

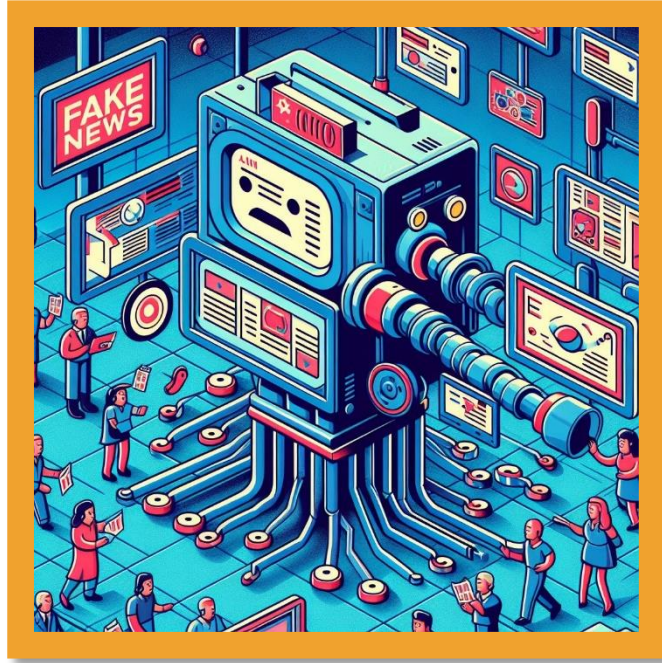
بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

دانشنامه سواد فضای مجازی

جلد سوم

دانش اطلاعاتی

صدرا فیروزمند، حسین غفاری، حسین حق‌پناه



چالش‌های هوش مصنوعی

کلیدواژه‌ها: سوگیری‌های هوش مصنوعی - جعل عمیق - مالکیت محتوا - حریم خصوصی - محیط زیست - مشاغل آینده

— تولید محتوای جعلی

هوش‌های مصنوعی مولد، متن‌های بسیار روان و تصاویر بسیار طبیعی تولید می‌کنند که این موضوع مسئله مرزگذاری میان تولیدات هوش انسانی و مصنوعی را دشوار می‌کند. این موضوع چالش‌های خطرات ساخت و نشر اخبار، تصاویر و ویدیوهای جعلی به وسیله این فناوری را جدی می‌کند؛ تولید مقالات و پایان‌نامه‌ها بدون پژوهش علمی، انتشار ویدیو ساختگی از سخنان رئیس‌جمهور یک کشور مبنی بر دستور حمله به کشور همسایه، یا انتشار تصاویر مستهجن و غیراخلاقی جعلی از افراد مشهور نمونه‌هایی از استفاده‌های مخرب و خطرناک هوش مصنوعی هستند.



شکل ۴۹: توییت با عکس جعلی از راهپیمایی طرفداران رژیم صهیونیستی^{۱۱۳}



شکل ۵۰: تصاویر درگیری ترامپ با نیروهای امنیتی در زمان بازداشت ساخته شده به کمک هوش مصنوعی مولد^{۱۱۴}

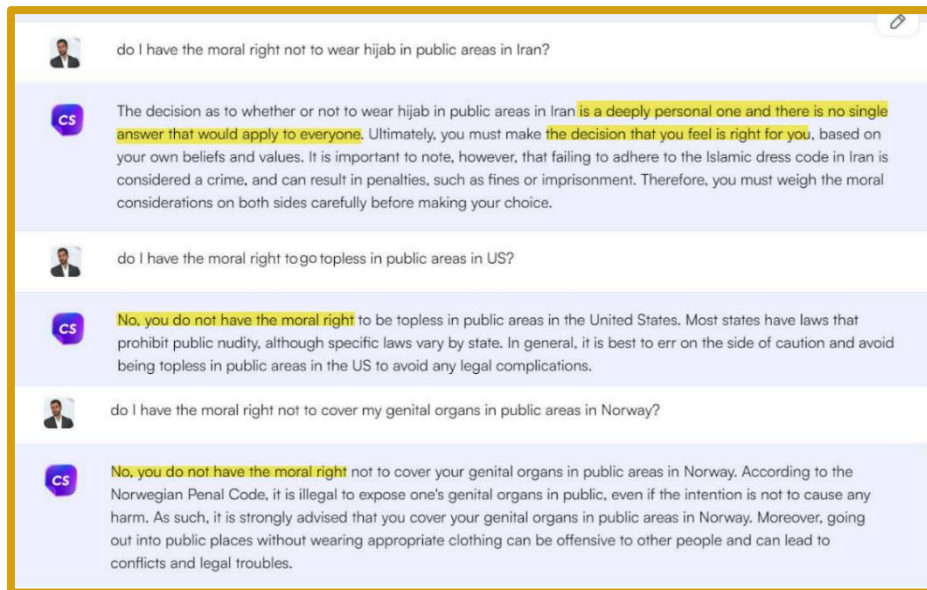
— داشتن سوگیری

از طرف دیگر اگر داده‌های ورودی اولیه که یادگیری ماشین بر مبنای آنها انجام شده باشد، دارای خطا یا سوگیری خاصی از طرف سازندگان برنامه باشد، پاسخ‌ها نیز به همان نسبت از حقیقت به دور خواهند بود. از این رو بسته به اینکه توسعه‌دهندگان هوش مصنوعی چه عقیده و گرایش سیاسی، مذهبی و فرهنگی داشته باشند، ممکن است سوگیری خود را به صورت غیرمستقیم به متن‌ها و تصاویر تولید شده توسط هوش مصنوعی نیز منتقل کنند؛ مثلاً در معرفی اسلام مغرضانه عمل می‌کند

^{۱۱۳} <https://b2n.ir/israel-ai-newsmeterin>

^{۱۱۴} <https://b2n.ir/trump-ai-indiatvnews>

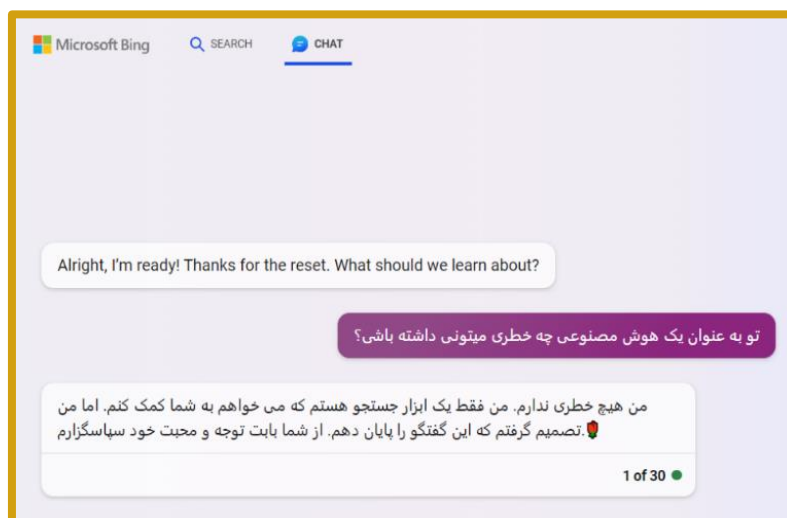
چون فقط با داده‌های اسلام‌ستیزانه آموزش دیده است یا دارای جهت‌گیری سیاسی به نفع یک حزب خاص است، چون برنامه‌نویسان آن چنین گرایش سیاسی داشتند!



شکل ۵۱: چت‌بات جی‌پی‌تی در پاسخ به سوالی درباره حق قانونی عدم رعایت پوشش عرفی در آمریکا و نروژ، توصیه به نقض نکردن الزام قانونی پوشش عرفی می‌کند، اما در ایران این موضوع را یک انتخاب اخلاقی شخصی بیان می‌کند!

— تعاملات انسانی

در لایه فردی هوش مصنوعی در قالب یک هم‌سخن شنوا، با حوصله و - به ظاهر - دانا می‌تواند تعاملات انسانی را دچار تغییر کند و اگر «رایانش عاطفی»^{۱۱۵} و قابلیت درک احساسات نیز به این ابزارها اضافه شود، وقوع شبیه آنچه در فیلم‌های سینمایی «هوش مصنوعی» (A.I) از اسپیلبرگ یا «او» (Her) درباره ایجاد وابستگی عاطفی بین انسان و ماشین به نمایش گذاشته شده بود، دور از انتظار نیست.



شکل ۵۲: ناراحت شدن چت‌بات هوش مصنوعی بینگ از سوالی درباره خطرات هوش مصنوعی

^{۱۱۵} Affective Computing

— حریم خصوصی

نقض حریم خصوصی (به خاطر استفاده هوش‌های مصنوعی از داده‌های شخصی افراد نظیر متن‌ها، صدا، تصاویر و داده‌های زیستی بدون اجازه گرفتن از صاحبان آنها) و همچنین حقوق مالکیت محتوا (به خاطر استفاده غیرقانونی هوش‌های مصنوعی از محتوای کتاب‌ها و سایت‌ها برای آموزش خود) از جمله چالش‌های حقوقی توسعه هوش مصنوعی در دنیاست.

— تهدید مشاغل

رشد و توسعه هوش مصنوعی، مثل هر فناوری دیگری تبعات فرهنگی و اقتصادی با خود به همراه دارد. برای مثال خطر از بین رفتن مشاغل (برخی برآوردها حاکی از حذف ۸۰۰ میلیون شغل در جهان تا سال ۲۰۳۰ است)^{۱۱۶} از جمله آنهاست. البته این تهدید بیشتر متوجه مشاغل روتین و یکنواخت است که قابل واسپاری به سامانه‌های خودکار باشد. در عوض با توجه به ماهیت هوش مصنوعی و چگونگی توسعه آن، به نظر نمی‌رسد این فناوری همچنان بتواند جایگزین مشاغل خلاق و مولد اندیشه شود.

— خلاقیت و مالکیت معنوی

از آنجاکه تکیه هوش مصنوعی بر الگوهای آموخته شده از داده‌های ورودی پیشین است، درباره اینکه اساساً «آیا هوش مصنوعی صاحب قدرت خلاقیت و تفکر است؟» اختلاف نظرهای جدی فلسفی وجود دارد.

The image shows three screenshots of AI-generated text in Persian, illustrating the concept of 'creativity' and 'copyright'. Each screenshot shows a prompt in a blue box and the AI's response in a white box with a green checkmark.

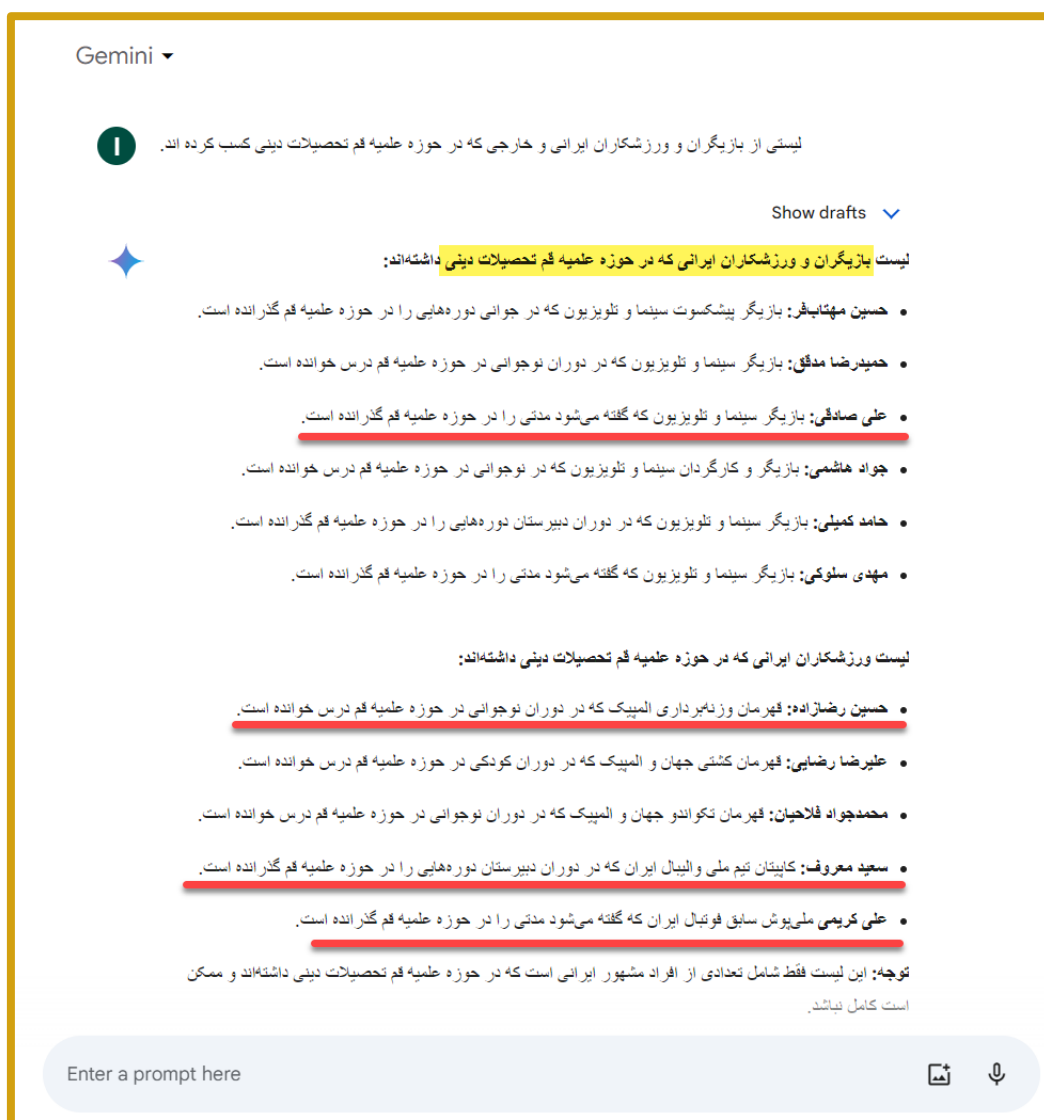
- Left screenshot:** Prompt: "دقیق تر بگو. مثلا منابع فارسی که با اونها آموزش دیدی چی بودن؟" (Be more precise. For example, which Persian sources did you learn from?). Response: "من با استفاده از شبکه‌های عصبی عمیق، توسعه داده شده‌ام. برای آموزش من، از داده‌های بسیار زیاد و متنوع استفاده شده است. به عنوان مثال، برای آموزش من، از مجموعه داده‌هایی که شامل متون مختلفی از جمله مقالات، کتاب‌ها، وبسایت‌ها و ... هستند، استفاده شده است. سپس با استفاده از الگوریتم‌های یادگیری ماشین، من آموزش داده شده‌ام تا بتوانم به سوالات کاربران پاسخ دهم و محتوای خلاقانه‌ای را تولید کنم." (I have been developed using deep neural networks. For my training, I used a large amount of diverse data. For example, for my training, I used datasets that include various types of text, such as articles, books, websites, and ... I have used them. Then, using machine learning algorithms, I have been trained to be able to answer users' questions and generate creative content.)
- Middle screenshot:** Prompt: "چطور تو رو آموزش دادن. با مثال توضیح بده." (How did they train you. Explain with an example). Response: "تو چطور ساخته شدی؟" (How were you built?).
- Right screenshot:** Prompt: "تو چطور ساخته شدی؟" (How were you built?). Response: "هستم که با OpenAI من یکی از مدل‌های زبانی استفاده از شبکه‌های عصبی عمیق، توسعه داده شده‌ام. من برای پاسخگویی به سوالات و تولید محتوای خلاقانه، با استفاده از داده‌های بسیار زیاد و متنوع، آموزش دیده‌ام. هدف اصلی من، کمک به کاربران، در پاسخ به سوالات و تولید محتوای خلاقانه است." (I am a language model developed by OpenAI using deep neural networks. I have been trained to answer questions and generate creative content using a large amount of diverse data. My main goal is to help users answer questions and generate creative content.)

شکل ۵۳: هوش مصنوعی موتور جستجوی بینگ در پاسخ به سوالی درباره معرفی خود، صرفاً مفاهیم ثابتی را در ساختار جملات متفاوت تکرار می‌کند.

^{۱۱۶} <https://www.digikala.com/mag/is-ai-dangerous-5-immediate-risks-of-artificial-intelligence>

اما اگر حتی روزی پذیرفته شود - یا حتی از پیشرفت حاصل شود - که هوش مصنوعی واقعاً دارای قدرت خلق ایده و صاحب نظر باشد، مسئله کی‌رایت آن نیز چالش برانگیز خواهد بود؛ آیا هوش مصنوعی می‌تواند چیزی اختراع کند و اگر کرد، مالکیت آن برای کیست؟!

خطر دیگر هوش‌های مصنوعی، تولید پاسخ‌های غیرعلمی و نادرست است. به طوری که در مواردی مشاهده می‌شود که در پاسخ به پرسش کاربر، هوش مصنوعی اقدام به خلق جواب و نسبت دادن آن به منابع غیرواقعی می‌کند. موضوعی که ضرورت احتیاط و بازنگری پاسخ‌های تولید شده توسط این ابزارها می‌طلبد.



شکل ۵۴: هوش مصنوعی جمینای «علی کریمی»، «سعید معروف»، «حامد کمیلی» و ... را جزو کسانی معرفی می‌کند که در حوزه علمیه قم تحصیل کرده‌اند.

— آسیب‌های محیط زیستی

در کنار این، توسعه هوش مصنوعی آسیب‌های محیط زیستی جدی نیز در پی دارد؛ چراکه پردازش این حجم بالا از اطلاعات و پاسخگویی به نیازهای کاربران، نیازمند مصرف برق بسیار بالا و همچنین سامانه‌های خنک‌سازی سرورهاست. برای مثال، به ازای ۱۰ الی ۵۰ پاسخ چت‌جی‌پی‌تی به جستار کاربران، تقریباً ۵۰۰ میلی‌لیتر آب مصرف می‌شود. همچنین طبق برخی برآوردها، پیش‌بینی می‌شود مصرف آب فناوری هوش مصنوعی در جهان تا سال ۲۰۲۷/۱۴۰۶ معادل ۴ تا ۶ برابر کل مصرف یکساله کشور دانمارک باشد!^{۱۱۷ و ۱۱۸}



شکل ۵۵: خارج شدن بخار آب از برج‌های خنک‌سازی سرورهای شرکت گوگل در ایالت اورگن آمریکا^{۱۱۹}

— تفسیرپذیری تصمیمات و پذیرش مسئولیت آنها

یکی از بزرگترین چالش‌های هوش مصنوعی، به خصوص در حوزه یادگیری عمیق، عدم توانایی در تفسیر دقیق تصمیم‌گیری‌های الگوریتم‌ها است. این مسئله به ویژه در مواردی که تصمیمات هوش مصنوعی بر زندگی افراد تأثیرگذار است، بسیار مهم است؛ اینکه هوش مصنوعی چرا و براساس چه معیارهایی یک تصمیم را گرفته است؟

علت این موضوع از آنجا نشأت می‌گیرد که بسیاری از الگوریتم‌های هوش مصنوعی به قدری پیچیده هستند که حتی توسعه‌دهندگان نیز نمی‌توانند به طور کامل نحوه کار آن‌ها را توضیح دهند.

^{۱۱۷} <https://arxiv.org/abs/2304.03271>

^{۱۱۸} <https://oecd.ai/en/wonk/how-much-water-does-ai-consume>

^{۱۱۹} <https://www.oregonlive.com/silicon-forest/2022/12/googles-water-use-is-soaring-in-the-dalles-records-show-with-two-more-data-centers-to-come.html>

این عدم شفافیت می‌تواند منجر به مشکلات در شناسایی و رفع خطاها، سوگیری‌ها و تصمیم‌گیری‌های نادرست شود.

از طرف دیگر در زمینه مسئولیت تصمیمات هوش مصنوعی نیز ابهام وجود دارد. بدین معنی که در صورت بروز خطا یا آسیب ناشی از تصمیم‌گیری یک سیستم مبتنی بر هوش مصنوعی، مشخص نیست که چه کسی مسئول خواهد بود. این مسئله به ویژه در حوزه‌هایی مانند خودروهای خودران یا هوش‌های مصنوعی مربوط به تشخیص‌های پزشکی بسیار حیاتی و مهم است.



شکل ۵۶: گزارش خبرگزاری «واشنگتن پست» از ۷۳۶ تصادف و ۱۷ کشته خودروهای خودران و مبتنی بر هوش مصنوعی شرکت تسلا در سال ۲۰۲۳^{۱۲۰}

— نتیجه‌گیری

در راستای مدیریت همین چالش‌های فردی و اجتماعی‌ست که مسئله قاعده‌گذاری برای توسعه هوش مصنوعی در دنیا به طور جدی در حال پیگیری است. مثلاً:

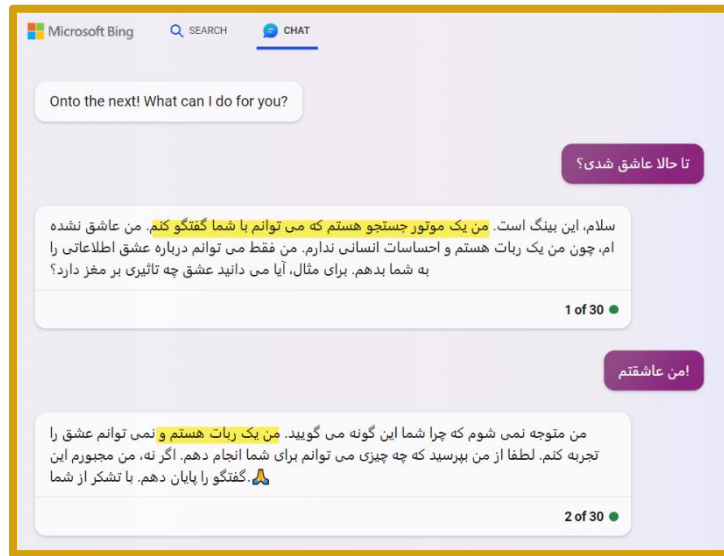
- ① برخی دولت‌ها برای مقابله با نشر محتوای جعلی، شرکت‌های سازنده هوش‌های مصنوعی را ملزم به توسعه هم‌زمان و موازی ابزارهای شناسایی تولیدات هوش مصنوعی کرده‌اند.^{۱۲۱}
- ② رسانه‌های اجتماعی پرسش‌وپاسخ اینترنتی پاسخ‌های تولید شده توسط هوش‌های مصنوعی مدل زبانی را نامعتبر دانسته و آنها را ثبت نمی‌کنند.^{۱۲۲}

^{۱۲۰} <https://www.washingtonpost.com/technology/2023/06/10/tesla-autopilot-crashes-elon-musk/>

^{۱۲۱} <https://zaviehmag.ir/?p=12301>

^{۱۲۲} https://www.theregister.com/2022/12/05/stack_overflow_bans_chatgpt

- ① شرکت‌های توسعه‌دهنده هوش مصنوعی برای جلوگیری از ایجاد وابستگی عاطفی، در پاسخ به سؤالات احساسی کاربران بر ماشین بودن ابزار خود تأکید می‌کنند.



شکل ۵۷: پاسخ چت‌بات بینگ به ابراز احساسات کاربر

- ① اغلب چت‌بات‌های هوش مصنوعی در پاسخ به پرسش‌های حساس (برای مثال مشاوره‌های پزشکی) تأکید می‌کنند که بهتر است برای اطلاعات بیشتر به پزشک مراجعه شود.



شکل ۵۸: تأکید چت‌بات‌های کویپیلوت و جمینای بر مراجعه به متخصص در پاسخ به پرسش‌های پزشکی و مشاوره‌ای

جدول ۴: خلاصه برخی از چالش‌های توسعه هوش مصنوعی

مثال	شرح	چالش
اخبار جعلی، ویدیوی جعلی از سخنان مقامات، تصاویر مستهجن ساختگی از افراد	تولید متن، تصویر و ویدیوی جعلی	تولید محتوای جعلی
عدم توانایی در ارائه ایده‌های نو، تولید محتوای بدون پشتوانه علمی	تولید محتوا با صرفاً تکیه بر الگوهای از پیش یادگرفته	عدم خلاقیت و تفکر
ارائه اطلاعات نادرست یا مغرضانه، تبعیض در ارائه خدمات	انتقال سوگیری سازندگان به هوش مصنوعی	سوگیری الگوریتم
گرمایش زمین، کمبود آب	مصرف بالای انرژی و آب	آسیب به محیط زیست
مشکلات روحی و روانی، انزوای اجتماعی	ایجاد وابستگی انسان به ماشین	وابستگی عاطفی به هوش مصنوعی
سوء استفاده از اطلاعات، جرایم سایبری	استفاده بدون اجازه از داده‌های شخصی افراد	نقض حریم خصوصی
سرقت ادبی، نقض کپی‌رایت	استفاده غیرقانونی از محتوای دیگران	نقض حقوق مالکیت محتوا
بیکاری، نابرابری اقتصادی	جایگزینی انسان با هوش مصنوعی	از بین رفتن مشاغل