل قرار می‌دهد.

مجموعه‌های زیادی که توسط هوش مصنوعی استفاده می‌شوند، بدون شک حاوی اطلاعاتی در مورد افراد هستند که توسط آنها ارائه نشده، ساخته نشده، یا حتی تصوری از وجود آنها نبوده‌است. سوابق عمومی مرتبط با ازدواج‌ها، رهن مسکن و مشخصات رای‌دهندگان همگی در این مجموعه‌ها قرار دارند، همچنین خبرها، اطلاعات کارمندان و صفحات ویکی‌پدیا. این مجموعه همچنین شامل میلیون‌ها عکس و ویدیو می‌شود؛ به عنوان مثال، Dall-E بر روی تصاویری آموزش داده شد که از رسانه‌های اجتماعی، موتورهای جستجو و سایت‌ها جمع‌آوری شده‌اند. بنابراین، اگر در پس‌زمینه یک عکس از سال ۲۰۰۷ در فلیکر (flickr) باشید، تصویر شما ممکن است برای آموزش یک الگوریتم استفاده شود. به نظر نمی‌رسد که کسی بداند چه چیزی وارد این مجموعه داده می‌شود، و هیچ راهی برای نظارت یا کنترل آن وجود ندارد. وقتی چت‌جی‌پی‌تی یک زندگینامه نادرست در مورد من می‌نویسد، من نمی‌دانم اطلاعات غلط از کجا نشأت گرفته، همانطور که نمی‌دانم اطلاعات صحیح از کجا آمده است. ما عادت داریم که حریم شخصی را به عنوان کنترل فردی بر روی اطلاعات بدانیم، اما اگر حتی منشأ اطلاعات شخصی‌تان را هم ندانید، امکان نظارت بر نحوه استفاده از آن وجود ندارد.

آنتروپولوژیست‌ها و دانشجو های حقوقی سال‌هاست که می‌دانند حریم شخصی نمی‌تواند توسط افراد کنترل شود، بخشی به دلیل اینکه ما اطلاعات را در داخل شبکه‌ها به اشتراک می‌گذاریم. به عبارت دیگر، مردم درباره یکدیگر صحبت می‌کنند، هم در دنیای مجازی و هم در دنیای واقعی. راهی آسان برای تعیین محدودیت‌ها وجود ندارد؛ می‌توانید از دوستان خود بخواهید که عکس‌های فرزندان شما را در اینستاگرام نگذارند یا از شما در تیک‌تاک یاد نکنند، اما حریم شخصی شما، هر چقدرهم که سعی بکنید، نقض می‌شود. نقض‌های حریم شخصی شبکه‌ای اغلب به دلیل این اتفاق می‌افتد که اطلاعاتی که در یک محیط با قوانین و انتظارات خاص ارائه شده‌اند به جاهای دیگر منتقل شده و به شکل متفاوتی تفسیر می‌شوند. تیک‌تاک‌های ساخته شده برای مخاطبان کوییر(queer) به منبعی برای کمپین‌های ضد تغییر جنسیت تبدیل می‌شوند؛ سخنرانی‌های سیاسی ارائه شده برای مخاطبان همدرد زمانی که توسط مخالفان تماشا شوند، به محتوایی کاملا بی منطق تبدیل می‌شود.

فناوری‌های جدید به طور فزاینده‌ای حریم شخصی شبکه‌ای را تهدید می‌کنند. به عنوان مثال، شجره نامه قانونی به پلیس این امکان را می‌دهد که با بررسی شواهد ژنتیکی جمع‌آوری شده از خویشاوندان دور اشخاص مظنون را شناسایی کند. شما می‌توانید انتخاب کنید که از Ancestry.com استفاده نکنید، اما نمی‌توانید از یک فامیل دور که احتمالاً حتی اسمش را هم نشنیدید جلوگیری کنید که همین کار را انجام دهد. داده‌های بزرگ، که از مجموعه‌های داده‌ای عظیم به همین روش‌ها استفاده می‌کند، به طور متداول دوستان، خویشاوندان و حتی آشنایان دور را نیز به هم مرتبط می‌کند، که وقتی وارد الگوریتم‌های پیش‌بینی پلیسی یا الگوریتم‌های ارزیابی ریسک می‌شود، مشکل‌آفرین می‌شود. افراد نمی‌توانند اقدامی برای جلوگیری از چنین تجاوزاتی به حریم شخصی بکنند.

هوش مصنوعی مولد نگرانی‌های حریم شخصی شبکه‌ای را تشدید می‌کند؛ از انجام هر گونه «کار شخصی»، روش‌ها و استراتژی‌هایی که برای حفظ حریم شخصی استفاده می‌کنیم، باز می‌دارد. و خروجی‌های هوش مصنوعی تولیدی به طرقی کاملاً از منبع اصلی خود جدا می‌شوند که تا قبل از این تصور نمی‌شد. خود همین که پیام‌های شخصی به راحتی فاش می‌شوند مشکل ساز است، چه برسد به اینکه از کل محتوای ردیت (reddit) به عنوان یک داده برای شعرهای رباتی و مقاله‌های غیر حرفه‌ای دانشگاه استفاده شود.

اطلاعاتی که در یک موضوع ارائه می‌شود، می‌تواند کاملاً بازسازی و بازمخلوط شود، معنی آن را تغییر داده و با نقض آنچه که فیلسوف هلن نیسنبوم آن را «یکپارچگی موضوعی» می‌نامد، مواجه شود. چگونه کسی می تواند از این جلوگیری کند؟

علاوه بر این، هوش مصنوعی مولد می‌تواند انواع مختلفی از نقض‌های حریم شخصی خلاقانه را ممکن سازد. من نتوانستم ChatGPT را متقاعد کنم که آدرس خانه دوستم را به من بدهد، اما با خوشحالی توضیح داد چگونه می‌توان آدرس خانه‌ای را آنلاین پیدا کرد. دیپ فیک ها مسئله دیگری است؛ چه چیزی می‌تواند از تقلید سبک، صورت یا حتی سخنان شخص دیگر حریم خصوصی را بیشتر نقض کند؟ من تعدادی ضبط قدیمی از صدای خودم را به ابزاری به نام Descript آپلود کردم و چند ساعت بعد یک نسخه سنتز شده از صدای خودم داشتم که می‌توانستم از آن برای گفتن هر چیزی استفاده کنم (free trial). در حالی که برخی از ابزارهای معروف هوش مصنوعی مولد حاوی محافظ حریم شخصی هستند تا نقض‌های حریم شخصی خیلی جدی‌تری را جلوگیری کنند، برخی دیگر ندارند. هوش مصنوعی‌ها به داده‌ها دسترسی دارند و اگر با ناجوانمردانگی و بی وجدانی استفاده شوند، می‌توانند شبکه اجتماعی کامل یک شخص را نقشه‌برداری کنند، وضعیت مالی یا مشکلات سلامتی او را تشکیل دهند و شناسایی کنند و آیا در مقابل کلاهبرداری های اینترنتی میتوانند از خود محافظت کنند یا خیر.

با اینکه از دوره ی هیجانی بودن هوش مصنوعی آگاه هستم، اما تفاوتی بین هوش مصنوعی مولد و دیگر فناوری‌هایی با پیامد عمیق بر حریم شخصی وجود دارد. پلتفرم‌های اجتماعی توسط الگوریتم‌های مشخص کنترل می‌شوند و ما نحوه عملکرد آنها را می‌فهمیم ولی هیچ کس حتی افرادی که در حوزه ساخت و توسعه هوش مصنوعی مولد تحقیق می‌کنند و برنامه می‌نویسند، نحوه عملکرد دقیق مدل‌های زبان بزرگ(LLM) را نمی‌فهمد. آنها با نرخ تغییر و نوآوری بسیار متفاوتی تغییر می‌کنند و می‌توانند توسط افرادی با استانداردهای اخلاقی متفاوت از خودمان مورد استفاده قرار گیرند. هوش مصنوعی مولد نقاط ضعف مدل قدیمی حفظ حریم شخصی ما را نشان می‌دهد؛ وقت این است که ما بی‌فایده بودن آن را در حفظ حریم شخصی تشخیص دهیم و به سوی یک مدل جدید حرکت کنیم