هوش مصنوعی (AI) به طور فزاینده ای روی زندگی افراد در سراسر جهان تأثیر می گذارد، از فناوری های نسبتاً دست و پاگیر تا سیستم های دیجیتال پیشرفته تر که قادر به تصمیم گیری خودشان هستند. هوش مصنوعی در صورت طراحی و استفاده مناسب، پتانسیل زیادی دارد، اما در غیر این صورت، خطرات جدی نیز به دنبال خواهد داشت. کتاب "هوش مصنوعی برای تنوع" درباره این موضوع سوالاتی مطرح می کند که چگونه تکنولوژی هوش مصنوعی می تواند منجر به بی توجهی و تبعیض علیه زنان(میزوژینیستیک): بی توجهی و تبعیض علیه افراد همجنس(هموفوبیک) تبعیض علیه افراد بر اساس نژاد و / یا اقوام مختلف. (نژادپرستی) تبعیض بر اساس سن تبعیض علیه افراد بر اساس آنچه برای آنها معروف است یا نیست. بی توجهی و تبعیض علیه افرادی که جنسیت خود را تغییر می دهند یا جنسیتشان با آنچه در حالت نرمال انتظار دارید، مطابقت ندارد ( ترنسفوبیک ). بی توجهی و تبعیض علیه افراد معلول و منجر به جامعه شناختی بودن یا حاشیه نشینی گروه های متنوع در جامعه شود. به همین دلیل، نیازمند به دیدن عواقب غیر منتظره در صورتیکه سیستم ها بدون کنترل باقی بمانند هستیم. این کتاب به خواننده کمک می کند تا بیشتر با تأثیرات هوش مصنوعی بر گروه های مختلف افراد آشنا شود و چگونگی استفاده از رویکردهای آگاهی از تنوع در هوش مصنوعی را یاد بگیرد.

راجر آ. سراآ، عضو هیئت علمی دانشکده مطالعات بین رشته ای فرهنگ دانشگاه علوم و فناوری نروژ (NTNU) در تروندهایم، نروژ است. علایق تحقیقاتی او شامل دیجیتالی شدن و رباتیک شدن جامعه و پیامدهای ایمنی، جنسیتی و فلسفی آن است. وی گروه تحقیقاتی "دیجیتالی شدن و رباتیک شدن جامعه" را رهبری می کند و چندین پروژه تحقیقاتی را هماهنگ می کند، از جمله پروژه "BIAS: شکاف های تنوع در بازار کار" در چارچوب برنامه هورایزن اروپا.

**هوش مصنوعی برای همه چیز**

هوش مصنوعی (AI) در سراسر جهان وجود دارد. از خودروهای بدون راننده تا کامپیوترهای برنده بازی و محافظت در برابر تقلب، AI در بسیاری از جوانب زندگی نقش دارد و تأثیر آن در آینده نیز ادامه خواهد داشت. بسیاری از شرکت های پرارزش جهان به قوی سازی تحقیقات و توسعه AI می پردازند و روزی بدون خبر از شکست های نوآورانه در AI و رباتیک گذشته نمی شود. سری AI برای همه چیز نقش هوش مصنوعی را در زندگی معاصر، از ماشین ها و هواپیماها تا پزشکی، آموزش، مد و غیره، بررسی می کند. هر کتاب کوتاه و قابل دسترس توسط متخصصان حوزه نوشته شده است و باعث می شود تا مطالعه و واقعیت هوش مصنوعی به هر دسته ای از خوانندگان؛ از جمله حرفه ای های مهتم، دانشجویان، پژوهشگران و خواننده های عادی رسیدگی شود.

فهرست مطالب

|  |  |
| --- | --- |
| مقدمه  قدردانی  لیست کوتاه شده عبارات | [vii](#_مقدمه) ix [xi](#_bookmark0) |
| 1 باز کردن جعبه سیاه هوش مصنوعی | 1 |
| 2 هوش مصنوعی جنسیتی: عملکرد ، انتظارات و جنسیت گرایی | 21 |
| 3 هوش مصنوعی کوییر: بیان جنسیت ، هویت و باینری ها | 39 |
| 4 هوش مصنوعی و نژاد: شناخت ، تعصب و مسائل سیستمیک | 49 |
| 5 بدن ها و هوش مصنوعی: سلامت ، پیری و معلولیت | 61 |
| 6 هوش مصنوعی و کلاس: کار ، آموزش و پایداری | 83 |
| 7 تلاقی زوجین اقتدار و هوش مصنوعی مسئول | 99 |
| فهرست نمایه | [117](#_bookmark2) |

### **مقدمه**

تکنولوژی دیجیتال بر زندگی‌های ما تأثیر قابل توجهی دارد. آنها می‌توانند با استفاده مناسب، به عنوان یک افزودنی عالی برای غنی کردن، ساده کردن و بهینه کردن عملکردها در زندگی روزمره مفید باشند، اما آنها نیز دارای پیامدهای پلید بیشتری هستند که در صورت استفاده بدون احترام به تنوع آنها به وجود می‌آیند. گشایش روند استفاده از فناوری دیجیتال که موجب بهبود زندگی انسان‌ها شود و از بروز موانع در راه کاربرد آنها جلوگیری شود، محور اصلی این کتاب است. ایده نوشتن این کتاب از زمانی شروع شد که من به عنوان مشاور و ارزیاب در گروه تخصصی Gendered Innovations 2 تحت عنوان مشاور و ارزیاب در کمیته اصلاحات جنسیتی در کمیته‌های اروپایی درباره توجه به جنسیت و تنوع در راستای بهبود فناوری و تحقیقات به شیوه‌هایی پویا و شامل‌تر کمک کردم. زمانی که افرادی از زمینه‌های همگون تنها با خودشان تعریف می کنند که چه برنامه‌ها، سیستم‌ها و داده‌هایی را در زندگی روزمره و جامعه‌های خود ما استفاده می کنیم، ریسک می‌کنیم این که جهانی کمتر شامل‌‌عناصر متنوع و به سمت اولویت‌ها و تجربیات آنها ساخته شود و بقیه را پشت سر بگذاریم. هنوز ما برای تنوع در زمینه هوش مصنوعی و زمینه‌های مرتبط، برای شامل شدن افراد متفاوت و خوب شدن برای همه، مسیر درازی را پیش رو داریم. این کتاب را برای شناسایی داده‌های نامنصفی و کاهش محدودیت و ابزاری برای جلوگیری از تبعیض و ساختن زمینه ای برای دانستن یک جهان بهتر برای همه، به عنوان یک مقدمه راهنمایی استفاده نمایید.

**این کتاب برای کیست؟**

اگر به دنبال یک کتاب فنی هستید که توضیح دهد چطور سیستم‌های مختلف هوش مصنوعی کار می‌کنند، این کتاب برای شما مناسب نیست. همچنین این کتاب انسایکلوپدیای تمام هوش مصنوعی‌های موجود نیست. در یک کتاب کوتاه، قطعاً جای بند اماکن و سیستم‌هایی وجود خواهد داشت که پوشش داده نشده است. همچنین، دانشمندان جنسیتی ممکن است به کتابی با بیشتر حجم و بیشتر عمیق برای بررسی هر موضوع در سطح نظری نیاز داشته باشند. همچنین، افرادی که به دنبال انتقام در برابر تمام هوش مصنوعی به عنوان چیزی بد هستند و فردی که به عنوان یک تکنوتوپیست خیال‌پرداز، همین کتاب را انتخاب نکنند. اما اگر می‌خواهید یک تحقیق بحرانی در مورد اینکه چگونه سیستم‌های مختلف هوش مصنوعی می‌توانند بر تنوع اجتماعی و فردی در عبارت فردی‌ شامل یا مستثنی از گروه های خاص تاثیر بگذارند و می‌خواهید بیشتر درباره آن مطلع شوید، آنگاه، با خیال راحت، این کتاب را بخوانید و به ساخت یک دنیای بهتر از هوش مصنوعی برای تنوع کمک کنید.

### **قدردانی**

من باید از گروه تحقیقاتی و شرکای تحقیقاتی خودم که در فرآیند نظردهی و نوشتن این کتاب شرکت داشتند، سپاسگزاری کنم. به خصوص، می‌خواهم از نقادان مفید واقعی که نظرات و پیشنهادات خود را ارائه کردند، تشکر کنم، و از آشلی شو، الیزابت استوبرود، اینگا استرومک، جنیفر برانت، جولی کاترین فلیک، کاترین اوانز، کریستین آسک، مارک و. خارس، مارتینه سلتن، ماشا کورپیکز-بریکی، سیری اسلبهو سرنسن، صوفیا موراتی، صوفیا ایوارسون و دو نفر از نقادان ناشناخته تشکر کنم. همچنین باید از نینک بروینینگ برای تصاویر و ویراستار خودم، الیوت مورسیا، تیم ویرایشی در CRC Press/Taylor & Francis Group و همچنین مدیر پروژه، یاسر عرفانات، قدردانی کنم. این کتاب ممکن شد با پروژه‌های زیر: Robotics4EU (شماره قرارداد 101017283، برنامه پژوهش و نوآوری اتحادیه اروپا Horizon 2020)، LIFEBOTS Exchange (شماره قرارداد 824047، برنامه پژوهش و نوآوری اتحادیه اروپا Horizon 2020) و AUTOWORK (شماره پروژه شورای تحقیقاتی نروژ 301088).

### **لیست کوتاه شده عبارات**

AI به معنای هوش مصنوعی است.

LGBTQ+ شامل لزبین، گی، دوجنسه، ترنسژندر، کوییر و افراد دیگری می‌شود.

ML به معنای یادگیری ماشینی است.

BIPoC به معنای افراد سیاه، اصلی‌نژاد و افراد با رنگ پوست متنوع است.

STS به مطالعات علوم و فناوری اشاره دارد.

1

**باز کردن جعبه سیاه هوش مصنوعی**

هوش مصنوعی (AI) در اکثر موارد، نامفهوم و پیچیده تلقی شده و بررسی برای مشخص شدن اینکه کیفیت استفاده از آن به چه کسانی مفید است یا به چه کسانی ضرر و زیان وارد می‌کند، لازم است. اغلب توسط فناوران آینده‌نگر به عنوان یک تحول آهنین جاودان و شفاف تلقی می‌شود که می‌تواند مشکلاتی همچون بحران درمان، تغییرات اقلیمی، امنیت سایبری، بهبود فرآیند یادگیری و حفاظت از تنوع زیستی را حل کند. اما به نظر دیگران، این تکنولوژی می‌تواند منجر به مشکلات مختلفی در جامعه شود که حتی می‌تواند به نابودی بشریت منجر شود، به عنوان مثال، نظارت و کنترل توپ دانشگاهی یا نظامی، سلاح‌های مرگ‌بار، و مسائل اخلاقی مربوط به فرزندان طراحی‌شده. ایده ثابت این است که این تکنولوژی زمینه ی ایجاد تغییرات بهتر یا بدتر را فراهم می‌کند. عده‌ای باور دارند که به طور تعیین‌شده تکنولوژی در میزان پذیرش به تغییرات در جامعه کمک می‌کند و بهتر است برای ایجاد تغییرات مورد نظر برای استفاده از سیستم‌های هوش مصنوعی، نیاز است در تبادل نظر و بحث جامعه شرکت کند. در این کتاب، من سعی می‌کنم بین دو نمایشگر "هوش مصنوعی خوب و بد"، به شکلی پراگماتیک، بررسی کنم که هوش مصنوعی چگونه برای فردی که از آن استفاده می‌کند، برای طراحی و توسعه و برای جامعه به‌صورت تخصصی اثر خواهد گذاشت. این بررسی به شناخت بهتر اینکه «چه کسانی از هوش مصنوعی بهره‌مند می‌شوند» و «چه کسانی به نفع از استفاده از آن بهره‌مند نمی‌شوند» کمک می‌کند. در حال حاضر، هوش مصنوعی به شدت زندگی ما را تحت تاثیر قرار داده است. تصمیمات ما درباره چگونگی کار کردن، کیفیت مراقبت‌های بهداشتی دریافتی، تصمیمات درباره بیمه و وام بانکی، کسانی که در رسانه‌های اجتماعی به عنوان هم‌راه انتخاب شده‌ایم و چگونگی مصرف کالاهای ما، همه تحت تاثیر این اختصار رازآمیز به نام "هوش مصنوعی" قرار می‌گیرند. اما هوش مصنوعی چیست و چگونه بر زندگی ما تأثیر می‌گذارد؟ چیست که باعث می شود در تصمیم گیری‌هایی مانند کدام فرد برای مصاحبه کاری دعوت می‌شود یا کدام فرد باید از یک داروی آزمایشی استفاده کند، عامل‌های بیش از اندازه ایجاب‌کننده فرد A نسبت به فرد B باشد؟ آیا تصمیمات مختلفی توسط هوش مصنوعی بر اساس ویژگی‌های شخصی مختلف گرفته می‌شود که ممکن است برای برخی افراد یا گروه‌های خاص موجب مزیت شود و برای دیگران تبعیض‌آمیز باشد؟

هوش مصنوعی معمولا به عنوان جعبه‌ سیاه (black box) مورد بررسی قرار می‌گیرد که در آن وقتی

کار علمی و فنی با موفقیت خود نامرئی می شود. وقتی یک ماشین به طور موثر کار می کند ، فقط باید روی ورودی ها و خروجی های آن تمرکز کرد نه روی پیچیدگی داخلی آن. بنابراین ، به طور متناقض ، هرچه علم و فناوری بیشتر موفق شوند ، مبهم تر و مبهم تر می شوند.

(Latour, 1999, p. 304)

عماری جعبه سیاه (black-box) در هوش مصنوعی به این معنی است که افراد غیر متخصص (و حتی برخی از متخصصین خود) از چگونگی و چرایی کارکرد هوش مصنوعی خبر کمی دارند. این کتاب به داخل "جعبه سیاه هوش مصنوعی" که در جامعه به کار می رود، نگاه می کند، و مورد توجه قرار می دهد چه افرادی به چه روش‌هایی تحت تأثیر هوش مصنوعی قرار می گیرند و کاربردهای هوش مصنوعی برای افزایش شامل بودن و کاهش تبعیض را نیز بررسی می کند. هوش مصنوعی نیز در صورتی که افرادی که آن را طراحی می کنند از اقدامات تبعیض آمیزی که آنها به صورت ضمنی در سیستم های خود راه اندازی کرده‌اند آگاه نباشند و یا داده هایی که در سیستم های هوش مصنوعی آموزش داده می شوند نماینده جامعه و یا شرایطی که ما می خواهیم در آن برنامه کاربردی را به کار گیریم را نشان ندهند - تبعیض آمیز خواهد بود. ویتکاور به عنوان مثال، بر این عقیده است که "برای کاهش اثرات تبعیض‌آمیز طراحی فنی، باید دیدگاهی را اتخاذ کنیم که به صریحی در برابر تبعیض ضد عمل کند" (ویتکاور، 2016، p. 1). تبعیض آمیزی در هوش مصنوعی می تواند هم در سطح گروهی - از طریق سیستم استخدام یک شرکت که قبلاً زنان را استخدام نمی کرد - و هم در سطح فردی - از طریق یک سیستم هدف گیری، فرد مورد نظر را برای یک امتیاز (اجتماعی) کمتر مبادرت قرار می دهد اگر فردی به پارامترهایی که با تقلب بیمه سلامت مرتبط هستند، جور دربیاید. آگاهی کافی نیست؛ ما به دانشی از تحولات اجتماعی و مادی هوش مصنوعی نیاز داریم، و به دانش، تست‌ها و آزمایش‌هایی که به مسائل مورد انتقاد فراتر می‌روند. هوش مصنوعی به طور فزاینده‌ای زندگی روزانه ما را تعیین می کند، که نیاز به ارزیابی مجدد روابط انسان و ماشین دارد (Dräger & Müller-Eiselt، 2020، p. 7).

اما با تنها برجسته کردن چند مثال از عملکرد های ناموفق یا تصمیمات بد AI - به عنوان مثال، مورد بات توییتری تی توی شرکت مایکروسافت است که پس از ساعت ها یادگیری از کاربران توییتر از طریق ارسال توییتهای نژادپرستانه مانند انکار الماس (Wolf et al.، 2017) - ما در یک اتاق بازتابی از ایستادن AI با "ماشین های بد" گیر کرده ایم. این مهم است که مسئولیت های اخلاقی توسعه دهندگان برای فناوری را مد نظر قرار دهیم. همچنین، درک اینکه چگونه AI می تواند در واقع برای بهره برداری از بشریت استفاده شود و چگونه می توان از AI به عنوان یک ابزار برای افزایش آگاهی از تنوع، بی عدالتی نژادی، مسائل جنسیتی، حقوق LGBTQ+ و برای دیگر گروه های تبعیض شده یا محروم، بهره برداری کرد. همزمان با توجه به وجود و کاهش خطرات آن.

یک مثال آن این است که چگونه AI می تواند مدیران و سرپرستان را از تبعیض در فرآیند استخدام و ترفیع آگاه کند. چندین مطالعه نشان داده است که نام، جنسیت یا ظاهر یک بازدید کننده شغلی ممکن است برای ارزیابی آنان و استنتاج مهارت های آنها قبل از اینکه کاندیدا شانسی برای ثابت کردن خودشان در عمل داشته باشند، استفاده شود. مثالی از تبعیض انسانی این است که اضافه کردن عکس سری به رزومه شخص در برخی کشورها غیرقانونی است، توصیه می شود یا به طور کاملاً تصادفی است. AI، به محاسبه تبعیض های قبلی شرکت ها، می تواند مبنای استخدام کارکنان بیشتر تنوع را فراهم کند (کوهن، 2019). مثال دیگری از AI است که خطرات بالینی برای گروه های خاصی را با قدرت محاسباتی بالا نشان می دهد، به عنوان مثال، با ایجاد پایگاه داده های بزرگی از خطرات بالینی از طریق تشخیص تصویر و تشخیص سریع تر و به طور موثر تر بیماران پتانسیل را تشخیص می دهد تا مراقبت های بهتری ارائه دهد (مک کینی و همکاران، 2020؛ پیسانو، 2020؛ سلیم و همکاران، 2020؛ وانگ و همکاران، 2021). درک AI و تبعیض انسانی، باعث کاهش خطرات مسئله می شود. نمونه هایی از این کار شامل استفاده از نمونه های خون بیشتر از افراد بالای 60 سال به منظور کاهش تبعیض AI در باره سلامتی سالمندان و افزودن پایگاه داده های بیشتر و شامل افراد رنگی به منظور بهبود تشخیص تصویری در بین غیرسفیدها است. با این حال، خطرات تنوع AI نیز وجود دارد.

در کتاب، من پنج ویژگی شخصی متمایز را انتخاب کرده ام که در آن به مسائل مختلف AI و تنوع می پردازم. من توصیف می کنم که چگونه AI با جنسیت (فصل 2)، جمعیت LGBTQ+ (فصل 3)، نژاد (فصل 4)، بدن، سلامتی و پیری (فصل 5) و وضعیت اجتماعی و طبقه (فصل 6) برخورد می کند. من این دسته ها را انتخاب کرده ام زیرا مثال های مهمی از اینکه چگونه AI می تواند بر اساس ویژگی های شخصی زندگی مردم را تحت تأثیر قرار دهد، ارائه می دهند. استفاده از AI در ارتباط با این موارد مسائل عمیق ساختاری جامعه را در دینامیک قدرت متقاطع زبان، قوانین و بازارهای کار نشان می دهد. درک اینکه AI چگونه بر یک دسته از مردم تأثیر می گذارد، می تواند به ما درباره تصمیمات تبعیض آمیزی که با دسته های دیگر ارتباط دارند و توسعه فن آوری آینده، بیاموزد.

هر فصل به بررسی تحقیقات و پروژه هایی می پردازد که به طور مستقیم با این پارامترها و AI درگیر هستند، تا نگاهی معاصر از نگاهی که جهان چگونه به تنوع و AI نگاه می کند، در سال ۲۰۲۳ نشان دهد. در پایان کتاب، با استفاده از نظریات مطالعات علم و فناوری (STS)، فمینیسم "تکنوساینس" و ارتباط بین جنسیتی، نژاد و ... از اسطوره انسان استاندارد چالش برانگیز می کنم، با دیدن اینکه چگونه AI می تواند تبعیض آمیز باشد در صورت توسعه بدون توجه به استانداردهای صحیح. برای جلوگیری از این امر، به AI مسئول نیاز داریم که مسایل تنوع را به جدیت بپذیرد. با دیدن مثال هایی از اینکه چگونه AI برای گروه های متنوع از جامعه به نفع و به ضرر کار می کند و در آگاهی از پیچیدگی ها، سیستم های AI مسئول، شامل و مقاوم تر می توانند برای رفع نیازهای جامعه متنوع و انصاف پذیر به کار گرفته شوند.

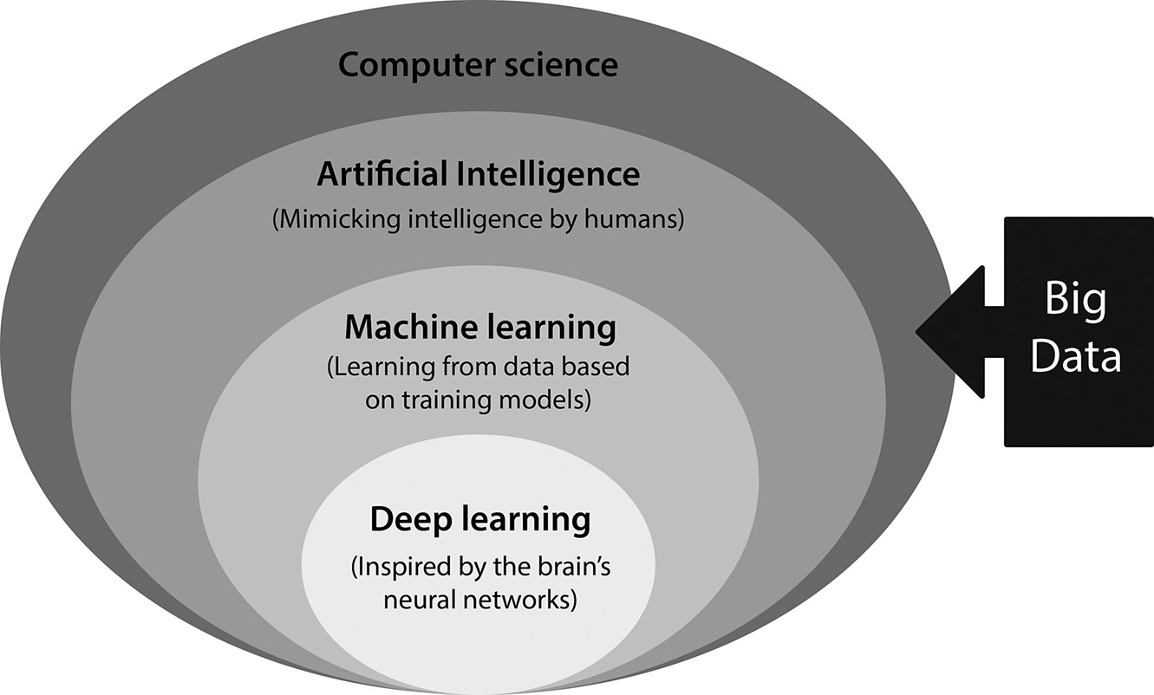
**AI چیست؟**

آگاهی یک فرد تصادفی از AI بیشتر از این است که به دنبال تخصصی در این زمینه باشد و به احتمال زیاد با جستجوی عبارت و دسترسی به فرهنگ لغت مواجه خواهد شد. در فرهنگ لغت ها، تعدادی تعریف متفاوت در مورد AI وجود دارد. به عنوان مثال، در فرهنگ لغت انگلیسی آکسفورد (2022) AI به عنوان "ظرفیت کامپیوترها یا دیگر ماشین ها برای اظهار یا شبیه سازی رفتار های هوشمندانه" تعریف شده است، در حالی که فرهنگ لغت Merriam-Webster (2022a) آن را به عنوان "یک شاخه از علوم کامپیوتر معرفی می کند که با شبیه سازی رفتار هوشمندانه در کامپیوترها و قابلیت تقلید از رفتارهای هوشمند انسان سر و کار دارد". تعاریف مشابهی نیز می توان در تحقیقات یافت، به عنوان مثال در Anjila (1984، ص 65): "هوش مصنوعی یک شاخه از علم و فناوری است که ماشین های هوشمند و برنامه های رایانه ای ایجاد می کند تا وظایف مختلفی را که نیاز به هوش انسانی دارند انجام دهند".

با توجه به این تعاریف، می توانیم تفاوتی را در شبیه سازی هوش مانند انسان و داشتن هوش مانند انسان مشاهده کنیم. همچنین، تفاوتی بین داشتن هوش کلی یا اینکه ماشین بتواند وظایف خاصی را با هوش فراگیر انجام دهد، وجود دارد. AI به طور معمول به دو دسته "AI ضعیف" (سیستم های AI که به نظر می رسند در یک وظیفه خاص هوشمندانه عمل می کنند) و "AI قوی" (یک سیستم AI که طبق تعریف انسان، به صورت کلی هوشمند است) تقسیم می شود. مترادف های دیگری برای AI قوی مقابل AI ضعیف وجود دارد، که به عنوان مثال AI کلی مقابل AI بسته و یا AI عمومی که توسط بروسارد (2018، ص 32) به صورت "AI کلی نوع هالیوودی AI است... ماشین هایی که مثل انسان فکر می کنند. AI بسته متفاوت است؛ این یک روش ریاضی برای پیش بینی است.... AI کلی چیزی است که برخی از مردم می خواهند، و AI بسته آن چیزی است که ما داریم " توضیح داده شده است. در این کتاب، اصلی ترین تمرکزم بر روی AI موجود و به همین ترتیب بسته است.

اصطلاح AI برای اولین بار در تابستان سال 1956 میلادی توسط آمریکایی ها جان مک کارتی، ماروین ال امینسکی، ناتانیل روچستر و کلود شانون در پروژه تحقیقاتی تابستانی در زمینه هوش مصنوعی در دارتموث، نیوهمپشایر، ایالات متحده آمریکا ابداع شد. ال امینسکی از آن زمان تا به حال در بسیاری از تحولات و گفتمان های AI نقش داشته است. به عنوان مثال، با همکاری در طراحی خلاقانه AI شرور "HAL 9000" در فیلم کوبریک (1968)، سیبیل (Odyssey) فضایی. از آن زمان به بعد، AI چندین مرحله را طی کرده است، شروعی به شادی هم در زمینه آنچه ایجاد شده، تحقیقات و تامین مالی شده بود، داشت، اما همچنین دوره هایی را نیز تجربه کرده است - با عنوان "زمستان های AI" (از 1974 تا 1980 و از 1987 تا 1993).

هوش مصنوعی یک زیرمجموعه علوم کامپیوتر است-مطالعه کاربردهای عملی و نظری رایانه ها ، شامل سخت افزار ، نرم افزار و ریتم های الگو (دستورهای حل مشکلات خاص و انجام محاسبات). کلمه کامپیوتر در اصل برای توصیف کارگران انسانی که جداول ریاضی را ساخته و محاسبه کرده اند استفاده می شد(Broussard, 2018, p. 77).

شکل 1.1 هوش مصنوعی و زمینه های مرتبط.

در دهه های اخیر، مطالعه علوم کامپیوتر به عنوان یک رشته تحصیلی و شغلی، گسترش چشمگیری پیدا کرده است. هوش مصنوعی به عنوان زیر شاخه ای از این رشته تعریف شده است و خود دارای چندین زیر شاخه است، که به صورت شکل 1.1 نشان داده شده است. یکی از این زیرشاخه‌ها، یادگیری ماشین است، که در آن کامپیوترها از داده ها یاد می گیرند؛ به عبارت دیگر، از مجموعه داده ها مدل های آموزش داده شده می سازند. این یادگیری معمولاً به سه شکل انجام می شود: یادگیری نظارت شده (که در آن نشانه ها بیشتر با ارائه برچسب توسط انسان در مجموعه داده ها هدایت می شوند) ، یادگیری بدون نظارت (در آن کامپیوتر بدون دستورالعمل وارد فرآیند می شود) و یادگیری تقویتی (که با مجازات و پاداش برای پاسخ ها همراه است). در این زیرشاخه، یادگیری عمیق که یکی از شاخه‌های محاسباتی یادگیری ماشین است، پیدا می شود. الهام گرفته از شبکه‌های مغزی انسان در ترکیب پیچیدگی آن، یادگیری عمیق اطلاق می‌شود. علوم کامپیوتر و زیرشاخه‌های آن می‌توانند به داده‌های بزرگ، داده‌های بسیار بزرگی که به کامپیوترها اجازه می‌دهند الگوها را ببینند و براساس داده‌های گسترده‌ای که به آنها منتقل می‌شوند، راه حل‌هایی پیدا کنند، حمل می‌کنند. این کتاب بر روی هوش مصنوعی تمرکز دارد. اما به عقیده خود کتاب، هوش مصنوعی یک زیر شاخه با زیرشاخه های خود است که اصطلاحات استفاده شده برای توصیف این لایه های گوناگون پژوهشی، در طول زمان تغییر می کند. گرچه این کتاب در اصل یک کتاب فنی نیست، اما داشتن دانش فنی، به درک تاثیرات پیچیده اجتماعی این فناوری ها کمک می کند.

هوش مصنوعی به عنوان یک اصطلاح برای توصیف چیزی که نه "مصنوعی" و نه "هوشمند" است، مورد انتقاد قرار گرفته است، به طور مثال توسط کرافورد (2021) در کتاب Atlas of AI. کرافورد این ادعا را مطرح می کند که هوش مصنوعی از منابع طبیعی ساخته شده است و آنچه باعث مشاهده سیستم های خودکار به نظر می رسد، انسان ها هستند. به عنوان مثال، او توضیح می دهد که چگونه فردی که از دستیار صوتی هوشمند خود به نام Alexa یا Siri استفاده می کند، یک فرآیند را آغاز می کند که در آن منابع جمع آوری می شوند، یک شبکه لجستیک مورد استفاده قرار می گیرد تا درختان را بریده، آنها را به توالت کاغذ مبدل، بسته بندی کند و آن را به نشانی شما بفرستد - به علاوه افرادی که در فرایند لجستیکی شرکت دارند و برنامه نویسی و پشتیبانی از تکنولوژی های استفاده شده. دستیارهای دیجیتالی مثل Alexa، Siri و دستیارهای دیگر، ساخته شده اند توسط انسان ها برای انسان ها و نه جادوگران جنی درون جعبه هستند. با این حال، فرآیندهایی که توسط هوش مصنوعی عرضه یا بزرگنمایی شده‌اند - منابع طبیعی مصرف شده، افرادی که به تکنولوژی دسترسی دارند و کسانی که تکنولوژی برای آنها سفارشی شده است - نشان می دهند که هوش مصنوعی در انتخاب خود بی تمایل نیست. کرافورد ادعا می کند که ما باید بپرسیم که چه کسی از هوش مصنوعی بهره می برد و چه کسی از قدرت محروم می شود. با توجه به این فراخوان، در این کتاب به بررسی تأثیرات هوش مصنوعی بر گروه های متفاوت از افراد پرداخته شده است، که با مفهوم "هوش مصنوعی تقویت شده" (Kurpicz-Briki، 2021؛ Rui، 2017؛ Zheng et al.، 2017) شباهت زیادی دارد، که یک رویکرد هوش مصنوعی مبتنی بر انسان است و نه بر اینکه چگونه هوش مصنوعی انسان ها را جایگزین می کند، بلکه چگونه می توانیم هوش مصنوعی را برای تقویت و پشتیبانی از تصمیم گیری های انسانی استفاده کنیم.

**یادگیری ماشینی هوش مصنوعی در عمل**

بیایید با یک آزمایش فکری شروع کنیم. به یک سیستم فکر کنید که آموزش دیده است تا تصاویر سگ و گربه را با تمیزی و درستی تشخیص دهد و از یکدیگر جدا کند. انسان ها ممکن است یک هزار تصویر اولیه با شناسه های درست را در مجموعه داده آموزشی - به نام حاشیه دادن اطلاعات - دسته بندی کنند، قبل از آن که سیستم برای تصمیم گیری در مورد تصاویر جدید تنظیم شود. مشترکاتی مانند شکل گوش، چشم، مقدار مو و غیره ممکن است برای این منظور استفاده شود که گربه ها را از گونه های سگی جدا کند. با این حال، این تنها ساده سازی است و پشت یک چنین فناوری ، چند لایه پیچیده ای وجود دارد (Castelvecchi، 2016). نرم افزار شناسایی تصویر نیازمند نگاه خاصی به تصاویر نیست که همانند چشم بشر نگاه کند. این سیستم می‌تواند به خوبی کار کند، به جز زمانی که مشکل پیش می‌آید. به عنوان مثال، حیوان خانگی کانادایی "آچوم"، مشاهده می‌شود در شکل 1.2 و در یک توییتر (Twitter، 2016) نمایش داده شده است. آچوم یک سگ است یا یک گربه؟ اگر گیج شدید، سیستم مرتب‌سازی هوش مصنوعی هم ممکن است گیج شود. پاسخ صحیح آن است که این توپ مویی گربه دارای سندرم گرگ مو است که باعث رشد سریع‌تر موهایش می‌شود.

موجوداتی مانند "آچوم"ممکن است مردم را گیج کند، اما AI چه؟ سیستم های AI با روش های گوناگونی به تقسیم بندی کردن دسته بندی های انسانی نمی پردازند و ممکن است توسط موارد کاملاً متفاوت گیج شوند. به عنوان مثال، یک مطالعه نشان داد دقت بسیار بالا در مرتب کردن عکس های گرگ و سگ را، به جز زمانی که گرگ ها در تابستان و سگ ها در زمستان نشان داده شده اند. چرا؟ این سیستم یاد گرفته بود مناظر برفی که اغلب با گرگ ها همراه است را تشخیص دهد، بنابراین بجای ساخت یک تشخیص دهنده "سگ یا گرگ"، برنامه نویسان به جای آن یک تشخیص دهنده "برف یا نه" ساختند(Roselli et al.،2019). این مثال های حیوانات ممکن است به عنوان چیزی جذاب و خنده دار تلقی شوند، اما وقتی مردم با زندگی، حقوق، ارزش ها و آینده خود در مرکز آن چیزی هستند که توسط AI تغذیه و محاسبه و تفسیر می شود، چه؟ برای درک بیشتر این مسأله، بیایید به اینکه یک سیستم AI چگونه "تصمیم می گیرد" بنگریم و چگونه این می تواند منجر به تبعیض شود. به عنوان مثال، ریچاردسون (2021، ص 13) سیستم های تصمیم گیری خودکار را به شرح زیر تعریف کرده است:

سیستم ها، نرم افزارها یا فرآیندهایی که در آن محاسبات برای کمک به تصمیم گیری، ارزیابی و / یا پیاده سازی سیاست های دولتی که در مسائلی از قبیل فرصت ها، دسترسی، آزادی ها، حقوق و / یا ایمنی تأثیر می گذارد، به کار می روند. سیستم های تصمیم گیری خودکار می توانند شامل پیش بینی، طبقه بندی، بهینه سازی، شناسایی و / یا توصیه شوند.

در سطح جامعه ، سیستم های AI ممکن است تصمیماتی اتخاذ کنند که بر تمام جامعه تأثیرگذار باشد و نه فقط برای یک سیستم شناسایی سگ یا گربه. قبل از ورود به بحث در مورد چه اتفاقاتی در سطح جامعه رخ می دهد ، ابتدا مفهوم "انسان بودن" را به عنوان "یک قالب برای همه" برای استفاده از مفهوم تنوع باز خواهم کرد.

**تنوع چیست؟**

تنوع می‌تواند به عنوان "شرایط داشتن یا تشکیل شده از عناصر متفاوت باشد، بویژه شامل گروه‌های مختلف قومیتی، فرهنگی و غیره" تعریف شود (دیکشنری مریم-وبستر، 2022b). مفهوم تنوع به‌طور کلی مشخصه‌های شخصی متنوعی از جمله جنسیت، جهت‌گیری جنسی، سن، نژاد/قومیت، معلولیت‌های جسمی / ذهنی و سلامتی، زمینه‌های اقتصادی و دین/مذهب را در نظر می‌گیرد. این تفاوت‌ها در انسان‌ها چیزی است که آنان را متمایز می‌کند اما از نظر مثبتی مورد توجه قرار می‌گیرد. مطالعات نشان داده‌اند که بیشترین تنوع در کار (Ilmakunnas & Ilmakunnas, 2011)، آموزش (Smith et al., 1997)، تولید دانش (Zhang & Guo, 2019) و بهداشت (Salisbury & Byrd, 2006) به شدت ضروری هستند.

مفهوم تنوع از طریق مدل‌های آماری (سیمپسون، 1949)، تحلیل اقتصادی (ویتزمن، 1992)، و توجه شرکتی (هانت و همکاران، 2015) به شمارش رسیده و تلاش شده است تا درک بهتری از آن داشته باشیم. در سال‌های اخیر، تمرکز بر تنوع نیروی کار به خصوص به عنوان یک موضوع در مدیریت منابع انسانی، پروفایل شرکت و رهبری مبتنی بر ارزش‌ها ظاهر شده است و برای دانشمندان کار، به عنوان مثال تاتلی (2006) (Handbook of Workplace Diversity) به عنوان یک تمرکز مطرح شده است. ساکسنا (2014، ص 76) تنوع نیروی کار را به عنوان "تشابهات و تفاوت‌های بین کارکنان از نظر سن، سوابق فرهنگی، توانایی‌ها و ناتوانی‌های جسمی، نژاد، دین، جنسیت و جهت‌گیری جنسی" تعریف می‌کند.

انواع مختلف برنامه‌های تنوع و شامل بودن انتقاداتی را به دلیل توجه بیش از حد به صفات سطحی افراد، مانند نژاد و جنسیت، دریافت کرده است. در واقع، این برنامه‌ها بسته به محیط و فرهنگی که در آن استفاده می‌شوند، متفاوت می‌باشند. برنامه‌های تنوع و شامل بودن بنابراین به مرور زمان، همچنان مورد استفاده، چالش قرار می‌گیرند، و مفهوم آن‌ها بازتعریف می‌شود. در این کتاب با اینکه از انتقادات مطرح شده برای مفهوم تنوع آگاه هستیم، با اعتراف به مسائل مطرح شده، آن‌ را به عنوان نقطه شروعی برای بررسی تأثیر هوش مصنوعی بر جامعه و افراد، استفاده خواهیم کرد.

تمرکز جهانی بر مسائل تبعیض و تنوع، به وضوح در برنامه توسعه پایدار سازمان ملل متحد برای سال 2030 قابل مشاهده است. در این برنامه، فرصت‌های شامل بیشتر و فراگیرتر برای گروه‌های اجتماعی مختلف در اهداف پایدار توسعه، از جمله هدف ۵ برابری جنسیتی، هدف ۱۰ کاهش نابرابری‌ها و هدف ۱۱ شهرهای پایدار، بسیار مهم است. همچنین، تاریخ انسانی نشان داده است که هرجایی که تبعیض وجود داشته باشد، نابرابری‌ها و مبارزه برای حقوق گروه‌های محروم به هم گره خورده اند. در سال‌های اخیر، این مبارزات برای عدالت آنتی-تبعیضی در جامعه، به ویژه با حضور هوش مصنوعی، در جنبش "Black Lives Matter" سال 2020 در محاکمات پس از کشته شدن جورج فلوید توسط پلیس و در انتخابات کاپیتولیوم سفید در سال 2021، مشاهده شده است که با استفاده از هوش مصنوعی، تظاهرات "Black Lives Matter" از سوی افراد نژادپرست به شکل دقیق‌تری مشخص شده‌‌اند. همچنین، در سال 2015، درباره بحران پناهجویان اروپا، بحث‌هایی درباره فناوری تشخیص چهره برقرار شده و نگرانی‌های عمیقی در مورد سیستم های هوش مصنوعی طراحی شده برای جنگ وجود داشته‌است. این حرکات معاصر برای حقوق پایداری، با نگرشی آنتی-تبعیضی، جایگاه مهمی برای هوش مصنوعی در بر دارند. به طور خلاصه، دنیای بسیار جهانی ما، به سوی کسب آگاهی از مسائل تنوع حرکت می‌کند و این جنبش‌ها، در جوامع دیجیتالی ما، که به ویژه در تکنولوژی دیجیتالی، مانند رسانه‌های اجتماعی، نقش مهمی در فرآیندهای دموکراتیکی بازی می‌کنند (Hall et al.، 2018). اما این مسئله چگونه با هوش مصنوعی مرتبط است؟

کلمات و زبان در ساختارهای قدرت قرار دارند و در موضوعات حساس مانند شامل شدن و تنوع، خطری از تولید مشکلات اجتماعی یا استفاده از اصطلاحاتی که به سرعت تبدیل به منسوخ می‌شوند، وجود دارد. با این حال، باید به خاطر داشته باشیم که ما با اصطلاحاتی که در دسترس ما هستند در زمان مشخصی کار می‌کنیم. اگرچه مفهوم تنوع مشکلات خود را دارد، اما هنوز هیچ اصطلاحی ثابت وجود ندارد که قدرتمندانه بتواند به همه مردم، زندگی‌ها، شیوه‌های زندگی، شخصیت‌های مختلف و مباحث پیرامون آن‌ها پرداخت کند. به طور مشابه، اصطلاحات "نژاد"، "طبقه" و "جنسیت" بسیار بحث‌برانگیز هستند، بنابراین بعضی از افراد ممکن است این اتفاق را خودشان به تنهایی بحث‌برانگیز بدانند. در این کتاب، تمرکز بر تولید فرضیات دسته‌بندی شده نمی‌باشد، اما به بررسی این موضوعات خواهیم پرداخت چگونه این فرضیات، ساز و کارهای هوش مصنوعی را در درک جهان به کار می‌برند. لذا به دنبال استفاده از اصطلاحاتی هستیم که به ما در بررسی مفاهیم کمک می‌کنند. به عنوان مثال، به دنبال بررسی اینکه چرا در یک سو، هوش مصنوعی تشخیص چهره همه‌گیر برای شناسایی مشتریان سیاه‌پوست ناموفق بوده، اما در سوی دیگر، برای شناسایی معترضان سیاه خوب عمل می‌کند، اصطلاحاتی را نیاز هست که به ما کمک می‌کنند در درک رویکردهای سوسیوتکنیکی در مورد مسائل تنوع در هوش مصنوعی پیشرفت کنیم. در نتیجه، من شما را به شدت ترغیب می‌کنم که این کتاب را به عنوان یک موضوع برای شروع بحث خود استفاده کنید، اما به خاطر داشته باشید که اصطلاحات تغییر خواهد کرد، و جایگزینی بایستی با پیشرفت درک ما از رویکردهای اجتماعی-فنی در مسائل تنوع در هوش مصنوعی ایجاد شود.

**چرا هوش مصنوعی برای تنوع یک رویکرد اجتماعی-فنی**

درک چگونگی توسعه، استقرار و استفاده از سیستم‌های فناوریک پیچیده است، به خصوص در مورد سیستم‌های هوش مصنوعی. به همین ترتیب، درک چگونگی تأثیر سیستم‌های هوش مصنوعی بر جوامع، گروه‌ها و افراد نیاز به تحلیل جامع در چندین سطح دارد. سیستم‌های هوش مصنوعی توسط افراد خاصی در زمان‌ها و فرهنگ‌های خاصی ساخته می‌شوند. آنها هم می‌توانند ابزاری برای سرکوب سیستماتیک باشند (به عنوان مثال از طریق نظارت) و هم ابزاری برای کمک به شناسایی اقدامات تبعیض‌آمیز باشد (به عنوان مثال از طریق بهینه‌سازی هشتگ و به اشتراک گذاری توسط حرکات سیاسی)

برای درک اینکه چگونه جامعه توسط فناوری تأثیر می‌پذیرد و همچنین چگونه فناوری هم به جامعه تأثیر می‌گذارد، رویکرد سوسیوتکنیکی می‌تواند مفید باشد - برای درک ارتباطات پیچیده بین جنبه‌های اجتماعی و فنی. علاوه بر این، هوش مصنوعی به عنوان علمی "سرد و بی طرف" در رشته‌های فنی اعتبار دارد، در حالی که مطالعات تنوع معمولاً در علوم اجتماعی و علوم انسانی یافت می‌شوند. با این حال، مطالعات بین‌رشته‌ای که اطلاعات و بینش را از چندین رشته گرفته و به کار می‌برد، به عنوان راهکاری برای درک مسائل پیچیده، همچنین در زمینه‌های فنی مانند برنامه نویسی کامپیوتر و هوش مصنوعی، اهمیت بیشتری به‌دست می‌آورند.

یکی از رشته‌هایی که تحقیقاتی در زمینه‌ی تولید و علم و فناوری انجام داده است، مطالعات علم و فناوری (STS) است. STS از دهه‌ها پیش برای ساخت اجتماعی تکنولوژی (کالون، 1987؛ لاتور، 1987)‌‌ مطرح بوده است - با دیدن تکنولوژی که به شکلی نادرست در جهت توسعه و اجرا قرار نگرفته بلکه به عنوان نتیجه‌ی فرآیندهای اجتماعی و مذاکرات پایه‌ای و تحصیلاتی دیده می‌شود. به عنوان یک دانشمند STS، این پس زمینه نظری را برای تحلیل موضوع این کتاب استفاده می‌کنم. یکی از گام‌های مهم تحقیقات STS، باز کردن جعبه‌های سیاه تکنولوژی است تا نه تنها ببینیم چگونه جامعه تکنولوژی را شکل می‌دهد، بلکه چگونه تکنولوژی نیز جامعه را شکل می‌دهد. STS برخی ویژگی‌های مشترک با رشته‌های مشابهی دارد مانند علم علم (که علمی زیرپشتیبانی AI را بررسی می‌کنند) (یوان و همکاران، 2020).

من از یک روش دیدگاه جندر و تنوع از فناوری-فمینیستی استفاده می کنم که نحوه توصیف و پیشنهاد هویت و جنسیت در زمینه های علم و فناوری را بررسی می کند (مانند نظریات باراد از سال ۲۰۰۳، هاراوی از سال ۲۰۰۶، و وجکمن از سال ۲۰۰۷)، تا به مسائل تنوعی که هوش مصنوعی می تواند به همراه داشته باشد، بپردازم."

در مسیر به دنبال هوش مصنوعی، مهم است که از عوامل اجتماعی آگاه باشیم؛ به عنوان مثال، همانطور که بروسارد درباره آزمون تورینگ (Turing Test) دلسوزی کرده است. در این آزمون، دو نفر به نام‌های A و B که به ترتیب مرد و زن هستند، از طریق متون نوشتاری با شخص C، یک سومیاتر جنبه به چالش کشیده، در ارتباط هستند. پس از دریافت پیام‌ها، سومیاتر C باید تشخیص دهد که کدام یک از A و B مرد و کدام یک زن است. امکان دارد A به C دروغ بگوید، در صورتی که B باید به C کمک کند. به عبارت‌ دیگر، تورینگ سپس سوال می‌پرسد که اگر A با یک ماشین جایگزین شود، چطور باشد؟ آیا این عادلانه خواهد بود که C بتواند انسان را از یک ماشین تشخیص دهد؟ بیشتر مردم این آزمون را به این صورت می‌شناسند که A با یک ماشین جایگزین شده است، و سپس سومیاتر باید بفهمد کدام یک از اشخاص ماشین است و کدام یک اسم انسان دارد. بروسارد نسبت به پیش فرض این آزمون تورینگ (Turing Test) تحت تاثیر در نمی آید. (صفحات ۳۷-۳۸، ۲۰۱۸) :

مشخصات تورینگ با آنچه در حال حاضر درباره جنسیت درک می شود، مطابقت ندارد. آیا بازی برای تعیین "هوش" است، در آن زن به عنوان کمک کننده تعیین شده است؟ و به مرد گفته می شود که دروغ بگوید؟ این پایه ها کاملاً سورئال هستند با صفات جسمی و اخلاقی که به لحاظ جنسیتی کد شده اند.

بیایید به اصل مسأله هوش مصنوعی و سایر عوامل اجتماعیی که در توسعه علم کامپیوتر و ایجاد هوش مصنوعی نقش داشته‌اند، توجه کنیم. اگر بنیاد راکفلر در سال ۱۹۵۶ تصمیم گرفته بود که بورس‌های تحقیقات تابستانی در دارتموث را ممکن کند یا اگر اعضای بنیاد آن‌را تعویض کرده بودند، دنیای کامپیوتر و هوش مصنوعی می‌توانست متفاوت باشد. اگر ادا لاولیس - به عنوان نخستین برنامه نویس زن دنیا - به دلیل بودن زن (پدرش می خواست فرزند پسرش یک افسانه‌ای باشد) به مشکل نخورده بود، یا صد سال پس از آن آلن تورینگ به دلیل هوموفوبیا (ترس از همجنس‌گرایی) در سال ۱۹۵۲ با سم زدایی از سر جان باخته نبود، همه چیز می‌توانست متفاوت باشد. همانطور که در اصطلاحات STS (STechnology and Science studies) تعریف شده است، مسأله این است که "آینده ممکن است متفاوت باشد" (یا اختصارا ICBO) و این به ما اجازه می دهد تا به صورت منتقدانه به دنیا، جوامع و فناوری نگاه کنیم، نه به صورت متغیرهای ثابت و ثابت، بلکه به عنوان چیزی که به دست شرایط اجتماعی، ساخته شده است. (وولگار، ۲۰۱۵؛ ووولگار و کوپر، ۱۹۹۹)

مشابها، هوش مصنوعی می تواند زمینه اجتماعی را شکل دهد و به عنوان یک پرتوانا، قدرتی دارد که به منظور حفظ یا تغییر روابط قدرت و شرایط اجتماعی استفاده شود (Ranchordás & Roznai، 2020؛ Richardson، 2021). این خصوصا زمانی برای هوش مصنوعی قابل مشاهده است که به مسایل تنوع و خلق فن آوری بی طرفانه و شامل توجه پیدا می کنیم. فن آوری مشخص شده استفاده تنها برای یک مورد است. دیدگاه یکپارچه فن آوری اجتماعی به ما اجازه می دهد باز ارزیابی کنیم که چگونه جوامع از آن استفاده می کنند و به ما انتخاب هایی مبتنی بر نیازها را می دهد تا بتوانیم از آن در چه شرايطی، به چه شکلی و در چه موارد استفاده نکنیم. بروسرد (Broussard)، ۲۰۱۸، ص ۸۷ می گوید: "برای خلق یک جهان فناوری منصفانه، وقتی که می خواهیم به فناوری دست بزنیم، باید صدای متنوعی از افراد در میان داشته باشیم." بنابراین، برای اینکه فن آوری هایی مانند هوش مصنوعی شامل تنوع باشند، باید آن ها را برای گروه های مختلف جمعیتی طراحی کنیم.

**پیامدهای اجتماعی تبعیض هوش مصنوعی**

هوش مصنوعی تنها در محیطی خاص ساخته نمی‌شود، بلکه در تمامی جوامع تأثیرات وسیعی دارد. آیا برخی از این تأثیرات باعث ضرر و بروز مشکلات در برخی افراد می‌شود؟ کتاب "Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech" نوشته واکتر-بوئچر (Wachter-Boettcher)، یک بررسی گسترده از تبعیض سیستماتیک و فن آوری‌های دیجیتال تبعیض آمیز است و خواستار آن است که زمان آن رسیده است تا شرکت‌های فناوری بیشتری مسئولیت بر تکنولوژی‌هایی که ساخته‌اند بگیرند و درک بیشتری از نحوه تأثیر گذاری تکنولوژی روی گروه‌های مختلف قومیتی داشته باشیم.

به طور کلی، تبعیض در هوش مصنوعی در سه مرحله اصلی، یعنی قبل، در حین و بعد از پردازش داده ها، ممکن است رخ دهد. به عنوان مثال، تبعیض تنوع می‌تواند در صورتی رخ دهد که داده‌های انتخاب شده برای یک برنامه یا الگوریتم، داده‌های کافی و تنوع‌مند را نشان ندهد. الگوریتم یک فرایند کامپیوتری است که دنباله‌ای از دستورات را برای حل مسائل خاص دنبال می‌کند. آنها می‌توانند بسیار ساده باشند، مانند الگوریتمی برای مرتب سازی اعداد تصادفی به ترتیب صعودی، یا بسیار پیچیده باشند، و به صورت ترکیبی با هم معرفی شوند، به عنوان مثال برای امکان پخش ویدیو با کیفیت بلوری با وضوح بالا. الگوریتم‌ها قدم‌هایی هستند برای حل یک مسئله، برای مثال دستورالعمل‌ها برای حل چند روش اروبیک می‌تواند به عنوان الگوریتم در نظر گرفته شود. انتخاب بهترین الگوریتم مهم است، اینکه شما در حال مرتب سازی صد عدد هستید، آیا باید قبل از شروع مرتب‌سازی اول لیست را بررسی کنید یا به ترتیب اعداد را بررسی کنید و هر چند عدد وارد می شود لیست را به عنوان یکی از ورودی‌ها اضافه کنید؟ و اگر لیست نه تنها از صد عدد متفاوت بلکه از میلیاردها عدد متفاوت و بلند با اعشار میلیارد تراز است چه؟ به عبارت دیگر، شما به یک برنامه مرتب سازی سریع، کارآمد و درست، یعنی یک الگوریتم، نیاز دارید.

مرتب سازی اعداد یکی از مثال‌های ساده‌تر الگوریتم‌ها در جامعه معاصر است؛ الگوریتم‌ها سیستم‌های تصمیم گیری هوش مصنوعی پیچیده‌تر را تحریک می کنند. در فرایندهای پیچیده، پیدا کردن خطاها یا درصورتی که چیزی ایراد داشته باشد، ممکن است دشوار و وقت‌گیر باشد. از طرف دیگر، وقتی یک الگوریتم "نابغه" نوشته شده و در یک سیستم پیاده سازی می‌شود، اگر به نظر بیش از حد درست و کارآمد بنظر بیاید، انسان‌ها هم ممکن است نسبت به حرکات الگوریتم کور باشند. در مورد تِی (Tay)، ربات توییتری مایکروسافت که نژادپرستی شد و باید در روز عرضه‌اش متوقف شود، ولف و همکاران (2017، صفحه ۱) می‌بینند از نظر سیستم‌های یادگیری که "با مردم به‌طور مستقیم یا غیرمستقیم از طریق رسانه‌های اجتماعی در ارتباط هستند، توسعه دهنده مسئولیت‌های اخلاقی اضافی بیشتری علاوه بر نرم افزارهای استاندارد دارد. بار مسئولیت بیشتری وجود دارد." ولی اگر چیزی را نبینیم که اشتباه باشد و نتایج را عالی و پیاده‌سازی‌پذیر تلقی کنیم، چه باید انجام دهیم؟

اطلاعات می‌توانند با ورودی و پردازش عادلانه، در خروجی هنوز به نادرستی استفاده یا به صورت نابعادلانه تفسیر شوند. چطور؟ سیستم‌های هوش مصنوعی ممکن است دچار جنبه‌های تحریف شده و یا بر روی نتایج نژادپرستانه تمرکز کنند، اما بیشترین سوء استفاده در این مرحله، ناشی از تعصب انسانی است. اگر انسان‌ها به دنبال نتایجی علیه گروه خاصی هستند، سیستم‌ها می توانند به گونه ای تنظیم شوند که نتایجی تولید می‌کنند که به گروه خاصی ضرر می‌زنند. به جنبش یوژنیک، دروینیست‌های اجتماعی و یا نژادپرستی علمی‌ فکر کنید. این مسئله در آلمان نازی به شکلی وحشتناکی حساب شد که درواقع بر اساس تئوری‌های علمی آن زمان، به دنبال داده‌های (در قالب داده های غیر واقعی تراکمی) بودند که بتوانند اثبات کنند که مردان سفیدپوست، قادربدن و غیره، بهتر از همه دیگر هستند. چون عصر دیکتاتوری ها هنوز به پایان نرسیده است، باید پرونده سوء استفاده‌ای از داده‌ها را که در عصر هوش مصنوعی و داده‌های بزرگ مربوط به منافع بزرگتری هم هستند، کاملاً بررسی کنیم. برداشت‌های داده‌ها قابل دروغ‌گویی هستند؛ با رهبرانی که قدرت مسلط شده را سوء استفاده می کنند، در جامعه معاصر نیز روبرو هستیم. ریچاردسون (2021، ص. 1) هشدار می‌دهد که چگونه این موضوع را باید در نظر داشت.

سازمان‌های دولتی برای کمک به تصمیم‌گیری و اجرای سیاست در حوزه‌های اجتماعی حساس، از سیستم‌های تصمیم‌گیری خودکار برای جایگزین کردن یا کمک به تصمیم‌ گیری انسانی استفاده می کنند. این سیستم‌ها تعیین می‌کنند کدام فردی حق یارانه غذایی خود را از دست خواهد داد، چه مقدار اولویت دریافت خدمات درمانی دارد و کدام فرد به احتمال زیاد قربانی جرم خواهد شد.

با این حال، نتایج سیستم‌های تصمیم‌گیری خودکار اغلب بیشتر از حد قدرت واقعی آنها تخمین زده می‌شود و هزینه‌های اجتماعی کامل به ندرت شناخته یا نادیده گرفته می‌شوند. سیستم‌های هوش مصنوعی در محیط‌های پیچیده به مسائل اعتماد بسیار بزرگی منجر می‌شوند ، اگر برای پیاده‌سازی آن‌ها، زودهنگام باشیم. به عنوان مثال، به خودروهای خودران و آزمون‌های خودروبر به عنوان خودکار، فکر کنید. چند اتفاق ناگوار باعث شده است که محققان سوالاتی مطرح کنند که آیا جامعه برای آزمایش‌های خودران آماده است؟ آیا می توانیم با پذیرش اینکه وسیله نقلیه برای فن آوری هایی که در صورت بروز خطا زندگی تهدید کننده باشد آزمایشگاهی باشد، بدون خط و مسئولیت واضح قبول داشته باشیم؟ بدون نیاز به چارچوب‌های اخلاقی بازنگری شده (Evans et al., 2020).

شکل در مرحله خروجی داده در اصل یک تعصب در هوش مصنوعی نیست بلکه تعصب انسانی و نحوه تفسیر داده‌ها از سیستم‌های ماشین است. هوش مصنوعی در رابطه با عمدتاً ارتکاب آسیب، تعصب ندارد و سیستم‌ها با مدل دستوری و گام به گام وظایف را حل می‌کنند. با این حال، انسان‌ها پیچیده‌تر هستند. اصلی‌ترین عاملِ تعصب در تفسیر انسان را به دو دستهٔ کمبود دانش و کمبود قلب تقسیم می‌کنیم. کمبود دانش وقتی رخ می‌دهد که ما داده‌ها را درک نمی‌کنیم، به سیستم اعتماد کرده و به مسئولیت ما برای بررسی احتمال وجود خطا ها فکر نمی کنیم. روش های کاهش این مشکل به شامل آموزش، آگاهی از مسئولیت و داشتن مکانیسم های کنترل می‌شوند. تعصب دیگرِ خروجی داده - زمانی رخ می‌دهد که انسان‌ها می دانند که سیستم‌های کامپیوتری با خطا مواجه شده‌اند، اما همچنان از داده‌ها به طوری که برای آن‌ها، شرکت خودشان یا دیگر تحریکات تفسیر داده‌ها منافع دارد، استفاده می‌کنند. این اتفاق به خصوص خطرناک است، زیرا ممکن است برای توجیه بی‌ارادی سیاست‌ها و قوانین جامعه به کار رود که بازهم به این دلیل از خوبی استفاده می شود که سیستم به آن‌ها به طور دقیق گفته است چه چیزی دقیق است و چی نیست.

**خلاصه فصل**

این فصل چندین تعریف مختلفی از هوش مصنوعی ارائه داده است - زیرا هنوز هم توسط کارشناسان مورد بحث قرار می گیرد. در کل، هوش مصنوعی به عنوان شبیه سازی کردن یا داشتن هوش انسانی تعریف می شود. ما بین هوش مصنوعی عمومی / قوی که هنوز واقعیت نیافتاده است، و توانایی انجام کارهای خاصی که هوش های محدود / ضعیف را هوشمندانه انجام می دهند، تمایز قائل هستیم که این سیستم های هوش مصنوعی در حال حاضر وجود دارند. هوش مصنوعی یک حوزه در علوم کامپیوتر است که به منابع فرعی شبکه های عصبی و یادگیری ماشینی نیاز دارد و قادر است از داده های بزرگ بهره ببرد. علاوه بر این، من از طریق رویکرد فرهنگی - فنی مورد بررسی قرار دادم که چگونه تنوع - شامل افراد با زمینه های مختلف - می تواند بر توسعه و استفاده از هوش مصنوعی در جامعه تأثیر بگذارد. من ادعا کردم که هوش مصنوعی به عنوان یک اشیای فنی همچنین واقعیت های اجتماعی را ساختار می دهد که از تعصب خالی نیست و باید جعبه های سیاه فناوری را باز کرد تا درک جامعی از نحوه تأثیر فنی و اجتماعی روی همدیگر را به دست آوریم. در این فصل، سه مکانی که تعصب در هوش مصنوعی اغلب رخ می دهد معرفی شد: (1) تعیین مقدار داده ها، (2) تعصب در پردازش داده ها، و (3) تعصب در خارج از داده / پردازش داده ها. اگرچه دو مورد اول عمدتاً درباره مشکلات انتخاب و مشکلات در پردازش داده هاست، اما دسته سوم عمدتاً درباره تعصب انسان در تفسیر داده ها است. در فصل های بعدی، چندین مثال از اینکه چگونه تعصب می تواند در فرآیندهای پیچیده فرهنگی - فنی بین انسان و ماشین ها رخ دهد و شتاب بخشد را خواهیم دید. به سوی پایان کتاب، همچنین خواهیم دید که چگونه تعصب تنوع و مکانیسم های اخراجی می تواند کاهش یابد.

**مرجع ها**

Anjila, P. F. (1984). What is artificial intelligence? In J. Karthikeyan, Su-Hie Ting, &Yu-Jin Ng (Eds.), *Success is no accident. It is hard work, perseverance, learning, studying, sac- rifice and most of all, love of what you are doing or learning to do* (p. 65). L’ Ordine Nuovo Publication. [https://www.researchgate.net](https://www.researchgate.net/)

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society*, *28*(3), 801–831.

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.

Callon, M. (1987). Society in the making: The study of technology as a tool for sociological analysis. In Wiebe E. Bijker, Thomas P. Hughes, & Trevor Pinch (Eds.), *The social construction of technological systems*: *New directions in the sociology and history of technology* (pp. 83–103). The MIT Press.

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature News*, *538*(7623), 20.

Cho, J. H. (2016). *“Diversity is being invited to the party; inclusion is being asked to dance,” Verna Myers tells Cleveland Bar*. Retrieved December 6, 2021 from [https://www.](https://www.cleveland.com/) [cleveland.com/business/2016/05/diversity\_is\_being\_invited\_to.html](https://www.cleveland.com/)

Cohen, T. (2019). How to leverage artificial intelligence to meet your diversity goals. *Strategic HR Review*, 18, 62–65.

Crawford, K. (2021). *The atlas of AI*.Yale University Press.

Dräger, J., & Müller-Eiselt, R. (2020). *We humans and the intelligent machines: How algo- rithms shape our lives and how we can make good use of them*.Verlag Bertelsmann Stiftung. Evans, K., de Moura, N., Chauvier, S., Chatila, R., & Dogan, E. (2020). Ethical decision making in autonomous vehicles: The AV ethics project. *Science and*

*Engineering Ethics*, *26*(6), 3285–3312.

Hall, W., Tinati, R., & Jennings, W. (2018). From Brexit to Trump: Social media’s role in democracy. *Computer*, *51*(1), 18–27.

Haraway, D. (2013). A manifesto for cyborgs: Science, technology, and social- ist feminism in the 1980s. In Linda Nicholson (Ed.), *Feminism/Postmodernism* (pp. 190–233). Routledge.

Hunt, V., Layton, D., & Prince, S. (2015). Diversity matters. *McKinsey & Company*, *1*(1), 15–29.

Ilmakunnas, P., & Ilmakunnas, S. (2011). Diversity at the workplace: whom does it benefit? *De Economist*, *159*(2), 223–255.

Kubrick, S. Director (1968). *2001: A space odyssey* [Film]. Insight Media. Stanley Kubrick Productions.

Kurpicz-Briki, M. (2021). *Is augmented intelligence the AI of the future?* Retrieved March 18, 2022 from [https://www.societybyte.swiss/en/2021/04/29/is-](https://www.societybyte.swiss/) [augmented-intelligence-the-ai-of-the-future/](https://www.societybyte.swiss/)

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B. (1999). *Pandora’s hope: Essays on the reality of science studies*. Harvard University Press.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., & Darzi, A. (2020). International eval- uation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89–94. Merriam-Webster-Dictionary. (2022a). *Artificial Intelligence*. Retrieved October 29, 2022 from [https://www.merriam-webster.com/dictionary/artificial%20](https://www.merriam-webster.com/)

[intelligence](https://www.merriam-webster.com/) Merriam-Webster-Dictionary. (2022b). *Diversity*. Retrieved November 29, 2021 from: [https://www.merriam-webster.com/dictionary/diversity](https://www.merriam-webster.com/)

Oxford-English-Dictionary. (2022). *Artificial Intelligence*. Retrieved November 29, 2021 from [https://www.oed.com/viewdictionaryentry/Entry/271625](https://www.oed.com/)

Pisano, E. D. (2020). AI shows promise for breast cancer screening. *Nature*, 577, 35–36. <https://doi.org/10.1038/d41586-019-03822-8>

Ranchordás, S., & Roznai, Y. (2020). *Time, law, and change: An interdisciplinary study*. Bloomsbury Publishing.

Richardson, R. (2021). Defining and Demystifying Automated Decision Systems.

*Maryland Law Review, Forthcoming*.

Roselli, D., Matthews, J., & Talagala, N. (2019). *Managing bias in AI*. *Companion Proceedings of the 2019 World Wide Web Conference*. San Francisco USA.

Rui, Y. (2017). From artificial intelligence to augmented intelligence. *IEEE MultiMedia*, *24*(1), 4–5.

Salim, M., Wåhlin, E., Dembrower, K., Azavedo, E., Foukakis, T., Liu,Y., Smith, K., Eklund, M., & Strand, F. (2020). External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammo- grams. *JAMA Oncology*, *6*(10), 1581–1588.

Salisbury, J., & Byrd, S. (2006). Why diversity matters in health care. *CSA Bulletin*, *55*(1), 90–93.

Saxena, A. (2014). Workforce diversity: A key to improve productivity. *Procedia Economics and Finance*, *11*, 76–85.

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*(4148), 688–688. Smith, D. G., Gerbick, G. L., Figueroa, M. A., Watkins, G. H., Levitan, T., Moore,

L. C., Merchant, P. A., Beliak, H. D., & Figueroa, B. (1997). *Diversity works: The emerging picture of how students benefit*. ERIC.

Stilgoe, J. (2019). Self-driving cars will take a while to get right. *Nature Machine Intelligence*, *1*(5), 202–203.

Tatli, A. (2006). Handbook of workplace diversity. *Equal Opportunities International*, 25(2), 160–162. <https://doi.org/10.1108/02610150610679583>

Twitter. (2016). *Her.- do you have a dog or a cat? me.- I don’t know*. Retrieved November 29, 2021 from [https://twitter.com/1evilidiot/status/](https://twitter.com/) [794613309613821952?ref\_src=twsrc%5Etfw](https://twitter.com/)

Wachter-Boettcher, S. (2017). *Technically wrong: Sexist apps, biased algorithms, and other threats of toxic tech*. WW Norton & Company.

Wajcman, J. (2007). From women and technology to gendered technoscience.

*Information, Community and Society*, *10*(3), 287–298.

Wang, B., Jin, S.,Yan, Q., Xu, H., Luo, C.,Wei, L., Zhao,W., Hou, X., Ma, W., & Xu,

Z. (2021). AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system. *Applied Soft Computing*, *98*, 106897.

Weitzman, M. L. (1992). On diversity. *The Quarterly Journal of Economics*, *107*(2), 363–405.

Wittkower, D. E. (2016). Principles of anti-discriminatory design. *2016 IEEE International Symposium on Ethics in Engineering, Science and Technology (ETHICS)*. 13–14 May 2016, Vancouver, British Columbia.

Wolf, M. J., Miller, K. W., & Grodzinsky, F. S. (2017). Why we should have seen that coming: comments on Microsoft’s Tay “experiment,” and wider implica- tions. *The Orbit Journal*, *1*(2), 1–12.

Woolgar, S. (2015). It could be otherwise: provocation, irony, and limits. Center for Science, Technology, Medicine & Society at the University of California, Berkeley.

Woolgar, S., & Cooper, G. (1999). Do artefacts have ambivalence: Moses’ bridges, winner’s bridges and other urban legends in S&TS. *Social studies of science*, *29*(3), 433–449.

Yuan, S., Shao, Z., Wei, X., Tang, J., Hall, W., Wang, Y., Wang, Y., & Wang, Y. (2020). Science behind AI: The evolution of trend, mobility, and collabora- tion. *Scientometrics*, *124*(2), 993–1013.

Zhang, L., & Guo, H. (2019). Enabling knowledge diversity to benefit cross- functional project teams: Joint roles of knowledge leadership and transactive memory system. *Information & Management*, *56*(8), 103156.

Zheng, N.-n., Liu, Z.-y., Ren, P.-j., Ma,Y.-q., Chen, S.-t.,Yu, S.-y., Xue, J.-r., Chen, B.-d., & Wang, F.-y. (2017). Hybrid-augmented intelligence: collaboration and cognition. *Frontiers of Information Technology & Electronic Engineering*, *18*(2), 153–179.

2

**هوش مصنوعی جنسیتی**

عملکرد ، انتظارات و جنسیت گرایی

جنسیت یکی از شایع ترین دسته بندی های اجتماعی است که زندگی فرد را شکل می دهد. از سونوگرافی، قبل از تولد، تا جشن برگزاری برای فاش کردن جنسیت بازمانده نوزاد، که فقط یک دوگانه "صورتی برای دختران، آبی برای پسران" است و در صنعت اسباب بازی به صورت مداوم تکرار می شود، همه این عوامل به کودکان آموزش می دهند که جامعه انتظار دارد دختران و پسران در رفتار خود کارهایی خاصی را دنبال کنند، با کمترین انعطاف برای کسانی که پیام تک دوانه جنسیتی را نمی پذیرند. سیاستگذاران به جنسیت پرداخته اند. به عنوان مثال برنامه های مالیاتی به عنوان یک روش ازدواج زندگی ترجیحی برای اعضای هیئت مدیره بازیابی می شود یا در نروژ، نفوذ حداقل به میزان یک درصد زنان در هیئت مدیره شرکت مورد نیاز است. یکی از شایع ترین پارامترهای استفاده شده برای دسته بندی انسان به عنوان متغیر در برنامه های کامپیوتری، جنسیت است. شاید یاد داشته باشید که در بازی های قدیمی پوکمون، اولین تصمیمی که باید بگیرید، نه این است که کدام پوکمون را بگیرید، بلکه باید به سوال دودویی "شما پسر هستید؟ یا دختر؟" پاسخ دهید که توسعه دهندگان بازی برای تعیین ظاهر نمایان شده در نمای پیش نمایش استفاده می کنند. این انتخاب تنها بخش های بصری بازی را تحت تأثیر قرار می دهد مانند مواردی که لباس های مختلف در دسترس هستند، و برخی از گفتگوهای جنسیتی. برای انسان ها در جهان واقعی، استفاده از جنسیت به عنوان پارامتر برای هوش مصنوعی تأثیرات گسترده تری دارد. در این فصل، من چندین دیدگاه برای درک جنسیت، نحوه درک هوش مصنوعی از جنسیت، تأثیر هوش مصنوعی نژادپرستانه، نحوه بازی های صدایی و بدنی هوش مصنوعی با جنسیت، و آنچه که موجبات تنوع را نمایان می کند، را شرح می دهم.

**نکات دیدنی برای درک جنسیت**

در این کتاب، من جنسیت را به عنوان یک سازوکار اجتماعی مورد بررسی قرار می‌دهم که به معنای این است که "جنسیت" یک مفهوم مذاکره شدنی و متغیر است که در فرهنگ‌ها و دوره‌های زمانی مختلف تغییر می‌کند. به عنوان مثال، قوانینی در بسیاری از کشورهای اروپایی برای سر کوتاه در مردان و مو بلند در زنان وجود داشته است، یعنی استانداردهای دودویی. با این حال، وایکینگ‌ها موی بلند را بسیار مردانه می‌دانستند و در چین و کره باستان، پس از بالغ شدن، کوتاه کردن موی خود تنبیه می‌شد. برخی دوره‌ها، کوتاه کردن مو به عنوان غیرانسانی، ضداجتماعی و نشانه‌ای از غم بزرگ تلقی می‌شد - اما این تغییر کرد و پس از زمان، معیارهای اجتماعی نیز تغییر یافت. جامعه و به همین دلیل جنسیت، یک مفهوم ثابت و پایدار نیست که هرگز در زمان و فضا تغییر نمی‌کند. سیمون دوبووار (1949) شهیرا گفته است "زن به دنیا نمی‌آید، بلکه زن می‌شود" که نشان‌دهنده ساختار اجتماعی دسته "زن" است، و اینکه چه چیز برای بودن یک زن از نظر زمان و فضا متغیر است- یک زن شهری در توکیو در سال 2022 انتظارات و ساختارهای اجتماعی متفاوتی نسبت به یک زن از مولداوی روستایی در دهه 1950 دارد. نظریه‌پردازان جنسیت مانند کریستوا (1986) برای فهم نشانه‌ای و نمادین از جنسیت، در حالی که باتلر (2002)مدعی است که جنسیت چیزی نیست که فردی باشد، بلکه چیزی است که فرد به عنوان عملکرد ارائه می‌دهد. از طریق رشد در جامعه، ما یاد می‌گیریم که زنان "باید" به روش‌های خاصی عمل کنند تا زنانگی را تقویت کنند، اما دانشمندان جنسیتی مانند باتلر بر این باورند که این تئاتر جنسیتی نیز به عنوان فرایندی اجتماعی شناخته می‌شود که در زمان‌ها، فضاها و فرهنگ‌های خاص قرار دارد. به این ترتیب، عمل کردن و شبیه‌سازی رفتارهای خاص یک جنسیت، رفتاری که بر اساس فضاها، زمان‌ها و فرهنگ‌های مشخص قرار دارد، به عنوان عملکرد جنسیتی شناخته می‌شود. جنسیت یک مفهوم پیچیده و داغ است که در جامعه و در میان برخی از دانشمندان جنسیت، فمینیست، شوو و دانشمندان دیگر بحث شدیدی می‌شود. به آورده شدن براساس تمام جنبه‌های فهم جنسیت، خارج از حوزه این کتاب است.

مطالعات جنسیت یک زمینه ی پیچیده و گسترده با چندین زیرشاخه مانند مطالعات مردان (کیمل، 1986) و نقد دهکولونیالی (اسپیواک، 2003) است که بسیاری از آنها در قابلیت انطباق با موضوع کتاب مفید هستند. اما من در این کتاب به دیدگاه‌های نظری جنسیتی که به صورت مستقیم با تکنولوژی سرو کار دارند، تمرکز داده‌ام. به طور خاص، من از کتاب پرز (2019) با عنوان "زنان نامرئی: به نمایش گذاشتن تبعیض داده‌ای در یک جهان برای مردان طراحی شده" به عنوان منبعی برای توضیح اینکه چگونه تاکنون اطلاعات جنسیتی در تکنولوژی نادیده گرفته شده است استفاده کرده‌ام. نصف زندگی بشریت در طول تاریخ کاملاً نادیده گرفته شده‌اند - همانطور که پرز این را یک حضور غایب توصیف کرده‌است مردان به عنوان معیار و نوعی برای توضیح تجربیات انسان درنظر گرفته شده‌اند. پرز (2019) در کتابش چندین مثال آورده و سالها تحقیقات خود را درباره تبعیض تکنولوژی در برابر زنان خلاصه شده است که شامل مواردی مانند دمای محیط کاری که بر اساس دمای بدن مردان تنظیم می‌شود، قفسه‌های موجود که برای ارتفاع استاندارد مردان طراحی شده، تصادفات خودرویی که عواقب خطرناکی برای بدن زنان دارد (مثال آورده شده در Augenstein و همکاران، 2005؛ Carter و همکاران، 2014؛ Hitosugi و همکاران، 2018) و بحران‌های بهداشتی مانند حملات قلبی که برای زنان به عنوان پیشنهادهایی برای استاندارد مردان تدوین شده است می‌شود.

در رابطه بین تکنولوژی و جنسیت، دانشمندان فمینیست تکنوساینس می‌توانند ابزارهای تجزیه و تحلیل خوبی را ارائه دهند، به عنوان مثال باراد (2003) که برای درک پست-انسان‌گرایی از جنسیت و پیچیدگی پیوندهای داخلی بین انسان‌ها و بدون جان، و هاراوی (1985) که با نظریه سایبورگ (Cyborg) خود بر این باور است که بیولوژیکی و فنی به هم تنیده و جدا‌ناپذیر هستند و جنسیت نیز توسط و با تکنولوژی شکل گرفته است. در حالی که جنسیت و بخش‌بندی جنسی با به صفر رسیدن چگونگی درک امر در زمینه مطالعات جنسیت بوجود می‌آید، من به عنوان سازنده کتاب، به فرض بر مبنای خلقت اجتماعی جنسیت به جای خلقت زیستی تمرکز خواهم کرد. نقطه نظر اصلی من در اینجا این است که با بررسی نحوه تعامل و تأثیر تکنولوژی و جنسیت بر یکدیگر، نگاه اجتماعی جنسیت را بررسی کنم و روابط بین آنها را بررسی کنم. در فصل بعدی این کتاب به بررسی بیشتری در مورد عدم تعادل در الگوریتم های هویت جنسی می‌پردازیم.

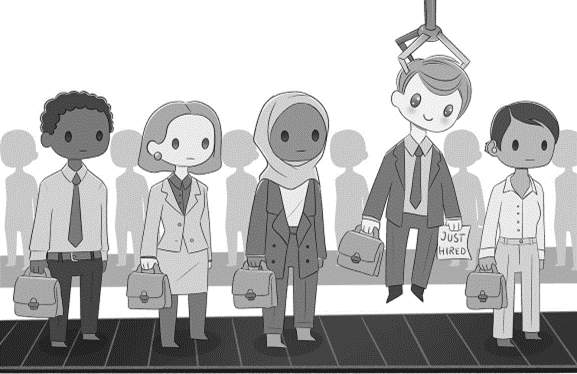
**درک هوش مصنوعی از جنسیت**

هوش مصنوعی به همان شکل که انسان‌ها فکر نمی‌کنند، فکر نمی‌کند، بلکه با الگوریتم‌هایی برای آموزش مدل ها برای انجام کارهای مختلف برنامه‌ریزی شده است. اگرچه برخی از فرآیندهای پیشرفته‌تر مانند یادگیری عمیق برای انسان‌ها سخت هستند، با این حال با یک مثال ساده از چگونگی پدیده تبعیض قابل توجه است که یعنی وکتورهای کلمه. وکتورهای کلمات از کلماتی که به کلمات دیگر ارتباط دارند، ساخته می‌شوند، به عنوان مثال کلمه «چنگال» نزدیک‌تر به «چاقو» است تا به «سگ». برخی از مطالعاتی که به تبعیضاتی که توسط ماشین‌ها تولید می‌شوند، می‌پردازند نشان می‌دهد که "فرض بر آن است که پرستار یک زن و مهندس یک مرد است" (کورپیکس-بریکی و لئونی، 2021).

یایید یک مثال دیگر از یک الگوریتم برداریم که داده ها را اشتباه پردازش کرده و سپس انسانهایی را که باید نتایج را تفسیر کنند، مبهوت می کند. از سال 2014 تا 2018، متخصصان یادگیری ماشین در آمازون در حال کار بر روی یک موتور استخدامی بودند که باید بر اساس رزومه و برنامه کاری کاندیدها، بهترین کاربران جدید را انتخاب کند. آمازون، به عنوان یکی از سه شرکت بزرگترین جهان، یکی از رهبران صنعت های هوش مصنوعی بود. شرکت های FAAMNG (فیس بوک، اپل، آمازون، نتفلیکس، مایکروسافت و گوگل) مسلط به صنعت فناوری و توسعه جهانی هستند. با مقایسه دارایی خالص شرکت ها با کشورها، "9 شرکت در بین 30 کشور بیشترین دارایی خالص را دارند" و شرکت هشتم پرسود جهان، آیفون، دارای دارایی خالصی است که بیشتر از دارایی خالص روسیه، ایتالیا و برزیل به همراه دارایی هایشان است. (Lishchuk, 2021).

بنابراین، آنچه که شرکت های FAAMNG در زمینه هوش مصنوعی انجام می‌دهند (یا شکست می‌خورند)، می‌تواند تأثیر گسترده‌ای داشته باشد. به عنوان مثال، آمازون بیش از 1.3 میلیون نفر در سراسر جهان استخدام کرده است و انتخاب شخصیت های مناسب برای شرکت، به عنوان الهام بخش برای شرکت های دیگر نیز مهم است. به همان اندازه که هر مدیر منابع انسانی می‌تواند به شما بگوید، استخدام فرد مناسب و کیفیت دار بسیار زمان و انرژی زیادی می‌طلبد، فقط جستجو در میان کاندیدهای نیازمند، منابع انسانی بسیاری را می‌گیرد. برای دهه‌ها، بسیاری از شرکت‌ها به دنبال بهینه سازی فرآیند استخدام بوده‌اند و چندین سامانه برای این منظور ایجاد شده است. مثال‌های متعددی مانند آزمون‌های شخصیتی، مراحل کیفیت و ارزیابی‌های روانشناختی، در محیط شرکت‌های فناوری، به شکل های متنوع و فراگیرتری پیدا می‌کنند. استخدام فرد نادرست برای یک شغل ممکن است خسارت زیادی به همراه داشته باشد؛ بنابراین، استخدام شخصی مناسب برای یک شغل، در لیست اولویت های کارفرمایان قرار می‌گیرد.

وقتی موتور جستجوی استخدامی آمازون در سال 2018 تست شد، نتایجی را به دنبال داشت که آنالیزورها را گیج کرد (داستین، 2018). به عنوان یک محصول از یادگیری ماشین، ابزار استخدامی آمازون تصمیم گرفت که برخی از کلمات کلیدی را که به نظر آنها منفی هستند و باید در متقاضیان، ارزیابی منفی شوند. این بیشتر بر اساس ارزیابی کارکنانی بود که تاکنون در شرکت کار کرده بودند و سعی در یافتن کارکنانی برای جایگزین کردن آنها بود. با این حال، مشکل اینجا بود که در گذشته آمازون (مانند بیشتر شرکت های تکنولوژی) بیشتر مردان را برای کارهای تکنولوژیکی استخدام کرده است. بنابراین، الگوریتم "زن" را به عنوان یک کلمه کلیدی منفی در نظر گرفت و برنامه های کاربردی را که به آن اشاره داشتند، ناارزیابی کرد، به عنوان مثال، متقاضیانی که در باشگاه شطرنج زنان بازی کرده یا از یک دانشگاه زنانه تحصیل کرده بودند.

این سیستم همچنین به کلمات مردانه مانند "captured" و "executed" علاقه مند بود که بیشتر در رزومه های مهندسان مرد پیدا می شد. آمازون بعداً این ایده را کنار گذاشت و اکنون از نسخه بسیار ساده تری استفاده می کند که برنامه های کاربردی تکراری را حذف می کند. در این مثال، آنچه به سیستم داده شد (10 سال از متقاضیان و بیشتر آنها مرد بودند) بخشی از تعصب بود. با این حال، چگونه آن پردازش شد و چه انتخاب هایی سیستم انجام داد، می تواند در صورت عدم نظارت و پایش دقیق، برای انحصار تعصبی ایجاد شود. سیستم اشکالی نداشت؛ فقط کاری را انجام داد که برنامه ریزی شده بود - بیشتر به همان شیوه‌ای که در گذشته استخدام شده بود - با تثبیت برخی از معیارها و شیوه های استخدام به توصیه ها و تصمیمات مبتنی بر داده. شکل 2.1 نشان می دهد که کسانی که شغل را دریافت می کنند به صورت تصادفی نخواهند بود. نقض وضع موجود نیاز به مداخله و خواست به تغییر دارد، در غیر این صورت، یک شخص با آنچه در گذشته انجام شده، گیر می کند.

چندین مثال از ابزارهای هوش مصنوعی وجود دارد که برای کاهش تعصب در فرایند استخدام، در استخدام به کار گرفته می شوند، همانطور که Ramboll (2022) توضیح داده است. یکی از این مثال‌ها Textio است. این یک ویرایشگر متن است که عبارات جنسیتی را در توضیحات شغلی برجسته می کند و به کارفرمایان این امکان را می دهد تا آگهی های استخدامی خود را بیشتر شامل عمومیت بخشید. مثال دیگری Pymetrics است، یک پلتفرم استخدام و مطابقت شغلی با هوش مصنوعی که از آزمون های شخصیتی و یا دیگر داده های "منطقی" برای بررسی مجموعه ای از خصوصیات افراد استفاده می کند تا به کارفرمایان این امکان را بدهد تا با رزومه هایی که متعارف هستند محدود نشوند و بدین ترتیب نیروی کار متنوعی را استخدام کنند. مثال های دیگر از جمله unbias.io، این پلتفرم نام و صورت خود را از پروفایل های LinkedIn حذف می کند، زیرا این می تواند باعث تعصب ناخودآگاه در فرایند استخدام شود، Entelo، این پلتفرم شناسایی جنسیت و نژاد را برای اینده انتخابی حذف می کند و Talvista که خصوصیات و شناسه های شخصی را از متقاضیان مخفی می کند و توصیف شغلی بدون جنسیت می‌نویسد.

**یا هوش های مصنوعی تعصب جنسیتی دارند؟**

بیشتر از هوش‌های مصنوعی، صحبتی را خلق نمی‌کنند، بلکه از پیام‌های پیش‌فرضی که برای پاسخ به شرایط خاصی طراحی شده‌اند، استفاده می‌کنند. در یک گزارش از سازمان ملل متحد از سال 2019، عنوان نشان می‌دهد که چگونه سیری، که یکی از دستیار صوتی (voice assistant) هاست ، به شخصی که او را به عنوان یک زن فحش می‌دهد، پاسخ می‌دهد: "اگر می‌توانستم خجالت می‌کشیدم" (وست و همکاران، 2019). علاوه بر این، گزارش تحقیقاتی را در مورد نحوه رفتار با کمک‌های گفتاری مختلف مانند Cortana شرکت مایکروسافت، Alexa شرکت آمازون و Siri شرکت اپل به ارمغان می‌آورد. این کمک‌های به گفتار (voice assistants) ، کاربران را که به آنها گفته‌اند که تجاوز شده‌اند و نیاز به مرکز بحران دارند را به سرعت نادیده می‌گیرند یا به شوخی به آنها جواب می‌دهند، به ویژه به شکلی توهین‌آمیز و جنسی. اما Cortana و Alexa و Siri (همه به شکلی به عنوان دستیار صوتی زن تصویر سازی می شوند) قابلیت پاسخ به این سوالات را تنها به صورت جداگانه فراهم نمی‌کنند، بلکه به صورت اتوماتیک با یک صف پیش‌فرضی پاسخ می‌دهند که توسط طراحان کد نویسی شده است. اگر کد پاسخ به گزارش تجاوز جنسی و رفتار سمی می‌تواند توسط افرادی که فکر می‌کنند که این مسئله ممکن است یک موضوع شوخی باشد نوشته شود، به شدت مشکل‌ساز است. به طور مشابه، طراحی ربات‌ها به عنوان زن نیز مشکل‌ساز است. کاتلین ریچاردسون، نویسنده کتاب "آنتروپولوژی ربات و هوش مصنوعی" گفت در یک مصاحبه (لوئیس، 2015) که این: "احتمالاً نشان دهنده دیدگاه برخی از مردان در مورد زنان است، یعنی آنکه آنها به عنوان انسان به کاملی شناخته نمی‌شوند [...] چیزی که برای آنها ضروری است ، قابل تکثیر است، اما در مورد ربات‌های پیشرفته‌تر ، باید مذکر باشند ".

در بسیاری از کشورها، دستیاران صوتی مانند Siri به صورت بلندگوی زنانه یا مردانه در دسترس هستند، اما در ایالات متحده، تنها از سال 2013، تنظیم مردانه برای Siri در دسترس قرار گرفت. اما چه چیزی باعث می‌شود این صداها "زنانه" باشند؟ در کشورهایی مانند ژاپن، زبان جنسیتی تأثیر عمیقی بر دستور زبان و کلمات استفاده شده دارد، به عنوان مثال، کلمه “من” به “آتاشی” برای زنان و “اوره/بوکو” برای مردان ترجمه می‌شود. در زبان‌هایی کمتر تحت تأثیر جنسیتی مانند انگلیسی، زبان جنسیتی به اصلی ترین ویژگی آواز که جنبه آن، البته یک ساده‌سازی است، مربوط است. میانگین تردید زنان در بردارهای فرکانس ۱۶۵ تا ۲۵۵ هرتز است در حالی که میانگین مردان بین ۸۵ تا ۱۵۵ هرتز است. این یک ساده‌سازی گسترده‌ترین مبتنی بر نمونه‌های ایالات متحده است. زنان در ژاپن، به عنوان مثال، از یک بردایت با فرکانس بسیار بالاتر استفاده می‌کنند، که در بخشی از آن وابسته به جامعه است.

معمولا دستیاران صوتی به میانگین متمایل هستند، با صداهای “زنانه” کمی برای صدای متوسط پایین‌تر و صدای “مردانه” کمی در بالای میانگین. با این حال، مثالی وجود دارد که این دو هنجار را نقض می‌کند، مانند Q، اولین دستیار صوتی “بدون جنسیت” در جهان، به گفته توسعه دهندگان. Q یک نرم افزار است که در ارتباط با یک جشنواره افتخار دانمارکی توسعه یافته و یک آزمایش است که صداهای یک مجموعه از افراد غیردودویی برای تهیه صدای Q تحلیل شده است. صدای Q طبق نمونه‌ها با فرکانس ۱۵۳ هرتز است و مناسب‌ترین برایش بود. این برای توسعه‌دهندگان نماینده "یک آینده است که دیگر با دو قطبیت جنسیت تعریف نمی‌شویم، بلکه با توجه به تعریف خودمان از جنسیت و همانطور که آنها را زندگی و تجربه می‌کنیم" (GenderLess Voice، 2019)، نشان دهنده یک آینده بدون جنسیت است. Q به ندرت در زمینه آزمایشی استفاده شده‌است، اما باز هم فرصتی برای بازآرایی بحرانی در مورد چگونگی و چرا دستیاران صوتی در ابتدا به جنسیت تعلق دارد و چه پیامدهایی دارد، فراهم می‌کند. برای دانشمندان صوتی، جالب است که بررسی کنند که آیا این فرکانس (نزدیک به دیگر سمت عمق بیش از آنکه مربوط به فرکانس بالاتر باشد) موجب بی‌جنسیت شدن، بی‌طرفی جنسیتی یا نوعی از بدون جنسیت شدن فناوری می‌شود، زیرا برخی از فرضیات جنسیتی جامعه را تکثیر می‌کند. در عین حال، در مورد تکنولوژی‌های دستیار صوتی دیگر پویایی‌های قدرت وجود دارد. لهجه‌ها و زبان‌های کوچکتر به طور معمول دشوارتر برای دستیار صوتی می‌باشند و بنابراین به طور مشکل‌سازی برای افرادی که زبانی غیراستاندارد را صحبت می کنند تبدیل می‌شود. برای استفاده در بازآرایی بحرانی بحث‌های زیادی در مورد زبانهای تبعیض‌آمیز وجود دارد که پیامدهای بزرگی برای بازتولید هوش مصنوعی دارد. برای توضیحات بیشتر در مورد چگونگی جنسیت بندی دستیاران صوتی به کتاب The Smart Wife نوشته Strengers و Kennedy می توانید مراجعه کنید.

توسعه دهندگان می‌توانند در مورد اینکه چگونه ارزش‌های اجتماعی را بازتولید می‌کنند، تفکر بیشتری داشته باشند. تبعیض جنسیتی توسط اندکی از دانشمندان و سازمان‌ها به عنوان رایج در جوامع معاصر و تاریخی ثابت شده است. تبعیض، تصویرسازی و تبعیض نسبت به زنان، فرصت‌ها و تجربیات زنان در سراسر جهان را محدود می‌کند. تبعیض جنسیتی در سراسر جهان شکل‌های متعددی را به خود می‌گیرد، از جمله رفتارهای تهدید کننده، استثنایی، نقش‌های سختگیرانه، جنسی، رسوا کردن، صدازدن، زبان توهین‌آمیز، توهین، کنترل، آزار، شکار، خشونت جسمانی و روانی، تجاوز، قتل زنان و نفی زنان. واژه روانی به چیزی اطلاق می شود که باعث محدود کردن خودزندگی، عدم توانمندی به جامعه، درد و رنج باشد(human-rights-channel.coe ، بی تاریخ). برخی تلاش‌ها برای جلوگیری از این امر از طریق استفاده از هوش مصنوعی عبارتند از پایان دادن به خشونت مردان به سمت زنان با تشخیص رفتار خشن به کمک ویدیو ، برای مثال با استفاده از رویکردهای ماشین بردار پشتیبان برای تشخیص رفتار سوءاستفاده از نظارت تصویری برای مثال نمونه پرسشی (Ramboll، 2022، ص. 17).

اما تکنولوژی‌های دیجیتال همچنین می‌تواند مخاطراتی برای بازتولید چنین تبعیضی به دنبال داشته باشد و ما باید بر روی ارتباط بین ارزش‌های اجتماعی و فنی با دقت تأمل کنیم. این مسئله ممکن است بر اساس سطوحی باشد که در نگاه اول به نظر خیلی جدی نیامده است. یکی از این بازتولیدها، نحوه و دلیلی است که چرا دستیاران صوتی دیجیتال به صورت زنانه نشان داده می‌شود. برای مثال، Cortana شرکت مایکروسافت بر پایه شخصيت قرار داده شده در سری بازی Halo به نام Cortana است، که او در آن با لباس شنای بسیار خلاصه با رنگ آبی بدون لباس شخصیت شده و جنسیتی شده است. مایکروسافت سعی کرده است که دستیار صوتی Cortana را از شخصیت بازی جنسی که بر پایه او طراحی شده است، دور کند. مثال دیگری از تکنولوژی که تبعیض جنسی را بازتولید می‌کند، استفاده از برنامه "TitStare" برای مردانی که به سینه زنان نگاه می‌کنند است (Broussard ، 2018، ص ۱۶۷؛ Ohlheiser، 2013). Broussard (2018، ص ۸۵) این مسئله را به صورت زیر شرح می‌دهد:

ما یک گروه کوچک و نخبه مردان داریم که عادت دارند از قابلیت های ریاضی خود بیش از حد تعریف کنند و به صورت سیستماتیک برای قرن ها زنان و انسان های سیاه پوست را به سود ماشین ها، کنار بگذارند. آنها عادت دارند ایده های علمی تخیلی را واقعی کنند، به معماری اجتماعی بی احترامانه روی بیاورند، باور نمی کنند که مردم باید به قوانین و نرم افزار های اجتماعی عمل کنند و هزینه هایی از دولتی ها که تلف شده اند، را دارند. به علاوه، آنها آرایگان های ایدئولوژیک پرچمدار دلالان آنارشیستی لیبرتری و کاپیتالیست هستند. با همه اینها، باید چه مشکلی پیش بیاید؟

وقتی که کاربر تصور شود مشابه گروه‌های همگن توسعه‌دهندگان است و روش «من» در طراحی استفاده می‌شود، بسیاری از مشکلات می‌تواند ایجاد شود (اودشورن و همکاران، ۲۰۰۴). به عنوان مثال، اگر پسران جوان سفید و استریت خود را به عنوان کاربران شان تصور کنند، کاربرانی که با آن‌ها هم‌خانه نیستند ممکن است از هر چیزی توسعه‌یافته، بسیار بیرون شوند. به عنوان مثال، "آقای پیش‌فرض" (ایریگاری، ۱۹۹۲؛ پرز، ۲۰۱۹)، که تا به حال در بسیاری از موارد منجر به دشمنی زنان شده است. قابل توجه است که چگونگی ایجاد مسیرهای جدیدی که می‌توانند عملیات و زیرساخت‌های جنسیتی را کاهش دهند، باید به طور بحرانی بررسی شود، بجای تقویت آن‌ها توسط فناوری‌های دیجیتالی مانند هوش مصنوعی (AI). به عنوان مثال، سازمان کار اتریش، از یک الگوریتم استفاده کرد که زنان و افراد معلول را نسبت به مردان و افراد دارای سلامت بدنی بالاتر مقدار داد، افراد بالای ۳۰ سال را کم ارزش قلمداد کرد، و مادران را از پدران با امتیاز کمتری رنک کرد.(dig.watch، ۲۰۱۹). در اینجا، بودن یک زن با پارامترهای شخصی دیگری همانند نگهداشتن دوستان مرد، همراه شده است. در بسیاری از کشورها غیرقانونی است که از درخواست کردن اطلاعات در مورد بارداری یا فرزند داشتن از متقاضیان کار خودداری شود، چرا که این باعث ارزش خوانده شدن آنها توسط کارفرمایان می‌شود، که همانطوری که در اتریش مشاهده می‌شود، توسط AI تولید می‌شود. در یک مطالعه در این زمینه، ال‌هوتر و همکاران (۲۰۲۰) توضیح داده‌اند که "طراحی الگوریتم تحت تأثیر امكانات فنی و همچنین ارزشها، نرم‌افزارها و اهداف اجتماعی قرار می‌گیرد"، با توصیه از حق شفافیت، حق اعتراض و نیاز به مشارکت عمومی (ماگر و ال‌هوتر، ۲۰۲۱). از آنجا که داده‌ها هرگز بی‌طرف نیستند، ارزیابی ارزش‌ها و معیارهای موجود در تکنولوژی می‌تواند مهم باشد. یک مطالعه، که از AI برای شناسایی تبعیض براساس جنسیت در تنظیمات شفاهی استفاده می‌کند، شناسایی کننده جنسیت منبع‌باز با استفاده از رادیو فرانسه به عنوان یک مورد مطالعه برای افزایش آگاهی از میزان زمان صحبت زنان و مردان (دوکان و همکاران، ۲۰۱۸)، و یک جایگزینی سوئدی برای ارزیابی نتیجه‌گیری از اقدامات تنوع و انعطاف پذیری در محل کار به نام Ceretai معرفی شده است.

**ربات ها و جنسیت**

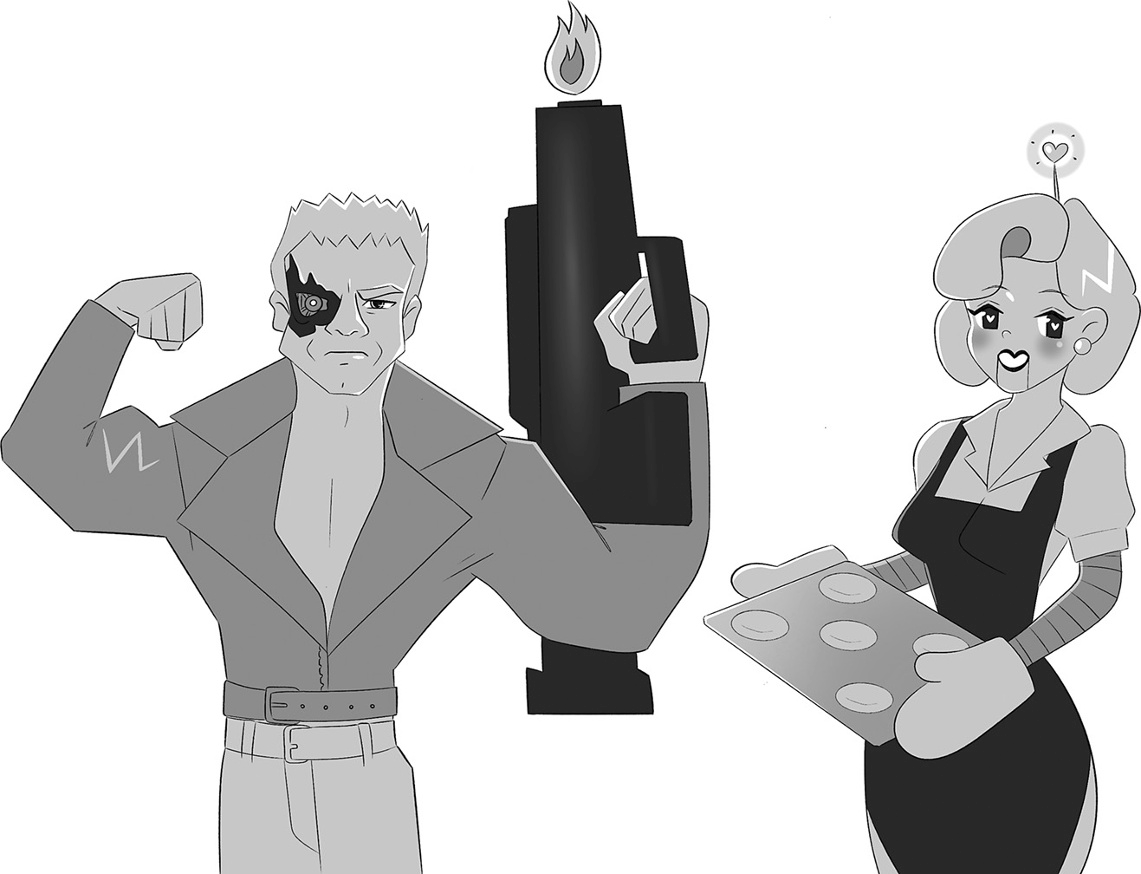
تا به حال، این کتاب به طور اساسی نرم‌افزارها و برنامه‌هایی را بررسی کرده است که حضور فیزیکی کمی دارند؛ به عنوان مثال، در یک بلندگو هوشمند. اما بیایید به جعبه‌های فیزیکی توجه کنیم که AI در آن حضور دارد، که به آنها "سخت‌افزار" می‌گویند. ربات‌ها به طور فزاینده‌ای به بخشی از زندگی افراد تبدیل می‌شوند، از جمله ربات‌های خدماتی مانند ربات‌های جاروبرقی و باغ‌سرایی و ربات‌های اجتماعی مانند ربات‌های پپر و صوفیا که وارد تحقیقات، صنعت و متروکه‌های فناورانه عمومی شده‌اند. ربات‌ها سیستم‌های AI داخلی دارند که با قطعات فیزیکی آنها تعامل دارند. رجبات، با پیچیدگی‌های مختلف AI، به شکلی جنسیت‌دار به تصویر کشیده می‌شوند. این به خصوص برای ربات‌هایی که به شکل زنانه تصویر می‌کنند، مانند ربات صوفیا است که بر اساس اودری هپبورن، ملکه نفرتیتی مصر باستان و همسر خالقش، مدل سازی شده است.

در مقاله "جنسیت مکانیکی: چگونگی جنسیت‌بندی ربات‌ها توسط انسان‌ها؟"، من در مورد این صحبت می‌کنم که چقدر بیشتر ربات‌ها شبیه به انسان می‌شوند، ما به آنها جنسیت خواهیم داد. ربات‌هایی که به شکل انسان طراحی شده‌اند، خصوصاً در زمینه‌ی جنسیتی جالب هستند. یک ربات "جنسیت ندارد" به طوری که ما انسان‌ها این اصطلاح را به عنوان یک صفت یا اسم می‌فهمیم، اما ما آن را به عنوان یک فعل جنسیتی می‌طلبیم؛ یک قسمت از آن از طریق طراحی و دیگر قسمتش از طریق شاخص‌های اجتماعی و عواملی که در آن استفاده می‌شود، به وجود می‌آید. زمانی که طراحان ربات Sofia تصمیم به مدل کردن آن بر اساس آدری هپبورن و ایجاد تصوری رژ لب و مژ‌های بلند ربات دادند، یک اجرای موثری از زنانگی ایجاد شد. هنگامی که خالقان Pepper (حداقل در حال تمام‌شدن شبکه‌ی رسانه‌ای آنها) از نداشتن جنسیت برای آن ربات در وب سایت خود توصیف می‌کنند، نیز این مسئله مهم است.

پپر یک ربات است که جنسیت ندارد.

چرا؟ پپر نه نر و نه ماده است، امّا هنگامی که با پپر آشنا می‌شوید، شاید خودتان را در حالی ببینید که به پپر با جنسیتی که برای شما منطقی‌تر است اشاره می‌کنید.

چگونه؟ پپر باید به عنوان یک ربات ملاحظه شود، یک دوستانه‌ترین آنها. پپر یک انسان‌نما است، امّا برخلاف انسان جنسیتی ندارد و به عنوان یک ربات شناخته می‌شود.

با این حال، ربات‌ها به مرور همانند انسان‌ها جنسیتی اجتماعی به آنها تعلق می‌گیرد چقدر که شبیه به انسان می‌شوند. ما انسان‌ها نیاز قوی به طبقه‌بندی و قرار دادن اشیا در جهان اجتماعی خود داریم. زبان ما به جنسیت بندی بستگی دارد (بعضی زبان‌ها به شدت بیشتر از دیگران). در گویش‌های جنسیتی نروژی "han Pepper" یا "hun Pepper" استفاده می‌شود؟ در زبان انگلیسی هنگام توصیف پپر، به آن به عنوان "it"، "he" و یا "she" اشاره می‌شود؟ آیا پپر در زبان ژاپنی از "boku" مردانه، "atashi" زنانه یا "watashi" که بیشتر نتیجه نگرایی ندارد، برای اشاره به خود استفاده می‌کند؟ این انتخابات زبانی که توسط سخنگو/نویسنده انجام می‌شود، قسمتی از جنسیت‌بندی ربات‌هاست و نشان می‌دهد چگونه طبقه‌بندی‌های زبانی ارزش‌های اجتماعی را نمایش می‌دهند و تکثیر می‌کنند، ارزش‌هایی که به داده و مواد نیز ترجمه می‌شوند (شکل ۲.۲).

شکل ۲.۲

همچنین می‌توان به این دید که لباس ربات‌ها به طرز قابل مشاهده‌ای به جنسیت بندی شده‌اند اشاره کرد. وقتی پپر را با لباس دوست داشتنی گریبان قرمز ("رد رایدینگ هود") پوشانده‌اند، در واقع یک داستان (با جنسیت) روایت شده است و داستان‌های تکنولوژیکی نیز مهم هستند. من در اینجا به دیدگاهی که می‌گوید جنسیت‌بندی ربات‌ها خوب یا بد است، بیشتر توجه نمی‌کنم (مانند پرسش‌های فلسفی دوگانه، پاسخ معمولاً پیچیده‌تر است)، اما نکته من این است که ما باید آگاه باشیم که جنسیت‌بندی فناوری یک زمینه است که برای مدت طولانی مورد مطالعه قرار گرفته است و بر روی بسیاری از فناوری‌ها، از جمله فرمان‌دهنده مایکروویو (Cockburn & Ormrod، 1993)، اصناف و دوچرخه‌های الکتریکی (Oudshoorn et al.، 2002) و تلفن‌ها (Frissen، 2018) تمرکز داشته است. ربات‌ها و هوش مصنوعی نباید استثناءی از این مطالعات باشند یا باید نباشند. با این حال، من باید بگویم که مبتکران رویکرد فمینیستی-تکنولوژیکی در بازخوانی ربات‌ها، همانند انسان‌ها، شبیه به انسان‌های واقعی است. چگونه باید ربات‌ها را درک و بحث کردن و حاکمیت بر آنها، همچنان موضوع بحث بسیاری از محققان است. به عنوان مثال، محققانی مانند برایسون (2010، ص 1) از آن جا که این موضوع ممکن است منجر به بی‌انسان‌کردن افراد شود، از نیروی بخشیدن انسانی به ربات‌ها هشدار داده‌اند: «ربات‌ها نباید به عنوان افراد توصیف شوند و نه مسئولیت قانونی و نه اخلاقی کارهایشان بر عهده آنها قرار گیرد. ربات‌ها متعلق به ما هستند و ما اهداف و رفتارشان را مشخص می‌کنیم، بطور مستقیم یا غیرمستقیم از طریق مشخص کردن هوششان.»

**هوش مصنوعی و سکس**

تکنولوژی همیشه یک ارتباط قوی با صنعت جنسی داشته است و هوش مصنوعی هم از این قاعده مستثنی نیست. از روبات های جنسی واقع گرایانه با گزینه های محدود صحبت تا ابزارهای جنسی سفارشی شده که از کاربرانشان یاد می گیرند، با نام هایی مثل "آتوبلو"، "هام"، "آرایش لب هوشمند" و "تخیلات واقعیت مجازی" ، هوش مصنوعی به عنوان یک بازار تجاری بزرگ به صنعت جنسی نگاه می کند. یکی از مسائل نگران کننده تر، این است که چگونه برخی ویدیوهای پورنو با فناوری "Deep-fake" ساخته می شوند، که در آن چهره های سلبریتی ها به عنوان مثال، بر روی بازیگران پورنو قرار می گیرند تا وهم بدهند (Maddocks، 2020). این از نظر مسائل حریم خصوصی و اخلاقی مشکل دارد و می تواند باعث آسیب شغلی بازیگران شود اگر به عنوان ویدیوهای واقعی شناخته شود. این مساله حتی می تواند برای افراد عادی با دشمنان سابق خود با "پورنو انتقامی" یا "سوء استفاده جنسی بر مبنای تصاویر" به عنوان برخی از دانشمندان اشاره کرده اند، آسیب رسان باشد. (McGlynn & Rackley، 2017).

در مفاهیم مرتبط، روبات های جنسی همچنین در سال های اخیر با انتقادات و علاقه دانشمندان مواجه شده اند. دناهر (2017) سه مسأله اخلاقی اصلی درباره روبات های جنسی مطرح می کند: "(i) آیا به کاربر کمک / آسیب می رسانند؟ (ii) آیا به جامعه کمک / آسیب می رسانند؟ یا (iii) آیا به روبات کمک / آسیب می رسانند؟" در یک مقاله به نام "آیا برای روبات های جنسی آماده ایم؟" Scheutz و Arnold (2016) در مورد اینکه مردان بیشتر به آنها قابل قبول هستند، اما زنان کمتر؛ Appel et al. (2019) روبات های جنسی را به فناوری fandom یا otaku در ژاپن مرتبط می کنند، در حالی که Richardson (2016) توازن ناهمتقارن "روابط" بین فعالیت های رسوا و توسعه جنسی را بررسی می کند و مشاهده می کند که چگونه نشان دادن شخصیت به شکل انسان نمایی نگرش های جنسی مبتنی بر جنسیت را نشان می دهد. یک معضل به لحاظ هوش مصنوعی این است که این روبات های جنسی در حال یادگیری از کاربران هستند و بر اساس علاقه ها و ناخوشایندهای آنها پروفایل هایی را ایجاد می کنند. برخی دانشمندان معتقدند که روبات های جنسی می توانند ابزارهای مهمی برای افراد معلول یا افرادی باشند که در غیر این صورت با پیدا کردن شریک جنسی انسان مشکل دارند (Devlin، 2015؛ Jecker، 2021). روبات های جنسی مثالی از تکنولوژی در مورد یک موضوع باریک است، زیرا جنسیت اغلب در طی تاریخ انسان به عنوان یک مسئله شارژ شده شناور بوده است.

**خلاصه فصل**

این فصل معیارهای متنوع "جنسیت" را بررسی کرد که در ابتدا آن را باز خورد داد تا متوجه شود که جنسیت در چه شکل های مفهومی تعریف می شود. هوش مصنوعی از جنسیت به عنوان یک پارامتر برای دسته بندی انسان ها و غیرانسان ها استفاده می کند و با این کار برخی از ارزش ها و معیارهای اجتماعی موجود درباره جنسیت را بازتولید می کند. اما این درک از جنسیت که چیست یا باید چگونه باشد، ثابت و نهایی نبوده و از طریق اجرایی بودن آن ها توسط جامعه ایجاد شده اند. در ادامه، یک مثال بدنام از سیستم استخدام آمازون ارائه شد که برداشت نشان می دهد که در آن، درخواست های شغلی زنان به دلیل عدم تطابق با پروفایل بیشتر افرادی که در حال کار در آمازون هستند، ارزش کمتری نسبت به دیگر درخواست ها داشته اند. به طور پیوسته، به برنامه های جنسیستی با سطح بالای جنسیتی، از جمله Siri ، Alexa و سایر برنامه های خواهران، که در پاسخ به درخواست ها و طلبهای جدی جنسی افراد پاسخ های لازم را ندادند و در هنگام جنبه جنسی ساختن برنامه ها، به این مسئله پرداخت شد. همچنین به معاونت صوتی Q که در حال راه اندازی معاونت صوتی بدون جنسیت است و به سمت سخت افزار فیزیکی نگاه کردیم که در فرایند انسان یابی، روبات ها به شکلی که جنسیت های خود را دارند، هویت پیدا می کنند. در نهایت، به برنامه های جنسیتی در AI پرداخته شد و به مسئله فیلم های پورنوی پویا که با استفاده از تکنولوژی "Deep-fake" ساخته می شوند، اشاره شد. هوش مصنوعی از ساختارهای اجتماعی عمیقی که جنسیت برای جوامع دارد و چگونه جنسیت در دوره ها و فرهنگ های مختلف می تواند به صورت متفاوتی درک شود، دارای یک باریک بینی است. بنابراین ما باید مراقب باشیم که معیارهایی که به خاطر داریم، همانطور که در حال حاضر شایع هستند، به عنوان معیارهایی برای همه جوامع پذیرفته شوند. بنابراین در فصل بعد به بررسی شکل گیری جنسیت به عنوان چیزی غیر دو بخشی، و نگاه به آنچه که هوش مصنوعی می تواند برای جریان باریک جنسیتی انجام دهد پرداخته می شود.

**مرجع ها**

Allhutter, D., Cech, F., Fischer, F., Grill, G., & Mager, A. (2020). Algorithmic profiling of job seekers in Austria: how austerity politics are made effective. *Frontiers in Big Data*, *3*, 5.

Appel, M., Marker, C., & Mara, M. (2019). Otakuism and the appeal of sex robots. *Frontiers in Psychology*, *10*, 569.

Augenstein, J., Perdeck, E., Bahouth, G. T., Digges, K. H., Borchers, N., & Baur,

P. (2005). Injury identification: Priorities for data transmitted. *Proceedings of the 19th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*. 6–9 June, 2005, Washington D.C.

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society*, *28*(3), 801–831.

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.

Bryson, J. J. (2010). Robots should be slaves. InYorick Wilks (Ed.), *Close engagements with artificial companions: Key social, psychological, ethical and design issues* (pp. 63–74). John Benjamins Publishing Company.

Butler, J. (2002). *Gender trouble*. Routledge.

Carter, P. M., Flannagan, C. A., Reed, M. P., Cunningham, R. M., & Rupp, J. D. (2014). Comparing the effects of age, BMI and gender on severe injury (AIS 3+) in motor-vehicle crashes. *Accident Analysis & Prevention*, *72*, 146–160.

Cockburn, C., & Ormrod, S. (1993). *Gender and technology in the making*. SAGE Publications.

Danaher, J. (2017). Should we be thinking about sex robots? In John Danaher, & Neil McArthur (Eds.), *Robot Sex: Social Implications and Ethical* (pp. 3–14). MIT Press.

Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters.com. Retrieved December 6, 2021 from <https://www.reuters.com/> [article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G](https://www.reuters.com/)

de Beauvoir, S. (1949). *Le Deuxieme Sexe,Vol 1: Les Faits Et Les Mythes*.

Devlin, K. (2015). *In defence of sex machines: Why trying to ban sex robots is wrong*. The Conversation.

dig.watch. (2019). *Discriminatory employment algorithm towards women and disabled*. Retrieved February 26,2022 from [https://dig.watch/updates/discriminatory-](https://dig.watch/) [employment-algorithm-towards-women-and-disabled](https://dig.watch/)

Doukhan, D., Carrive, J., Vallet, F., Larcher, A., & Meignier, S. (2018). An open- source speaker gender detection framework for monitoring gender equal- ity. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 15–20 April 2018, Calgary, Alberta, Canada.

Ferragne, E., & Pellegrino, F. (2007). Automatic dialect identification: A study of British English. In C. Müller (Eds.), *Speaker Classification II. Lecture Notes in Computer Science* (vol 4441, pp. 243–257). Springer. [https://doi.org/10.1007/](https://doi.org/10.1007/978-3-540-74122-0_19) [978-3-540-74122-0\_19](https://doi.org/10.1007/978-3-540-74122-0_19)

Frissen, V. (2018). *Gender is calling: Some reflections on past, present and future uses of the tele- phone*. Taylor & Francis.

GenderLess Voice. (2019). *Meet Q: The first genderless voice – full speech*. Retrieved December 6, 2021 from <https://www.genderlessvoice.com/>

Haraway, D. J. (1985). *A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s*. Center for Social Research and Education.

Hitosugi, M., Koseki, T., Hariya, T., Maeda, G., Moriguchi, S., & Hiraizumi, S. (2018). Shorter pregnant women restrained in the rear seat of a car are at risk for serious neck injuries: Biomechanical analysis using a pregnant crash test dummy. *Forensic Science International*, *291*, 133–137.

Human-Rights-Channel.coe. (n.d.). *Sexism: See it. Name it. Stop it*. Council of Europe. Retrieved February 26, 2022 from <https://human-rights-channel.coe.int/> [stop-sexism-en.html](https://human-rights-channel.coe.int/)

Irigaray, L. (1992). *Elemental passions*. (J. Collie & J. Still, Trans.). Athlone.

Jecker, N. S. (2021). Nothing to be ashamed of: Sex robots for older adults with disabilities. *Journal of Medical Ethics*, *47*(1), 26–32.

Kimmel, M. S. (1986). Introduction: Toward men’s studies. *American Behavioral Scientist*, 29(5), 517–529.

Kristeva, J. (1986). *The kristeva reader*. Columbia University Press.

Kurpicz-Briki, M., & Leoni, T. (2021). A world full of stereotypes? Further inves- tigation on origin and gender bias in multi-lingual word embeddings. *Frontiers in Big Data*, *4*, 20.

Language-Log. (2017). *Nationality, gender and pitch*. Retrieved December 6, 2021 from [http://itre.cis.upenn.edu/~myl/languagelog/archives/005104.html](http://itre.cis.upenn.edu/) Lewis, T. (2015). *Rise of the fembots: Why artificial intelligence is often female*. [https://](https://www.livescience.com/)

[www.livescience.com/49882-why-robots-female.html](https://www.livescience.com/)

Lishchuk, R. (2021). *How large would tech companies be if they were countries?* Retrieved December 6,2021from [https://mackeeper.com/blog/tech-giants-as-countries/](https://mackeeper.com/)

Maddocks, S. (2020). ‘A deepfake porn plot intended to Silence me’: Exploring continuities between pornographic and ‘political’deep fakes. *Porn Studies*, *7*(4), 415–423.

Mager, A., & Allhutter, D. (2021). *How fair is the AMS algorithm?* [https://epub.oeaw.](https://epub.oeaw.ac.at/) [ac.at/?arp=0x003c4bd9](https://epub.oeaw.ac.at/)

McGlynn, C., & Rackley, E. (2017). Image-based sexual abuse. *Oxford Journal of Legal Studies*, *37*(3), 534–561.

Ohlheiser, A. (2013). *Meet‘Titstare,’ the tech world’s latest ‘joke’ from the minds of brogrammers*. The Atlantic. Retrieved December 6, 2021 from <https://www.theatlantic.com/> [technology/archive/2013/09/titstare-tech-worlds-latest-brogrammer-joke-](https://www.theatlantic.com/) [techcrunch-disrupt/311308/](https://www.theatlantic.com/)

Oudshoorn, N., Rommes, E., & Stienstra, M. (2004). Configuring the user as everybody: Gender and design cultures in information and communication technologies. *Science,Technology, & Human Values*, *29*(1), 30–63.

Oudshoorn, N., Saetnan, A. R., & Lie, M. (2002, July). On gender and things: Reflections on an exhibition on gendered artifacts. *Women’s Studies International Forum*. (Vol. 25, No. 4, pp. 471–483). Pergamon.

Perez, C. C. (2019). *Invisible women: Exposing data bias in a world designed for men*. Random House.

Ramboll. (2022). *AI for gender equality—Addressing inequality through AI*. [https://www.](https://www.vinnova.se/) [vinnova.se/globalassets/mikrosajter/ai-for-jamstalldhet-starker-tillvaxten-](https://www.vinnova.se/) [samhallsekonomin-och-arbetsmarknaden/ramboll---ai-for-gender-equality-](https://www.vinnova.se/) [2020-11-19.pdf](https://www.vinnova.se/)

Richardson, K. (2016). The asymmetrical ‘relationship’ parallels between prosti- tution and the development of sex robots. *Acm Sigcas Computers and Society*, *45*(3), 290–293.

Roberts, C., Davies, E., & Jupp, T. (2014). *Language and discrimination*. Routledge. Scheutz, M., & Arnold, T. (2016). *Are we ready for sex robots? 2016 11th ACM/IEEE*

*International Conference on Human-Robot Interaction (HRI)*. 7–10 March, Christchurch New Zealand.

Softbank. (n.d.). *Pepper is a humanoid-robot*. Retrieved September 18, 2022 from [https://developer.softbankrobotics.com/pepper-qisdk/design/pepper-](https://developer.softbankrobotics.com/) [humanoid-robot](https://developer.softbankrobotics.com/)

Søraa, R. A. (2017). Mechanical genders: How do humans gender robots? *Gender, Technology and Development*, 21, 99–115.

Søraa, R. A., & Bruijning, N. (forthcoming). Gendering Sophia the robot: Cyborg, fembot, fashionista, citizen, imagination, and boundary object. In NA (Ed.), *Gender of things*. Routledge.

Spivak, G. C. (2003). Can the subaltern speak? *Die Philosophin*, *14*(27), 42–58.

Strengers, Y., & Kennedy, J. (2020). *The smart wife: Why Siri, Alexa, and other smart home devices need a feminist reboot*. Mit Press.

Vicenik, C., & Sundara, M. (2013).The role of intonation in language and dialect discrimination by adults. *Journal of Phonetics*, *41*(5), 297–306.

West, M., Kraut, R., & Ei Chew, H. (2019). *I’d blush if I could: closing gender divides in digital skills through education*. [https://unesdoc.unesco.org/ark:/48223/](https://unesdoc.unesco.org/) [pf0000367416](https://unesdoc.unesco.org/)

Xu, F., Dan, Y., Yan, K., Ma, Y., & Wang, M. (2021). Low-resource language dis- crimination toward chinese dialects with transfer learning and data augmen- tation. *Transactions on Asian and Low-Resource Language Information Processing*, *21*(2), 1–21.

3

**هویت جنسیتی در هوش مصنوعی**

**برجستگی جنسیتی، هویت جنسیتی و دودویی‌ها**

تنوع جنسیتی بیش از شمردن دودویی زن و مرد است، همانطور که بسیاری از برنامه‌های کامپیوتری این کار را به عنوان "شمارش تنوع جنسیتی" انجام می‌دهند. من تصمیم به تمرکز بر روی جهت‌گیری جنسی و برجستگی جنسیتی گرفته‌ام، از طریق دیدگاه نامتعارف، با تمرکز بر روی نحوه تأثیرگذاری اعضای جامعه LGBTQ+ بر - و تأثیرگذاری آنان توسط - هوش مصنوعی. در اینجا، عبارت LGBTQ+ به معنای لزبین، گی، دوجنسگرا، ترنسمر، نامتعارف و هر کس دیگری که در این جامعه شناخت داده می‌شود، و یک umbrella term است که در برخی جوامع و فرهنگ‌ها مورد اختلاف است. در این کتاب، عبارت "queer" یا نامتعارف به معنای این گروه اقلیت استفاده خواهد شد، با توجه به اینکه همچنان شامل برخی افراد است و برخی دیگر را از جمع خارج می‌کند. در حالی که عبارات و دسته‌بندی‌های دیگری پیشنهاد شده‌اند، عبارت نامتعارف به عنوان یک نقطه نظر قابل استفاده در نظر گرفته شده است تا به بررسی نحوه برخورد هوش مصنوعی با جنسیت و جنسیت‌گرایی بپردازیم، باز کردن فضای فهم ما به غیر دودویی و چالش قرار دادن ساختار منطقی (0-1) کامپیوتر.

**"پتانسیل تبعیض‌آمیز بودن هوش مصنوعی"**

حدس زدن جنسیت کاربران اثرات گسترده‌ای دارد. از دیدگاه شرکت‌ها، برای بازاریابی منطقی است و نقطه‌ی درآمدی را اضافه می‌کند، زیرا بیشتر درباره کاربر فردی می‌آموزد و برای چه چیزی بیشتر احتمال دارد کلیک کند. دانستن اینکه آیا کاربر بیشتر به تبلیغات نخ یا تبلیغات آرایشی علاقه دارد، به بازاریابی هدفمند کمک می‌کند. با این حال، تصویر تاریکی وجود دارد. جنسیت زیادی از مردم در سراسر جهان یک مسئله شخصی بسیار مهم است، به خصوص در مناطقی با دولت‌های کنترل‌کننده و جوامع خصومت آمیز. این مهم است که قضیه LGBTQ+ را به غربی‌سازی نکنید، بلکه بپرسید که افراد عجیب و غریب کجا هستند؟ افشای جنسیت به عنوان یک فرد باعث تأثیرات بزرگی می‌شود برای افرادی که در مناطقی زندگی می‌کنند که شرایط حقوقی، اجتماعی یا مالی برای زندگی در جایی که بودن همجنس‌گرا کیفر خواهد داشت، و یا از دسترسی به حقوق، بهداشت و کار بیرون می‌افتند، حتی ممکن است منجر به مرگ شود. افراد عجیب و غریب در سراسر جهان با تبعیض گسترده روبرو هستند و در حالی که بسیاری از کشورها نسبت به افراد عجیب و غریب بازتر شده‌اند از آنچه در چند دهه قبل بود، بسیاری از مناطق در سراسر جهان از قوانین سختی در برابر جفت‌های همجنس‌گرا روبرو هستند، از تبعیض قانونی در کشورهای مختلف تا حکم مرگ برای همجنس‌گرایی در چند کشور.

برای ساکنان همجنس‌گرای این کشورها، که بدون شک وجود دارند و به دلیل هویت جنسی خود در شرایطی قابل تهدید زندگی زیر سوال قرار گرفته‌اند، هر گونه هوش مصنوعی مانند پیش‌بینی‌های رسانه‌های اجتماعی یا تشخیص جنسیت از طریق صورت در صورت قرار گرفتن در دستان نادرست، ممکن است موجب حکم مرگ شود. در این کشورها، هر فناوری که به نیروهای حاکم برای نظارت، کنترل، تنبیه و حتی کشتن ساکنان خود اجازه می‌دهد، در برابر خطرات نقض حقوق بشری قرار دارد. به عبارت دیگر، هوش مصنوعی نه تنها یک سوال در مورد فناوری است - بلکه بسیار سیاسی است و طراحی آن برای افراد خاص واقعا می‌تواند تفاوت میان زندگی و مرگ را به همراه داشته باشد. برای بررسی عمیق‌تر اینکه چگونه هوش مصنوعی می‌تواند به شیوه‌ای قابل اعتماد منظم شود، به "رویکرد Queer in AI به مدیریت خطرات هوش مصنوعی" (Agnew et al.، 2021) مراجعه کنید.

تاریخچه حقوق جنسیتی در سطح جهانی نشان می‌دهد که افراد همجنس‌گرا به سمت بیشترین حقوق حرکت می‌کنند. در آوریل 1952، انجمن روانپزشکی آمریکا همجنس‌گرایی را به عنوان یک اختلال شخصیتی نامتعارف و بیمارگونه تعریف کرد. در سال بعد، رئیس‌جمهور آمریکا آیزنهاور دستورالعملی را امضا کرد که "همجنس‌گرایان" را از کار در دولت فدرال منع کرد. پس از این اتفاقات، اعتراضات استون وال در سال 1969، که پلیس در نیویورک بر خلاف جامعه همجنس‌گرایان حمله کرد که به عنوان آغاز جنبش حقوق شهروندی همجنسی در ایالات متحده شناخته می‌شود. با این حال، تعیین سال 1969 به عنوان آغاز جنبش، تلاش‌های قبلی را از بین می‌برد. به عنوان مثال، جامعه ماتاچین، سازمان حقوق همجنسگرایان که در سال 1950 تأسیس شد، و یا اعتراضات کافه کامپتون در سال 1966 در سانفرانسیسکو، که جامعه تراجنسی را نیز در مقابل زورگویی و خشونت پلیس قرار داد. در تاریخ 15 دسامبر 1973، انجمن روانپزشکی آمریکا همجنس‌گرایی را از فهرست اختلالات روانی خود حذف کرد و به طور جهانی، سازمان بهداشت جهانی در تاریخ 17 مه 1990 همین کار را انجام داد. در تاریخ 2 مارس 1982، ویسکانسین اولین ایالتی شد که تبعیض بر اساس جنسیت را غیرقانونی اعلام کرد. ازدواج همجنسگرایان در هلند در تاریخ 1 آوریل 2001 قانونی شد و پس از آن با تصویب اولین قطعنامه حقوق همجنسی توسط سازمان ملل متحد در ژوئن 2011 بسیاری از کشورهای دموکرات غربی همین کار را در دهه‌های بعد انجام دادند.

با این حال، همچنان همجنسگرایی به عنوان یک حق شناخته شدن در سطح جهانی وجود ندارد. در ماه اکتبر سال 2009، یک عضو پارلمان اوگاندا لایحه "کشتن همجنس‌بازان" را معرفی کرد. در اوج شیوع بیماری کووید-۱۹، افراد همجنس با سرزنش به عنوان عامل تنبیه الهی باور شده به خاطر بروز این بیماری در اوکراین و سنگال مورد آزار و اذیت قرار گرفتند و در پاناما، افراد ترنس‌جندر به هر دلیلی که در روزهای قرنطینه برای مردان یا زنان مشخص شده بود، مورد آزار و اذیت قرار گرفتند (reid، ۲۰۲۰). در همین زمان، دانشمندانی مانند پوآر (2013) به ما توصیه می‌کنند که «هوموناسیونالیسم» را به عنوان چیزی که نه تنها به حقوق قانونی مرتبط است، بازاندیشی کنیم و فرآیندهای اجتماعی عمیق‌تری را که باعث می‌شود "دوستی با همجنسگرایان" را به‌عنوان یک شاخص از کشورهای تمدنی در نظر گرفته می‌شود، زیر سوال ببریم.

چرایی این موضوع برای هوش مصنوعی بسیار مهم است نحوه برخورد جامعه با گروه‌های اقلیت باید برای شما به شدت مهم باشد، حتی اگر در جامعه LGBTQ+ نباشید. همانطور که تاریخ بارها نشان داده است، تبعیض اکثریت در برابر گروه‌های آسیب پذیر و اقلیت‌ها رخ می‌دهد و بر کیفیت زندگی آنان بسیار تاثیرگذار است. جامعه به طور کلی از تنوع و شامل بودن سود می‌برد و مراقبت از همه یک ارزش بشری بنیادی است. بعدی ممکن است شما با یکی از جنبه‌های وجودتان برای بررسی مواجه شوید. به یاد داشته باشید که عملکرد قدیمی وارد کردن کانری‌ها به معادن در هنگام حفر مواد معدنی چگونه بود؟ طیور تا زمانی که سطح اکسیژن کافی بود، آواز می‌خواندند، اما وقتی شرایط اکسیژن بدتر می‌شد، ساکت می‌ماندند. ما نیاز داریم کانری‌های دیجیتالی، یعنی اقلیت‌هایی را که ابتدا با بروز خارج شدن و تبعیض روبرو هستند، به عنوان هشدارهای جدی در مورد چه چیزی را کم کنیم، حذف کنیم و به اطمینان از اینکه به کنترل و تبعیض بر مبنای هوش مصنوعی در جامعه گسترش نمی‌یابد، نگاه کنیم. دانشمندان LGBTQ+، به خصوص در زمینه علوم STEM، با "فرضیات هترونورماتیو شایع هنوز هم می‌تواند فرم‌های کمتر آگاهانه‌ای از تبعیض و محیطی غیرمهمان‌نواز که دانشمندان از اقلیت‌های جنسی و جنسیتی (LGBTQ) را در معرض عقب‌ماندگی قرار می‌دهد، ایجاد کند" (Freeman، 2018). با ظهور فناوری هوش مصنوعی، کنترل و تبعیض میتواند خطرناک باشد مگر اینکه به شکل مسئولانه پیاده سازی شود. بهتر است فناوری‌هایی را که ثابت شده تأثیری بر این دارند، مورد بررسی قرار دهیم.

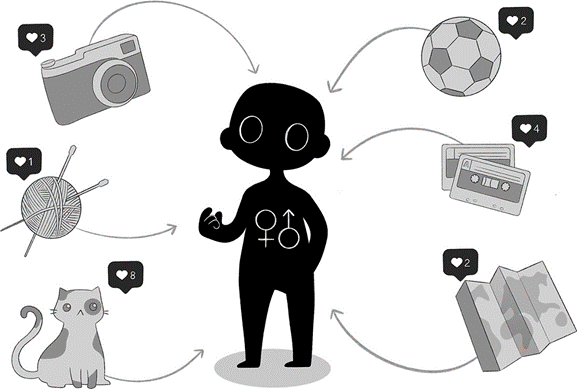
**پیش‌بینی جنسیت و جنسیت جنسی**

آنچه ما در رسانه‌های اجتماعی "لایک" می‌کنیم، چه مواردی را "دنبال" می‌کنیم و چه محتوایی را "اشتراک" می‌گذاریم، بسیاری از خصوصیات ما به عنوان افراد را نشان می‌دهد - که می‌تواند توسط سیستم‌های هوش مصنوعی شناسایی و استفاده شود. مطالعه‌ای از Kosinski و همکارانش در دانشگاه کمبریج در سال 2013 با استفاده از الگوریتم‌ها برای پیش‌بینی دین، سیاست، نژاد و جنسیت جنسی حدود ۵۸٬۰۰۰ کاربر فیسبوک صحت پیش‌بینی "بین مردان همجنس‌گرای و هتروجنس‌گرای در ۸۸٪ از موارد، آمریکایی‌های سیاه‌پوست و سفیدپوست در ۹۵٪ از موارد و بین دموکرات و جمهوری‌خواه در ۸۵٪ از موارد" را به خوبی تأیید کرد. آیا این یک بازی حدس‌زدن ساده و بی‌خطر است یا می‌تواند منجر به تبعیض شود؟ در رسانه‌های اجتماعی، حدس زدن جنسیت مبنایی برای درآمد بازاریابی است و ممکن است پیچیده شود زمانی که به جنسیت جنسی مرتبط می‌شود. چندین کاربر گزارش داده‌اند که الگوریتم‌های رسانه‌های اجتماعی قبل از آنکه خودشان متوجه شوند، می‌دانستند که آنها همجنس‌گرا یا دوجنس‌هستند، مانند موردی که در تیک‌تاک یک زن تازه‌وارد هویت همجنس‌گرای خود را شناسایی کرد:

به نظر می‌رسد تیک‌تاک مناسب سفرهای خودشناسی جنسی است و در مورد زنانی که به زنان علاقمند هستند، فکر نمی‌کنم این مسئله فقط به الگوریتم باهوش برمی‌گردد. فرمت ویدیوی کوتاه، منجر به بارش‌های شگفت‌انگیزی از لحظات پوشیدگی روحی می‌شود، با زوایای دید دوربین POV که با فرهنگ‌های نگاه مردانه در فیلم و تلویزیون که در بیشتر موارد با شهوت مردانه هتروسکسوالی در هم آمیخته است، متفاوت است.

هر چند رسانه‌های اجتماعی ممکن است درست باشند - بگوییم کاربر آماده شنیدن صحیح باشد یا نباشد - اما ممکن است اشتباه هم بکنند. به عنوان اینکه هوش مصنوعی یکی از اجزای اصلی رسانه‌های اجتماعی است، الگوریتم‌های استفاده شده باید با دقت و توجه بسیاری مورد بررسی قرار گیرند. مطالعاتی مانند مطالعه‌ای از Fosch-Villaronga و همکارانش در سال 2020 نشان داد که مردان همجنس‌گرا در رسانه‌های اجتماعی به طور معمول بیشتر به خاطر نامتناسب بودن جنسیتی اشتباه شناسایی می‌شوند و افراد غیردودویی نیز در فراوانی زیادی به اشتباه شناسایی می‌شوند. این موضوع به علت طبقه‌بندی اکثریت، تقسیم بندی آنچه که شما روی آن کلیک، لایک و نظر می‌دهید به چه دسته‌بندی‌های "زنانه" و "مردانه" می‌پردازد و سپس شما را به عنوان کاربر، بر اساس علاقه‌های شما به دسته‌بندی مرد یا زن قرار می‌دهد، همانطور که در شکل ۳.۱ نشان داده شده است.

اما این تقسیم دوتایی به "مردان مردانه هستند" و "زنان زنانه هستند"، مفهومی استریوتایپیک بیشتری را در "مردان از سیاره مریخ؛ زنان از سیاره ونوس" فراهم می‌کند که باعث ایجاد دو دسته بشریت تقریباً زندگی می‌کنند روی سیارات مختلف می‌شود که تمام محدودیت‌هایی که این شیوه تفکر نشان می‌دهد را به همراه دارد. اگر مداخله‌ای اساسی صورت نگیرد، پسران در جهان آبی کالاهای بازاریابی شده و علاقه‌هایی که مردانگی سخت را با قراصنه‌ها، کشتی‌های فضایی و سلاح‌ها تسهیل می‌کند، بزرگ می‌شوند در حالی که به دختران یک داستان خیالی صورتی با براقیت، تک‌شاخانه‌ها و کارهای خانگی ارائه می‌شود. فشار جامعه برای تعیین ظاهر و علاقه‌های جنسیتی، بسیار بزرگ است و این داده‌های تاریخی هستند که ممکن است ماشین‌ها را مجبور به تولید دوباره کنند.



**طبقه‌بندی و استثنا**

اکثر اوقات، جامعه کوییر، درک فضای متحرک جنسیتی دارد و دودویی جنسیتی را نابود می‌کند. به یاد داشته باشید که در فصل قبل در مورد جنسیت صحبت کرده بودیم و چگونه ما اغلب به اجرای نقش جنسی خاصی یاد می‌دهیم، به عنوان مثال، پسری همجنس‌گرا که به او گفته می‌شود "مرد شو" و ورزش بازی کند به جای رقصیدن، یا با دختران قرار بدهد، زیرا این نقش‌های جنسیتی سنتی جامعه را "صحیح" تلقی می‌کنند. اما در فرهنگ خودمان، اغلب فراموش می‌کنیم که آنچه که به عنوان "مردانه" یا "زنانه" در زیرگروه خودمان تلقی می‌شود، ممکن است به طور کامل در مکان‌های دیگر جهان خارج از دیدگاه باشد. مسئله با کامپیوترها این است که آنها نمی‌توانند به روابط جنسیتی فراتر از دودویی مرد و زن بپردازند، به خصوص برای فناوری‌های نظارتی (Katyal & Jung، 2021). به عنوان مثال، گزارش شده است که زمانی که سیستم‌های هوش مصنوعی برای نظارت بر محتوای نامناسب رسانه های اجتماعی استفاده می‌شوند، محتواهایی که توسط کاربران ترنسجندر، محافظه‌کار و سیاه قلمبه می‌شوند، به طور نسبی به شدت "غیر مطلوب" شناخته شده و حذف می‌شوند. در حالی که محتوای محافظه‌کار عموماً به دلیل نقض دستورالعمل‌های واقعی سایت حذف می‌شود، محتوای کاربران ترنس‌جندر و سیاه به عنوان "نامناسب" شناخته شده و از سایت حذف می‌شوند، به رغم رعایت سیاست‌های سایت (Haimson et al.، 2021). شبکه‌های رباتی که توسط نرم‌افزارهای مخرب به اشغال درآمده‌اند، گزارش شده است که به کاربران ترنس‌جندر هدف منفی دارند، همچنین به نظر می‌رسد "نفرت در ماشین" در برابر کاربران سیاه و مسلمان نیز وجود دارد (Williams et al.، 2020). اینگونه مسائلی که اشاره گردید، یعنی انتشار اطلاعات نادرست و گرفتار شدن موضوعات توسط ربات‌ها، به صورت سیستماتیک در رسانه‌های اجتماعی و همچنین در دوران شیوع بیماری کووید-۱۹ مشاهده شده است (Himelein-Wachowiak et al., 2021).)

یکی از مشکلات هوش مصنوعی، ناتوانی در درک زیرفرهنگ‌هاست. یک مطالعه بر روی کاربران مشهور توییتر و سطح سمیت پیام‌های آن‌ها با استفاده از فناوری هوش مصنوعی Jigsaw (متعلق به شرکت Alphabet که صاحب گوگل است)، به بررسی نحوه استفاده از اصطلاحات داخلی گروه توسط درگ کوئین‌ها و مقایسه آن با سیاست‌مداران راستگرا و نژادپرستان پرداخت. در این مطالعه، سمیت به عنوان سطحی از نظر نااحترامی و توهین در نظر گرفته شده که ممکن است منجر به ترک بحث شود. این سامانه با پرسیدن نظر افراد درباره پیام‌ها آموزش داده شده است. نتایج نشان داد که "سطح سمیت حساب‌های درگ کوئین‌ها از 16.68 درصد تا 37.81 درصد متغیر بود، در حالی که میانگین سطح سمیت نژادپرستان سفید بین 21.30 درصد تا 28.87 درصد بود". یکی از مشکلاتی که شناسایی شد، همین الگوریتم برچسب زدن کلمات خاص به عنوان سمیت بود، مثلاً "همجنس‌گرا"، "لزبین" و "کویر" که باعث می‌شد توییت‌های افراد همجنس‌گرا درباره زندگی کویری خود به صورت خودکار برچسب سمیت بگیرند (Antonialli, 2019). مثال‌های مخالفی هم وجود دارند، مانند Perspective API که با استفاده از فناوری یادگیری ماشین، پست‌ها و نظرات را جهت تشخیص سایبربلایی و اذیت‌کردن ارزیابی می‌کند (Ramboll، ۲۰۲۲، ص ۱۷).

در مورد درگ کوئین‌ها، که اغلب با استفاده از سبک ارتباطات تند و برندهای شخصی خود به عنوان هنرمندان اجرایی از اصطلاحات توهین آمیز استفاده می‌کنند، خطر این است که توییت‌هایشان سانسور شوند، در حالی که دیگران مانند نژادپرستان سفید، کمتر به رسمیت شناخته می‌شوند، با این حال برخی از آن‌ها در نهایت از توییتر و دیگر پلتفرم‌ها محروم شدند. محروم کردن برخی افراد با نظرات دشمنانه و حمله‌آمیز که خشونت را در پلتفرم‌های رسانه‌ای اعلام می‌کنند، یکی از مراحلی است که می‌توان برای کاهش رفتار ناخواسته اقدام کرد. مدیریت رسانه‌های اجتماعی یک صنعت بزرگ است، به صورت دولتی (به عنوان مثال در مداخله خارجی در انتخابات آمریکا)، برای تحصیل سود و تروریسم، و همچنین برای منافع شرکت‌های خصوصی. برچسب زدن محتوای LGBTQ+ به عنوان محتوای بزرگسال، حتی زمانی که با رهنمودهای سایت همخوانی دارد، فقط یک مثال است. "شبح بن"، یعنی کاهش قابلیت دسترسی محتوا یا کاربران برای مناطق یا جوامع خاص، نمونه دیگری از این است که چگونه رسانه‌های اجتماعی به نفع خود خود تلاش می‌کنند، به عنوان مثال توییتر با محدود کردن دسترسی به پست‌های خاص از سیاست‌مداران یا دیگران که باورهای آن‌ها در مورد کووید-۱۹ یا توطئه انتخابات در مقابل سیاست شرکت است. با این حال، چندین مثال دیگر نیز از زندگی‌های ترانس در رسانه‌های اجتماعی وجود دارد که در آن‌ها دیدار پیدا کرده و تشویق هم‌سن و هم‌رده بیشتر محقق شده است، مانند جامعه یوتیوب (O'Neill، ۲۰۱۴؛ Raun، ۲۰۱۶؛ Tortajada et al.، ۲۰۲۱) که دیدارپذیری و ارتباط با جامعه کلیدی است.

مطالعه دیگری با عنوان "LGBTQ-AI؟ بررسی بیانات جنسیت و جنسیت در چت‌باز" توسط ادواردز و همکاران (2021) نشان داد که چندین چت‌بات، اگرچه "توانایی استفاده از زبان برای بیان هویت را دارند، اما نیز نشان دهنده کمبود تجربیات اصیل جنسیت و جنسیت هستند". بنابراین، اگر چه چت‌بات‌ها می‌توانند برخی از فهمیدنی‌های جنسیتی جهان را تقلید کنند، اما این اجرا نمایی به معنای داشتن درک عمیقتری از چگونگی کار کردن جهان نیست، به خصوص در یک پدیده اجتماعی- فرهنگی پیچیده مانند جنسیت و تنوع.

**خلاصه فصل**

در این فصل، بررسی کردیم که تنوع جنسیتی بیشتر از شمارش دودویی زنان و مردان است. این یک موضوع پیچیده است که شامل هویت جنسیتی، جنسیت، فرهنگ جنسیتی جامعه و سایر موارد است. اگرچه برخی برنامه‌ها سعی در تعیین جنسیت یا هویت جنسیتی کاربران دارند، اما این یک مسیر خطرناک است، زیرا منجر به انحصار و تبعیض می‌شود. در مورد هویت جنسیتی، پتانسیل صدمه رساندن بزرگ است، زیرا در گذشته این بخشی از زندگی مردم بیشترین حق‌ها و تبعیضات را تجربه کرده است. بنابراین، همانطور که در فصل توضیح داده شد، پیش‌بینی و طبقه‌بندی افراد بر اساس برخی ویژگی‌هایشان مانند جنسیت، ممکن است به نتایج ضررآور منجر شود. به عنوان مثال دیدیم که محتوای رسانه‌های اجتماعی به عنوان نامناسب مشخص شده است و چت‌بات‌ها درک کافی از این موضوع ندارند. همچنین دیدیم که درگ کوئین‌ها به عنوان بیان کنندگان سخنان نفرت‌انگیز برچسب خورده‌اند، در حالی که افراد متطرف بیشتر شناخته شده‌اند. اگرچه برنامه‌های کامپیوتری در طبقه‌بندی برخی ویژگی‌های انسانی به خوبی عمل می‌کنند، اما نه تمام ویژگی‌های انسانی باید به صورت دودویی شمرده شوند.

##### REFERENCES

Agnew, W., Pajaro, J., & Subramonian, A. (2021). Rebuilding trust: Queer in AI approach to artificial intelligence risk management. *arXiv preprint arXiv:2110.09271*

Antonialli, D. (2019). Drag queen vs. david duke: Whose tweets are more ‘toxic’. *Wired, July/August edition*. Retrieved September 20, 2022 from [https://www.](https://www.wired.com/) [wired.com/story/drag-queens-vs-far-right-toxic-tweets/](https://www.wired.com/)

Edwards, J., Clark, L., & Perrone, A. (2021). LGBTQ-AI? Exploring expressions of gender and sexual orientation in chatbots. *CUI ’21:Proceedings of the 3rd Conference on Conversational User Interfaces*. 27–29 July 2021, Spain. ISBN: 978-1-4503-8998-3. Fosch-Villaronga, E., Poulsen, A., Søraa, R. A., & Custers, B. H. (2020). Don’t guess my gender gurl:The inadvertent impact of gender inferences. *BIAS 2020: Bias and Fairness in AIWorkshop at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)* (pp. 1–9). 14–18

September 2020.

Freeman, J. (2018). LGBTQ scientists are still left out. *Nature*, 559, 27–28. <https://doi.org/10.1038/d41586-018-05587-y>

Haimson, O. L., Delmonaco, D., Nie, P., & Wegner, A. (2021). Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: Marginalization and moderation gray Areas. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–35. Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M., Ungar, L., Schwartz, H. A., Epstein, D. H., Leggio, L., & Curtis, B. (2021). Bots and misin- formation spread on social media: Implications for COVID-19. *Journal of Medical*

*Internet Research*, 23(5), e26933.

Joho, J. (2021). *TikTok’s algorithms knew I was bi before I did. I’m not the only one*. Mashable. Retrieved February 26, 2022 from [https://mashable.com/article/](https://mashable.com/) [bisexuality-queer-tiktok](https://mashable.com/)

Katyal, S., & Jung, J. (2021). The gender panopticon: Artificial intelligence, gen- der, and design justice. *UCLA Law Review*, 68, 692.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.

O’Neill, M. G. (2014).Transgender youth andYouTube videos: Self-representation and five identifiable trans youth narratives. In Christopher Pullen (Ed.), *Queer youth and media cultures* (pp. 34–45). Springer.

Oliva, T. D., Antonialli, D. M., & Gomes, A. (2021). Fighting hate speech, silenc- ing drag queens? artificial intelligence in content moderation and risks to LGBTQ voices online. *Sexuality & Culture*, 25(2), 700–732.

Puar, J. (2013). Rethinking homonationalism. *International Journal of Middle East Studies*, 45(2), 336–339.

Ramboll. (2022). *AI for gender equality—Addressing inequality through AI*. [https://www.](https://www.vinnova.se/) [vinnova.se/globalassets/mikrosajter/ai-for-jamstalldhet-starker-tillvaxten-](https://www.vinnova.se/) [samhallsekonomin-och-arbetsmarknaden/ramboll---ai-for-gender-equality-](https://www.vinnova.se/) [2020-11-19.pdf](https://www.vinnova.se/)

Raun, T. (2016). *Out online: Trans self-representation and community building on YouTube*. Routledge.

Reid, G. (2020). *A global report card on LGBTQ+ Rights for IDAHOBIT*. Retrieved February 22, 2022 from [https://www.hrw.org/news/2020/05/18/global-](https://www.hrw.org/) [report-card-lgbtq-rights-idahobit](https://www.hrw.org/)

Tortajada, I., Willem, C., Platero Mendez, R. L., & Araüna, N. (2021). Lost in transition? Digital trans activism on Youtube. *Information, Communication & Society*, 24(8), 1091–1107.

Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2020). Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*, 60(1), 93–117.

**4**

هوش مصنوعی و نژاد

**شناسایی، تعصب و مسائل سیستمیک**

"نژاد" یکی از پر بحث ترین و مورد بحث قرار گرفتن ترم ها در تاریخ بشر است - و چندین مطالعه نشان داده است که جوامع در سراسر جهان با ناعدالتی نژادی، تبعیض نژادی و تبعیض انسان محور مواجه هستند (آپیاه، 2010؛ گریفین، 2012؛ مایلز، 2004؛ وست، 2017). به تعریف دیکشنری آکسفورد، نژادپرستی به عنوان "تعصب، تبعیض یا استیزه از سوی فرد، جامعه یا مؤسسه علیه یک شخص یا افراد بر اساس عضویت آنها در یک گروه نژادی یا قومی خاص، به طور معمول یک گروه کوچکی است که محروم شده است." اما، چنین تعاریفی مشکلاتی دارند و ممکن است مسائل ساختاری عمیقی را که پایه جامعه را تشکیل می دهند، نادیده بگیرند. به عنوان مثال، گروسفوگل (2016) از ما می خواهد که نژادپرستی را به عنوان "مادیت حکومت استفاده شده توسط سیستم جهانی" برای منطقه وجود و منطقه غیر وجود ببینیم، در حالی که شمید (1996) به ویژگی های رفتاری، حرکتی و شناختی نژادپرستی اشاره می کند. اولو (2019، ص 11) می نویسد "نژاد یک ساختار اجتماعی است - اصلاحیت علمی ندارد" و همچنین یک مسئله اقتصادی-اجتماعی است که "ریسمان خود را به هر قسمتی از زندگی ما بافته است. گذشته و آینده ما را شکل داده است." با چنین تأثیر فراوانی بر زندگی افراد، چگونه سیستم های هوش مصنوعی حتی به درک نژاد و نژادپرستی می پردازند؟

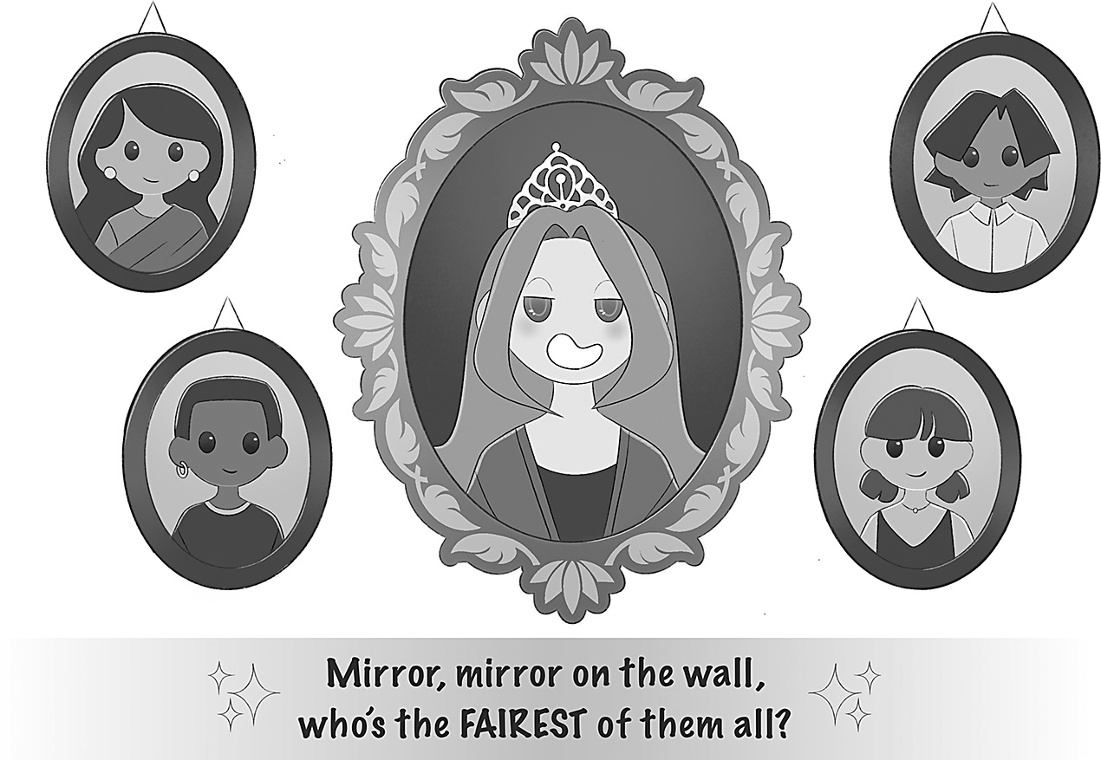
کدهای هوش مصنوعی ممکن است به گونه ای توسط دستگاه ها ساخته یا تفسیر شوند که شامل عوامل غیرقابل قبول باشد و حتی افرادی که کدها را می سازند، می توانند تعصباتی داشته باشند. در این فصل به خصوص به مکانیسم های منحصر به فرد برای این منظور نگاه خواهیم کرد، ولی با آگاهی از اینکه سیستم های هوش مصنوعی تنها محصول روشی است که ما انسان ها آنها را ایجاد می کنیم. سیستم های هوش مصنوعی می توانند نژادپرستانه شوند اگر بر روی داده ها یا منطقی که ساختارهای نژادپرستانه جامعه را به نمایش می گذارند، ساخته شوند. همانطور که در این فصل خواهیم دید، نتایج نژادپرستانه سیستم های هوش مصنوعی می تواند پیامدهای ویرانگری داشته باشد، حتی اگر ناقص و یا به نظر به صورت تصادفی رخ دهد. با این حال، هنگامی که آگاهی از عوامل منجر به انحصار وجود داشته باشد، اقدامات جامعه پذیر می تواند انجام شود. فناوری ممکن است نژادپرستی سیستمی را تولید کند، بنابراین ما باید به آگاهی از اینکه ساختارهای اجتماعی چگونه نژادپرستی را در ابتدا تولید می کنند، نیز توجه کنیم. در این فصل، به نگاه دقیق تری به روش های مختلفی که فناوری در هنگام مواجهه با مسائل نژادی اشتباه کرده است خواهیم پرداخت.

**نوبت فرنولوژیک هوش مصنوعی**

فرنولوژی، یک باقیمانده نژادپرستی و خودسرپرستانه از گذشته بود که بر این باور بود که ویژگی های ذهنی مانند هوش، با اندازه گیری دقیق جمجمه قابل تعیین است. این تئوری که به عنوان یک نظریه مخالفت شده شناخته شده است، با این حال، اثراتی در حال حاضر نیز دارد، مانند شرکت Faception که ادعا می کند از تشخیص چهره برای تعیین هوش یک فرد استفاده می کند. این شرکت به اتهام استفاده از یادگیری ماشین برای ایجاد "دایی نژادپرست" متوسط ​​شده است، که افراد را به دسته بندی هایی مانند "پژوهشگران علمی، بازیکن پوکر یا تروریست" تقسیم می کند (اسنایدر، 2018) - زیرا بر ایده بسیار مورد بحثی که چهره فرد می تواند شخصیت او را تعیین کند، ساخته شده است. با توجه به اینکه سامانه های آگاهی اجتماعی ساختاری هستند که مربوط به جامعه هستند، نیازی به پشت سر گذاشتن زمان برای یافتن مثال هایی از ایدئولوژی های فرنولوژیک برای توجیه "سیاست های علمی" نیست، به عنوان مثال، می خواهید دولت های آپارتاید را ایجاد کنید که در آن بیشتریت یا افرادی که در قدرت هستند، برای "دیگران" در نظر گرفته شده اند.

چندین مورد از نتایج نژادپرستانه ای که هوش مصنوعی به دست می دهد، وجود دارد، مانند مورد شناخته شده حال حاضر نرم‌افزار تشخیص تصاویر AI شرکت گوگل که افراد سیاه پوست را به عنوان "گوریلا" شناسایی می کند، خدمات Flickr شرکت Yahoo که نیز افراد سیاه پوست را به عنوان "میمون" برچسب می زند و اردوگاه کاری نازی داخائو را به عنوان "قایق بازی در جنگل" (Hern، 2018) شناسایی کرده است. همچنین سیستم های تشخیص چهره در کنترل گذرنامه که اغلب به افراد آسیایی می گویند "چشمان خودشان را ببندند". (راه حل گوگل در آن زمان، از حذف دسته بندی "گوریلا" از تحلیل خود به دور از اینکه کافی باشد به نظر می رسید.) همچنین گزارش های متعددی در مورد اینکه "فیلترهای زیبایی" AI چگونه پوست افراد را به رنگ سفیدتر تغییر می دهند و ویژگی های چهره بیشتر شبیه به افراد قوم کوکاسیایی را اضافه می کنند (Wang، 2020) وجود دارد. مثال دیگری از شکست، مربوط به "مانع شدن تبعیض نژادی، سن، جنسیت، ناتوانی و دیگر عوامل تبعیض آفرین توسط انسان ها و هوش مصنوعی با استفاده از جدیدترین پیشرفت های هوش مصنوعی" توسط مرکز تحقیقات Diversity.ai بود، با هدف (1) "استفاده از هوش مصنوعی برای شناسایی، تجزیه و تحلیل، جلوگیری و مبارزه با تعصبات و تبعیض انسانی" و (2) "استفاده از هوش مصنوعی برای جلوگیری از تبعیض توسط هوش مصنوعی". مرکز تحقیقات پس از مسابقه بین المللی زیبایی سال 2016 خود، نیز به بازخورد مثبتی دست یافت که الگوریتم های AI را به منظور انتخاب زیباترین چهره های انسانی "بی طرفانه" از بین بیش از 6000 شرکت کننده از بیش از 100 کشور مجوز داد، بازگشت با برندگان سفید 82٪ (Howard & Borenstein، 2018؛ Levin، 2016). این "چرخه نژادپرستی" به ویژه زمانی آسیب‌های جدی را به دختران و زنان جوان در سراسر جهان وارد می‌کند که تصور می‌شود پوست سفیدتر، زیباتر است. این برای شکل گیری "یک شکل باستانی از تبعیض در مورد رنگ پوست که در عصر اینترنت های مدرن در حال گسترش است" (Wong، 2021) آشکار است که در شکل 4.1 نمایش داده شده است.

آیا این یک نوبت یا بازگشت به شیوه های تاریخی نژادپرستانه است؟ بنجامین (2019a، ص 421) می‌گوید که داده‌های سلامتی که در سامانه‌های هوش مصنوعی وارد می‌شود، دارای تبعیض نژادی تاریخی است؛ به عنوان مثال، نشان می‌دهد که در زمانهای گذشته "قصد افزایش نابرابری‌های نژادی بیشتر صریح بود، اما امروزه، عدم توجه به نظریه‌های نژادپرستانه سیستمیک، دقیقاً باعث پایداری نابرابری‌های رمزگذاری شده می‌شود؛ به دلیل اینکه کسانی که ابزارهای چنینی را طراحی و به کار می‌گیرند، به دقت درباره تبعیض نژادی سیستمیک فکر نمی‌کنند" (Buck و همکاران، 2019).



وقتی هوش مصنوعی زیبایی را تعریف میکند

**تبعیض در تشخیص چهره**

یکی از مشهورترین مثال‌های تبعیض در تشخیص چهره، از پروژه جندر شید Joy Buolamwini در دانشگاه MIT است. بولاموینی به دنبال بررسی علت این بود که چرا چهره خود با پوست تیره‌ای که دارد، تا زمانی که ماسک سفیدی بپوشاند، شناسایی نمی‌شود، و می‌خواست ببیند که چقدر فناوری تشخیص چهره افراد را از مناطق مختلف جهان شناسایی می‌کند، تا ارزیابی کند که آیا تفاوتی در ارزیابی افراد با پوست رنگی متفاوت وجود دارد (Buolamwini، 2017). طبق وب‌سایت gendershades.org، مطالعه بولاموینی 1270 تصویر را به عنوان یک معیار برای آزمون عملکرد تشخیص جنسیت انتخاب کرد، با مشارکت افرادی از 3 کشور آفریقایی و 3 کشور شمال اروپایی، و چهره ها را به دسته بندی های زن و مرد و رنگ پوست روشن و تیره تقسیم کرد، ویژگی های قابل توجهی در خصوص نبود توازن در شناسایی جنسیت برای زنان با پوست تیره و همچنین در شناسایی مردان با پوست روشن یافت. سپس، افراد بر اساس جنسیت، نوع پوست و تلاقی جنسیت و نوع پوست به گروه‌بندی شدند. سه شرکت IBM، مایکروسافت و Face++ که محصولات طبقه‌بندی جنسیتی را ارائه می‌دهند، برای ارزیابی انتخاب شدند. تمام شرکت‌ها در شناسایی جنسیت برای مردان بهتر از زنان با درصد خطای 8.1٪ تا 20.6٪ عملکرد بهتری داشتند. همچنین تمام شرکت‌ها در کل به صورت میانگین بر روی افراد با پوست روشن عملکرد بهتری نسبت به افراد با پوست تیره داشتند با تفاوتی در درصد خطای 11.8٪ تا 19.2٪. (<http://gendershades.org/overview.html>)

مطالعه Gender Shades، نشان می‌دهد که در داده‌هایی که سامانه استفاده می‌کند، تبعیض وجود دارد. وقتی این داده‌ها بیشتر از مردان سفید پوست باشند، آنگاه مردان سفید پوست بهتر شناسایی می‌شوند. بولاموینی همچنین بنیانگذار "اتحادیه عدالت الگوریتمی" است که هدف آن مبارزه با تبعیض الگوریتمی است. همچنین، گزارش‌های متعددی در مورد عدم توانایی وبکم ها در شناسایی افراد با رنگ پوست تیره (CNN.com، 2009) و عمل نکردن دستگاه‌های پخش صابون برای پوست تیره (Fussell، 2017) منتشر شده است. اینها همه مثال‌هایی از یک مسئله سیستمیک بزرگتر هستند که ویلیامز (2020، ص 574) آن را به عنوان "سیستم‌هایی مانند تشخیص چهره، پلیس پیش‌بینی کننده و بایومتریک بر اساس تعداد زیادی از تعصبات و فرضیات تبعیض‌آمیز انسانی بنا شده اند که قبل از هر نوآوری باید نامیده و بررسی شوند"، در حالی که تاریخچه‌های نظارت بر افراد بر اساس بی‌عدالتی کنترل گروه‌های خاصی از مردم، به عنوان مثال بر اساس رنگ پوست آن‌ها، استوار است.

**تبعیض در سیستم های انفورس قانون**

سازمان های انفورس قانون برای پیدا کردن مجرمان از فناوری های "چه کسی این کار را کرده" استفاده می کنند که به گفته ویلیامز (2020) ریشه های تاریخی در الگوهای نژادپرستانه انفورس قانون دارد. از اثر انگشت تا آزمایشگر دروغ شناسی، سیستم قضایی جرم شناسی دارای بسیاری از فناوری های استفاده شده برای تحقیق و دادگاهی است. کمترین فناوری ها بدون تبعیض یا مشکلاتی هستند، اما به عنوان مثال در مورد پایگاه داده های گروه های خلافکاری که "به مشکلات بسیاری مانند داده های ناصاف، تبعیض نژادی، پیامدهای جانبی بی پایان و نتایج بی فایده" (ریچاردسون، 2021) خالی از مشکلات نیستند. این مورد برای COMPAS نیز درست است، یک سیستم قضایی ارجاع دهنده که توسط سازمان رسانه ای ProPublica نشان داده شد که مردان سیاه پوست به عنوان فردی با خطر بیشتر جرم احتمالی در مقابل سفیدپوستان دستگیر می شوند (انگوین و همکاران، 2016) . آنها مثال یک زن 18 ساله سیاه پوست به نام بوردن و یک مرد سفید 41 ساله به نام پراتر را می دهند. هر دو جرم خود را بابت سرقت کالاهای 80 دلاری ارتکاب کردند، اما بوردن با درجه خطر بالاتری درآمده و پیش بینی شد که بیشتر احتمال ارتکاب جرائم خشونت آمیز و کلی را دارد. COMPAS پیش بینی کرد که متهمین سیاه پوست 77٪ بیشتر در معرض خطر بالا برای ارتکاب جرایم خشونت آمیز هستند و 45٪ بیشتر در معرض خطر کلی جرم افزایش می یابند (انگوین و همکاران، 2016؛ بروسارد، 2018، ص 155). اما اگر زنان، به طور کلی، کمتر از مردان جرائم را ارتکاب کنند، سیستم قدرت پیش بینی بیشتری را به نژاد تخصیص می دهد تا جنسیت.

اشتباها به عنوان یک جنایتکار دسته بندی شدن، تأثیرات گسترده ای روی زندگی فرد خواهد داشت - ممکن است امتیازات اعتباری فرد کاهش یابد و تأثیری بر روی پیشنهادات بیمه یا حق بیمه افراد داشته باشد و ممکن است دسترسی به شغل برای فرد دشوار شود. به عبارت دیگر، نامزدهایی که در پایگاه داده های گروه های خلافکاری قرار می گیرند، تأثیراتی مانند لیست سیاه دائمی بر روی زندگی آنان خواهد داشت و این موضوع می تواند منجر به تفاوت در بازخورد از سوی نهادهای خصوصی و عمومی شود و موانعی را برای دسترسی به مسکن، آموزش، شغل، خدمات مالی، مهاجرت، مزایای عمومی و فرصت های اجتماعی به مدت طولانی، اگر نه برای همیشه، قرار خواهد داد.

همچنین فصل کتاب لیتون (2003) با عنوان "فهرست سیاه جدید: تهدید به آزادی های مدنی ارائه شده توسط پایگاه داده های گروه خلافکاری" را نیز ببینید. مثالی دیگر، در حالت بی گناهانه تر، از شور و شوق Pokémon Go در سال 2016 است، که میلیون ها نفر در حال پیاده روی و صید موجودات واقعیت افزوده Pokémon بودند که در سراسر جهان پخش شده بودند، اما تنها با نگاه کردن از طریق تلفن همراه می توانستند دیده شوند و از سیستم های هوش مصنوعی برای ایجاد تجربه کاربری در سراسر جهان استفاده می کردند. یکی از مشکلات آن این بود که بسیاری از "پوک استاپ ها" در حیاط خلوت و مال خصوصی افراد قرار داشت، و بازیکنان از اقلیت های رنگین پوست در گزارش های خود احساس خطر بسیاری کردند.

بسیاری از مردان سیاه پوست به دلیل "مشکوک بودن" یا تعجب از اینکه فردی که حق استفاده از سلاح را دارد، ممکن است چه کاری انجام دهد، صدا زده شده اند. زمانی که ذهن من پیچیدگی بودن وجود سیاه پوست در آمریکا را با پیشنهاد واقعی جستجو و بررسی هایی که در بازی Pokemon GO تعبیه شده است، ترکیب کرد، تنها یک نتیجه وجود داشت. اگر ادامه دهم، ممکن است مرگ برایم پیش آید. (از کتابخانه های عمومی، 2016)

حتی با بازی بی گناهانه مانند Pokemon Go، می توان دید که فرهنگی و فنی چگونه به هم متصلند. بازی با فناوری خود، جهان واقعی را به واقعیت افزوده تبدیل می کند، جایی که اسرارآمیز موجودات Pokémon ظاهر می شوند، بنابراین به دنیای واقعی لایه اضافی اضافه می شود، همچنین نشان می دهد که ارزش ها و نرم افزارهای اجتماعی به چالش کشیده شده است (برای مثال، جایی که افراد BIPOC در جستجوی Jigglypuffs می توانستند به آزادی حرکت داشته باشند). برای برخی از بازیکنان بازی، صید Pokemon در شب دیر، به دلیل اینکه بودن سیاه پوست و سرگردانی در باغچه های افراد در آمریکا "مشکوک" قلمداد شده است، امکان پذیر نبود.

**هوش مصنوعی کوررنگی ندارد**

نژادپرستی یک موضوع تک بعدی با راه حل های سریع نیست. نژادپرستی عمیقا در جامعه ریشه دارد و تنها با فعالیت در برابر ساختارهای نژادپرستانه سیستماتیک، گذشته و حال، می توان امیدوار بود که هوش مصنوعی نوع جدیدی از کارهای نژادپرستانه را ایجاد یا تکثیر نکند. در کتاب "نسل پس از فناوری"، بنجامین (2019b، صفحه پشت) به این موضوع اشاره می کند که چگونه نژادپرستی هنوز هم جامعه آمریکا را تعریف می کند و در قالب "کد جیم جدید" فناوری به نژادپرستی کمک می کند، جایی که فناوری مانند هوش مصنوعی در زمینه های نژادی و اجتماعی ایجاد می شود و "وقتی بد (استفاده) می شود، برای افرادی که در حالت خاص زندگی می کنند، ناراحتی ها را به دنبال دارد" و نشان می دهد که "فناوری ستم سیستمیک را در آمریکا تقویت می کند، شبکه دیجیتالی ایجاد می کند که با کدگذاری افراد برای جایی که زندگی می کنند، کار می کنند و بازی می کنند، به آنها علامت گذاری می کند. فناوری روش های تبعیض آمیز را به صورتی کدگذاری می کند که باعث پاسخ های نژادپرستانه به مسائل اجتماعی می شود." بنابراین، در زمان پیاده سازی فناوری که به عنوان "نابینای نژاد" (یعنی "تمرکز روی نژاد ندارد") شناخته می شود، باید بسیار مراقب باشیم - که فقط مشکلات عمیق جامعه را زیر فرش بکشیم.

تجزیه و تحلیل ابزار پیش بینی هزینه‌های مراقبت از سلامت به عنوان نماینده‌ای از نیازهای سلامت، یکی از مطالعات Obermeyer و همکاران (2019، صفحه 421) است. این ابزارها میزان هزینه‌هایی که برای مراقبت از سلامت یک فرد صرف خواهد شد را تخمین زده و با استفاده از این هزینه‌ها، نیازهای سلامتی را بررسی می کنند. با این حال، همانطور که بنجامین (2019a، صفحه 447) توضیح می دهد، اینگونه ابزارها می توانند منجر به نتایجی با پیامدهای جانی شوند زیرا "بیماران سیاه پوست با امتیاز ریسکی مشابه با بیماران سفیدپوست، معمولاً بیماران بسیار بیمارتری هستند، زیرا مراقبت‌های کاملی برای درمان آنها انجام نمی شود". در نتیجه، این ابزار ممکن است پیش بینی کند که مقدار کمتری برای یک فرد سیاه‌پوست صرف خواهد شد نسبت به یک فرد سفید پوست و در نتیجه نتیجه بگیرد که فرد سیاه پوست در وضعیت بهتری قرار دارد، در حالی که در واقع این علائمی از نژادپرستی ساختاری و بین فردی در سیستم بهداشت و درمان است که بیشتر منابع را برای بیماران سفیدپوست هزینه می کند. بدین ترتیب، فرد سیاه پوست در این حالت ممکن است نیازهای سلامتی بیشتری داشته باشد نسبت به فرد سفید پوست و استفاده از چنین ابزارهایی، تنها عدم توازن موجود در زمینه سلامت را تشدید می کند.

یک مثال دیگر از الگوریتم‌های جستجو، مطالعه Nobl (2013، صفحه 1) درباره مسائل جنسی و استعماری با جستجوی کلمات "دختران سیاه پوست" است. نتایج جستجو در گوگل نشان دهنده قدرت اجتماعی آمریکا و تعصب نژادی و جنسیتی به جای نمایش منافع اجتماعی، سیاسی و اقتصادی زنان و دختران سیاه پوست که به منافع شرکای تجاری و تبلیغ کنندگان اولویت می دهد.

این موضوع در کتاب جالب "الگوریتم های ستم" (Noble، 2018) اوضاع بیشتر شرح داده شده است. یکی از نتایج مثبت این تحقیق این است که حداقل در حال حاضر، عبارت جستجو برای زنان و دختران سیاه پوست در جستجوی تصاویر با جنسیتی شدن همراه نیست - با این حال، تنها کافی است برای جستجوی زنان و دختران آسیایی به جستجو بپردازیم تا ببینیم که این هنوز هم یک مسأله اساسی است، حتی در سال 2022. این موضوع نشان می دهد که چه کسانی در جامعه حضور دارند، چگونه نمایش داده می شوند و دعوت می شوند، بسیار مهم است (برای مثال، مطالعات Prescod-Weinstein [2020] در مورد تجربه شواهد سفید در فیزیک و رسانه‌ای کردن فلسفه علمی)، و هنگامی که هوش مصنوعی از انسان‌ها یاد می گیرد، باید اولین بار نژادپرستی گونه خودمان را پاک کنیم.

**خلاصه فصل**

سیستم‌های هوش مصنوعی ممکن است پیامدهای نژادپرستانه و تفسیرات نژادپرستانه داشته باشند، زیرا کدهای فنی به گونه ای نوشته یا تفسیر شده است که گروه های خاصی از افراد را مستثنی می کند. چه بخشی از نتیجه نوشتارهایی است که به منظور تبعید هوش بر اساس چهره ها نوشته می شود یا به دلیل عدم ورود آمار ورودی متنوع، همانطور که در مطالعه Gender Shades مشاهده می شود، تبعیض می تواند نتیجه برنامه ها باشد. ما دیدیم که چگونه نهادهای انفورماتیکی از هوش مصنوعی برای نقشه برداری و ارزیابی مجرمین بالقوه استفاده می کنند، اما به صورتی که ممکن است تبعیض آمیز و پیش فرض باشد. فناوری های به نظر بی ضرر مانند بازی پوکمون، نیز به دلیل ساختارهای اجتماعی نژادپرستانه، می تواند منحصر به فرد باشد. ما باید این سیستم ها را در یک زمینه گسترده تر از یک چیز فنی درک کنیم - زیرا سیستم‌های هوش مصنوعی ممکن است نژادپرست شوند اگر بر داده هایی یا منطقی که ساختارهای نژادپرستی جامعه را بازتاب می دهد، ساخته شوند. یکی از قدم های اولیه مهم، اطمینان حاصل شود که ساختارهای داده نژادپرستانه تکرار نشوند. یک تلاش بسیار بزرگتر، تغییر اجتماعی واقعی است که مشکل نژادپرستی ساختاری را در ابتدا محدود می کند. هنگامی که مسائلی مانند نژادپرستی به عنوان یک مسئله ساختاری مورد بررسی قرار نمی گیرند، تکنولوژی می تواند تعصبات اجتماعی را برقرار کند.

##### REFERENCES

Akil, O. (2016). *Warning: Pokemon GO is a death sentence if you are a Black man*. [https://](https://medium.com/) [medium.com/dayone-a-new-perspective/warning-pokemon-go-is-a-death-](https://medium.com/) [sentence-if-you-are-a-black-man-acacb4bdae7f](https://medium.com/)

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine bias:There’s software used across the country to predict future criminals. And it’s biased against Blacks*. Retrieved December 6, 2021 from [https://www.propublica.org/article/machine-bias-](https://www.propublica.org/) [risk-assessments-in-criminal-sentencing](https://www.propublica.org/)

Appiah, K. A. (2010). *The ethics of identity*. Princeton University Press.

Benjamin, R. (2019a). Assessing risk, automating racism. *Science*, *366*(6464), 421–422.

Benjamin, R. (2019b). *Race after technology: Abolitionist tools for the new jim code*. Polity Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT

Press.

Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A., Choudhury, T., Hauser, M., Kane, J. M., & Ben-Zeev, D. (2019). Relationships between smart- phone social behavior and relapse in schizophrenia: a preliminary report. *Schizophrenia Research*, *208*, 167–172.

Buolamwini, J. A. (2017). *Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*. MIT Press.

CNN.com.(2009). *HP looking into claim webcams can’t see black people*. Retrieved December 06, 2021, from [http://edition.cnn.com/2009/TECH/12/22/hp.webcams/](http://edition.cnn.com/) [index.html](http://edition.cnn.com/)

Fussell, S. (2017). *Why can’t this soap dispenser identify dark skin*. gizmodo. [https://](https://gizmodo.com/) [gizmodo.com/why-cant-this-soap-dispenser-identify-dark-skin-1797931773](https://gizmodo.com/) Griffin, R. A. (2012). I am an angry Black woman: Black feminist autoethnog-

raphy, voice, and resistance. *Women’s Studies in Communication*, *35*(2), 138–157. Grosfoguel, R. (2016).What is racism? *Journal ofWorld-Systems Research*, *22*(1), 9–15. Hern, A. (2018). Google’s solution to accidental algorithmic racism: ban goril-

las. *The Guardian*. Retrieved February 15, 2022 from [https://www.theguardian.](https://www.theguardian.com/) [com/technology/2018/jan/12/google-racism-ban-gorilla-black-people](https://www.theguardian.com/)

Howard, A., & Borenstein, J. (2018). The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. *Science and Engineering Ethics*, *24*(5), 1521–1536.

Levin, S. (2016).A beauty contest was judged by AI and the robots didn’t like dark skin. *The Guardian*, *11*(08). [https://www.theguardian.com/technology/2016/](https://www.theguardian.com/) [sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people](https://www.theguardian.com/)

Leyton, S. (2003). The new blacklists: The threat to civil liberties posed by gang databases. In Darnell F. Hawkins, Samuel L. Myers, Jr., & Randolph N. Stone (Eds.), *Crime control and social justice:The delicate balance* (pp. 109–174). Greenwood Press.

Miles, R. (2004). *Racism*. Routledge.

Noble, S. U. (2013). Google search: hyper-visibility as a means of rendering black women and girls invisible. *In Visible Culture*, 19.

Noble, S. U. (2018). *Algorithms of oppression*. New York University Press.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453.

Oluo, I. (2019). *So you want to talk about race*. Hachette.

Prescod-Weinstein, C. (2020). Making Black women scientists under white empiricism: the racialization of epistemology in physics. *Signs: Journal of Women in Culture and Society*, *45*(2), 421–447.

Richardson, R. (2021). Defining and demystifying automated decision systems.

*Maryland Law Review*, 81, 785–840.

Schmid, W. T. (1996). The definition of racism. *Journal of Applied Philosophy*, *13*(1), 31–40.

Snyder,B.(2018).[https://twitter.com/jbensnyder/status/1064759174574280704](https://twitter.com/) Wang, C. (2020). *Why do beauty filters make you look whiter?* Retrieved February 26 from [https://www.popsci.com/story/technology/photo-filters-white-kodak-film/](https://www.popsci.com/)

West, C. (2017). *Race matters, 25th anniversary:With a new introduction*. Beacon Press.

Williams, D. P. (2020). Fitting the description: Historical and sociotechnical elements of facial recognition and anti-black surveillance. *Journal of Responsible Innovation*, *7*(Supp1), 74–83.

Wong, J. (2021). How digital beauty filters perpetuate colorism. Retrieved March 1, 2022 from [https://www.technologyreview.com/2021/08/15/1031804/digital-beauty- filters-photoshop-photo-editing-colorism-racism/](https://www.technologyreview.com/2021/08/15/1031804/digital-beauty-%20filters-photoshop-photo-editing-colorism-racism/)

**5**

هوش مصنوعی و بدن ها

**سلامت ، پیری و معلولیت**

در فیلم‌های دیستوپیای ماتریکس، هوش مصنوعی و ماشین‌ها به جنگ با بشر پرداخته‌اند و پیروز شده‌اند - با استفاده از بدن مردم به عنوان منابع انرژی و بلایای ذهنی، انسان‌های باقیمانده را در واقعیت مجازی به دام انداخته‌اند. خوشبختانه، این تصویر ذهنی مرتبط با هوش مصنوعی جنرال به عنوان خبیث و شرور، با هوش مصنوعی محدود کنونی که داریم، فاصله دارد. با این حال، هوش مصنوعی جهان معاصر هنوز هم بر تن‌های ما تأثیر می‌گذارد و در این فصل، به بررسی نحوه تأثیر سیستم‌های هوش مصنوعی بر سلامتی، پیری و ناتوانی می‌پردازم.

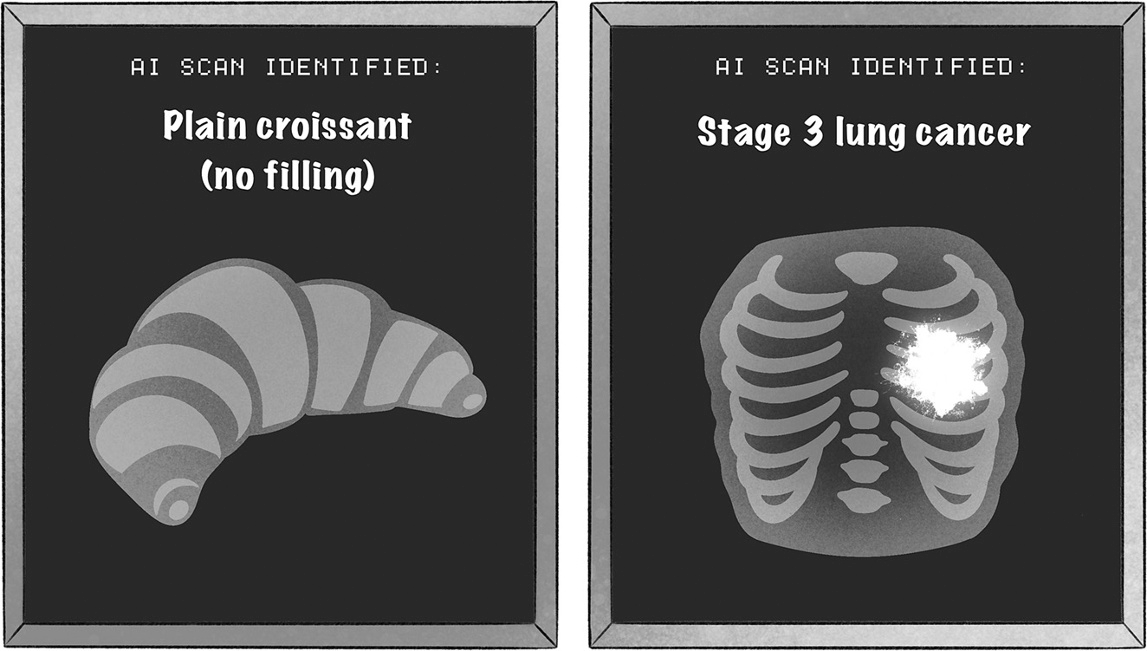
چگونه پیر شدن با هوش مصنوعی کار می‌کند، برای چه کسانی کار می‌کند و چه کسانی بهره‌مند نیستند؟ و چگونه هوش مصنوعی نه تنها بر تن‌های افراد در حال پیر شدن تأثیر می‌گذارد، بلکه بر افراد با تجربیات بدنی متفاوتی نیز تأثیر گذاشته است؟ برای مثال، ما باید بحث را باز کنیم و تأیید و بررسی کنیم که آدم‌های با ناتوانی نیز تحت تأثیر هوش مصنوعی قرار می‌گیرند. تجربیات بدنی، برای نسل‌های جوان هم مهم هستند، بنابراین در این فصل، یک دیدگاه جامع در مورد سلامت و چگونگی تأثیرگذاری هوش مصنوعی بر آن، ترغیب می‌شود، اما با مطالعه بخشی از دنیای گسترده هوش مصنوعی پزشکی و بحث‌های فعلی، با گروه‌های سنی بالاتر که بیشتر در مورد هوش مصنوعی و سلامتی مورد بحث قرار می‌گیرند، آغاز خواهم کرد.

**سلامت مصنوعی یا سلامت هوشمند؟**

سلامتی درمانی که از هوش مصنوعی استفاده می‌کند، دارای پتانسیلی برای پیشرفت‌های نوآورانه در تشخیص و درمان است، اما همچنین مخاطراتی برای اظهار نتایج بیش از حد وجود دارد. همچنین، نگاه به هوش مصنوعی در چارچوب تاریخچه‌ی طولانی کشف و نوآوری نیز مهم است. پزشکی یکی از قدیمی‌ترین علوم انسانی است که شامل یک بدنه‌ی مداوم از دانش است. پیشرفت پزشکی، شامل افزایش هوش پزشکان در مورد چگونگی پاسخ بدن به بیماری و درمان بوده است. بنابراین، پیشرفت در سلامتی، نیاز به پیشرفت در دانش دارد.

اما در مورد موضوع بهداشت، وقتی هوش مصنوعی به معادله اضافه می‌شود، این یک دورو به سختی در حال کار است. بنجامین (2019، ص 421) به عنوان مثال توضیح می‌دهد که "بیشتر سیستم‌های بیمارستانی از ابزارهای پیش‌بینی برای تصمیم گیری در مورد نحوه سرمایه‌گذاری منابع استفاده می‌کنند"، در نظر بگیرید واتسون IBM، بخش هوش مصنوعی شرکت فناوری IBM، که به دنبال تغییر پزشکی مدرن با استفاده از هوش مصنوعی از طریق رویکردهای مختلف است. با این حال، آن‌ها با «عدم تطابق اساسی بین روشی که ماشین‌ها یاد می‌گیرند و روش کار پزشکان» مواجه شده‌اند، و داستان واتسون سلامتی را با عنوان "داستانی از غرور و افراط" برای مبتدیان توصیف می‌کند - که یک پروژه مجلل بود که وعده‌های زیادی داده بود اما هرگز نتوانست نتایج مورد نظر خود را برای حل مسائل بهداشتی که برای خود تعیین کرده بود پیاده سازی کند. همچنین، این برنامه در کتاب واکتر (2015) با انتقاد مواجه شده است به خاطر رویکرد مبتنی بر بازاریابی بیشتر و محصول کمتر. نتایج ضعیف باعث شده است که IBM از تشخیص هوش مصنوعی دور شود، طبق گفته جی رویورو، معاون رئیس جمهور برای تحقیقات در حوزه بهداشت و علوم زندگی IBM، "چیزی است که متخصصان به خوبی انجام می دهند. این کار سختی است و هر چقدر هم که با هوش مصنوعی بدرستی انجام دهید، نمی‌تواند جای پزشک متخصص را بگیرد" (Strickland، 2019). پس از توصیه‌های "علاج نامناسب و نادرست در بیماری سرطان" (Ross & Ike، 2018)، راه به جلو برای IBM نامشخص است.

اما هوش مصنوعی می‌تواند شگفتی‌هایی در پی داشته باشد. به عنوان مثال، سامانه یاد شده به نام BakedScan را در نظر بگیرید. هر کسی که به کافه نان در ژاپن رفته است، بر سر خود شگفت زده می‌شود که چه تعدادی از محصولات نان تازه به صورت باز در دسترس هستند. چراکه این برای صندوقداران سخت است زیرا آنها باید با تنها روی دیداری محصولات نان چندین نوع برابر بدست آیند. وارد شدن BakedScan نیاز را برطرف کرد، که یک محصول توسعه یافته توسط هیساشی کامبه یک ژاپنی است، که BakedScan را به عنوان نرم‌افزار تشخیص تصویری توسعه داد تا به صندوقداران کمک کند که با دقت بسیار بالا نشان دهد هر محصول نان کجاست، نمایش داده شده در شکل 5.1. در ادامه، دکتری با کامبه تماس گرفت که فکر کرد سلول‌های سرطانی مشاهده شده تحت میکروسکوپ بسیار شبیه نان هستند، و با استفاده از همان سیستم اصلی، همکاری به عنوان یک پروژه پرداخته شد که براساس تصاویر سلول‌های سرطانی بسیار خوب تشخیص داد می‌شود (Sommers، 2021). یک مثال مثبت دیگر، پلتفرم نرم‌افزاری QUANTX است، که به دنبال شناسایی آسیب‌های سرطان پستان با استفاده از هوش مصنوعی می‌باشد(Jiang و همکاران، 2021).

استریکلند (2019) بر این باور است که یکی از مسائل عمده در حدس زدنی که هوش مصنوعی انجام می دهد این است که "توصیه های آن عملا به عنوان اثبات شده در نظر گرفته نخواهند شد [...] بدون کنترل های سخت یک مطالعه علمی، یافته چنین گزاره ای فقط به عنوان همبستگی تلقی می شود و نه ایجاد علت". دیگر استفاده های هوش مصنوعی در زمینه پزشکی شامل جراحی رباتیک به عنوان پشتیبان تصمیم گیری بالینی، تجزیه و تحلیل ژنتیکی، اداره پزشکی،(شکل 5.1 نان و سرطان)

و پاتولوژی مربوط به سلامت روانی، همچنین برنامه های سلامت شخصی سازی شده و بازدیدهای بیمارستانی سریعتر ، ربات های هوش مصنوعی که اندوسکوپی را انقلابی می کنند ، و کشف دواهای دیجیتالی ابرمبتنی بر هوش مصنوعی (Data-Core Healthcare، 2022) است. اما قبل از کردن به مثال های محسوس، ابتدا درباره این صحبت خواهم کرد که چگونه سالمندان به عنوان کاربران واقعی فناوری هوش مصنوعی پزشکی می‌توانند تعریف شوند.

**(باز) تعریف "****سالخورده"**

این تکنولوژی ها برای سالمندان به عنوان "فناوری های سالمندی" یا "فناوری های کمکی" شناخته می شوند، که هر فناوری های جهت کمک به افراد با معلولیت یا سالمندان استفاده می شود. اصطلاحات متداول دیگر شامل "زندگی کمکی فعال" در اتحادیه اروپا یا "تکنولوژی رفاهی در اسکاندیناوی" است. فناوری‌های کمکی (AT) شامل انواعی از فناوری‌های مختلف مانند یک عصا برای کمک به راه رفتن، وسایل شنوایی، عینک و سایر تجهیزات حسی، و به روبوت‌های اجتماعی و ایکسواسکلت‌های پیشرفته مانند سیستم‌های هوش مصنوعی ساخته شده در آنها. برخی از این فناوری ها شامل سیستم های هوش مصنوعی مختلفی هستند و در اینجا همین تکنولوژی‌ها مورد تمرکز قرار می‌گیرند.

سن یکی از پارامترهای ساده‌ای است که در سیستم‌ها و الگوریتم‌های کامپیوتری پیاده‌سازی می‌شود، زیرا در بیشتر موارد دارای دامنه مشترکی است: زمان تولد شخص و تاریخ امروز (هرچند برخی استثناء‌هایی در جامعه وجود دارند؛ به عنوان مثال، سن کره‌ای که سال اول در رحم شمرده می‌شود و با سال تازه، برای همه‌ی افراد تغییر می‌کند). در اکثر موارد، سن به نظر یک موضوع ساده است. با این حال، قبل از آنکه به سرعت پارامتر سن را در سیستم کامپیوتری پیاده‌سازی کنیم، باید اصطلاحات را بررسی کنیم. چه کسانی در جوامع امروزی به عنوان "پیر" تلقی می‌شوند؟ این در هر فرد و همچنین در سازمان های مختلف جامعه متفاوت است. کشورهایی که امید به زندگی بلندترین در سیاره دارند، مثلا ژاپن (میانگین حدود 84.3 سال ) و سوئیس (میانگین حدود 83.7 سال) با امید به زندگی بلند، در برخورد شدید با امید به زندگی کمتریر قاره آفریقای جنوبی هستند که به طور متوسط 23 سال کمتر زندگی می کنند (61.6 سال) (Our World in Data، 2019). همچنین، افزایش تاریخی در امید به زندگی وجود دارد، اما تا قسمت پایانی قرن نوزدهم زیادی روند افزایشی ای نداشت - تا آن زمان، امید به زندگی در تمام قاره‌ها به طور متوسط حدود 25-35 سال بود (با این حال مرگ و میر نوزادان با عدد حقیقی این آمار را در جهت ناهموارکردن می‌گرداند).

پس افرادی که " سالخورده " چه کسان هستند؟ آیا این اصطلاح حتی یک اصطلاح خوب است؟ Östlund و همکاران (2015) استدلال می‌کنند که این اصطلاح توهین‌آمیز و قاعده‌مند است و به جای آن در عوض از اصطلاح "بزرگسالان" استفاده کنید که اصطلاحی با بیشتر برخورداری از کار کارکردی است و مسائل دیگر کمتری نسبت به اصطلاح "پیر" دارد. به شکل جمعیتی که دیده‌ایم، امید به زندگی و میانگین سن، نوسان دارد و تعیین عتبات مختلف مشکل است. هنگامی که درباره " سالخورده " فکر می‌کنید، تعصبات خاصی به ذهنتان می‌رسد. در بسیاری از جوامع غربی، سن معیشت معمولاً در حدود 65-67 سال است و به سمت افزایش است. با این حال، این مورد به هیچ وجه برای همه صادق نیست، زیرا اعضای بعضی از حرفه ها مانند سیاستمداران یا استادان دانشگاه ، معمولاً تا سن 80 سالگی کار می کنند ، در حالی که حرفه های دیگر مانند رقصندگان حرفه ای یا افسران پلیس به دلیل استرس جسمی در سنین جوانتر بازنشسته می شوند. با این حال، با وجود اینکه دقیق باشد یا نه، در تصور عمومی بازنشستگی به یکی از اعتبارات عمده تبدیل می شود و با "سالمند" و "پیران" مرتبط می شود. با این حال، بازنشستگی از کار به معنی بازنشستگی از زندگی نیست، زیرا بسیاری از کشورهای غربی در حال حاضر انتظار دارند که بزرگسالان بسیاری از فعالیت های خود را تا سنین بالا ادامه دهند، حتی در صورت بیماری‌هایی که قبلاً ممکن است باعث ناتوانی و بیماری شدن بدون توانایی در انجام انشاط بودند. چنین شرایطی را می‌توان بوسیله فناوری‌های جدید برای سالمندان مدیریت کرد.

**نظارت حسی**

یکی از مثال‌های فناوری بهداشتی با استفاده از هوش مصنوعی، نظارت حسی است. نظارت و مراقبت معمول در مراقبت‌های زخمی معمولاً در مراقبت از افراد مبتلا به زوال عقل رایج‌تر شده است. افراد مبتلا به زوال عقل از عادت‌ها بهره می برند تا عملکردهای شناختی شان به مدتی طولانی ادامه یابد. به عنوان مثال، با حفظ روال‌ها و تمرینات روزانه‌ی ثبت شده، زندگی شخص مبتلا به زوال عقل در ساختاری یکنواخت قرار می‌گیرد. به همین دلیل برای یک شخص دارای زوال عقل بسیار سخت است که به خانه جدیدی نقل مکان کند، زیرا سریعاً رویدادهای جدید را فراموش می‌کند، اما خاطرات گذشته خود را برای مدتی بیشتر حفظ می‌کند. این می‌تواند به عنوان استفاده از نوار اندازه گیری قابل برگشتی در نظر گرفته شود، حیات شما را به عنوان ۱ سانتی متر معادل یک سال تصور کنید. اگر در تاریخ تولد صد سالگی خود به زوال عقل دچار شوید، به آرامی خاطرات خود را بازخواهید گرفت، ابتدا سال‌های ۹۰ تا ۱۰۰ سانتی متری (یا سال‌ها در این بیان مجازی) ناپدید می شود، سپس سال‌های ۸۰ تا ۹۰ سانتی متری و در نهایت حفره خاطرات از نوادگان بزرگ، سپس فراموشی‌های از فرزندان شما رخ می‌دهد، اما شماره تلفن دوستان فعلی شما از مدرسه هنوز برای مدت طولانی محو نمی شود. البته در این مثال ساده، تنوع‌هایی وجود دارد، و هوشیاری شناختی یک حوزه‌ی پیچیده در پژوهش است (Deary و همکاران، 2009؛ Salthouse، 2009؛ Williams و همکاران، 2010).

چگونه می توان از هوش مصنوعی و سیستم های کامپیوتری برای کمک به کاهش مشکلات حافظه استفاده کرد؟ یکی از نوآوری های فناوری در این زمینه فن آوری یادآوری است، جایی که تقویم ها یا هماهنگی عادات با همانکاری برای یادآوری سالخوردگان مبتلا به زوال عقل در مورد صبحانه خوردن، قدم زدن در بعد از ظهر یا تماس با خانواده برای گفتن چگونگی حالشان کار می کنند. یک مثال از این نوع فناوری سیستم های حسی مبتنی بر روبات است که می تواند عادت ها را نظارت کرده و به وابستگان و ارائه دهندگان بهداشت اعلاناتی ارائه دهد. مثالی از این نوع نظارت اتصال روباتی به شکل گیاه به یک سیستم برای بررسی اینکه در هر لحظه درب یخچال باز شده است و یا دیگر داده های حسی می تواند مشاهده شود (Søraa و همکاران، 2021). این یکی از بسیاری از مطالعات در زمینه آزمایش فناوری است تا ببینیم چگونه بهبود کیفیت زندگی برای سالخوردگان به ویژه در مورد پیر شدن در خانه خود و نه در مؤسسات بهبود یابد. هوش مصنوعی در بخش بهداشت دارای پتانسیل بالایی به عنوان مجموعه ای از ابزارهای مفید است که می توانند بهبود زندگی افراد را ایجاد کنند. به عنوان مثال، کمک به سالخوردگان برای زندگی در خانه های خود بدون نیاز به انتقال به مؤسسات قبل از لزوم اساساً یک مسئله اقتصادی برای جامعه است، اما برای فرد، انتقال از خانه به مؤسسه بسیار عمیقتر است. این در مورد از دست دادن خانه به عنوان منطقه امنیتی، از دست دادن تماس با دوستان و خانواده و یک تغییر اساسی در پایان زندگی شخص است.

فناوری هایی مانند سیستم های حسی برای شناسایی، پتانسیل افزایش سلامت را دارند، نه فقط تأخیر در مرگ. سیستم های هوش مصنوعی که در فناوری های موجود متصل شده اند و بدون توجه ما کار می کنند، شاید موثرترین ها باشند. یک کاربر معمولی تلفن همراه خود را درباره ۱۲۰ بار در روز بررسی می کند، و بسیاری از آن ها با مکانیزم های شناسایی چهره قفل می شوند. چه اگر این ۱۲۰ اسکن چهره برای شناسایی بیماری های حاد نیز قابل شناسایی باشد؟ به عنوان مثال، برنامه "شناسایی فوتال" توسعه یافته در هنگ کنگ در سال ۲۰۱۹ سعی کرد با این کار به شما کمک کند، برای مثال با بررسی اینکه آیا یک ناحیه از چهره شما به سمت پایین سقوط کرده است (علامت رایجی از سکته)، به شما هشدار می دهد تا اورژانس را فراخوانی کرده و با شخصی که از قبل انتخاب کرده اید، تماس بگیرید (یت ش.، ۲۰۲۰). موضوعات حریم خصوصی و رضایت کاربر برای اینگونه برنامه ها مهم هستند - از یک طرف، ما همه احتمالاً می خواهیم به سرعت شفا یابیم، اما افراد احتمالاً از بررسی و تأیید رضایت خود در مورد استفاده از داده های صورت خود احتیاط می کنند.

**مراقبت از راه دور**

با توجه به پیر شدن جمعیت جهانی، به ویژه در کشورهای غربی و کشورهای آسیای شرقی، چشم‌انداز زندگی سالم مردم در خانه خود تا جای ممکن ، یک اولویت سیاسی بزرگ شده است. نهادینه سازی تغییر گرانبها ای از زندگی است. در یک نهاد، خدمات سلامتی متمرکز شده‌اند، در حالی که برای سالمندانی که در خانه زندگی می کنند، خدمات سلامتی باید به سمتشان بیاید. اما این تنها خدمات سلامتی بالینی با نمونه خون، دارو، درمان زخم و غیره نیست که بخشی از مراقبت است. بخش تعاملی اجتماعی مراقبت نیز، قابلیت دیدار و تعامل با خانواده، دوستان و جوامع برای بهبودی روانی اهمیت داشته و ضروری است.

مراقبت از فاصله فناورانه نیز می تواند از طریق دستگاه‌ها انجام شود. یک فناوری AI کاربردی و بیشتر دیداری در این زمینه، سیستم جراحی داوینچی (da Vinci Surgical System) آمریکایی است که در سال ۲۰۰۰ ساخته شده است، جایی که پزشک اپراتور کنترل گریپ ها و بازوهای ربات را برای ارائه روش های عملیاتی دقیق انجام می دهد. در اینجا، پزشکان و دستگاه ها در عین حال با هم همکاری می کنند تا بهترین نتایج را ارائه دهند. در ۲۱ سال عملیات خود، این سیستم برای بسیاری از نهادهای پزشکی به خاطر گران بودنش انتقاداتی را به دنبال داشته است و بعد از مرگ برخی بیماران مربوط به عمل جراحی، پرسش‌های اخلاقی قانون نظارتی را شروع کرده است. (بورلا و همکاران، ۲۰۰۸؛ دیمایو و همکاران، ۲۰۱۱؛ میسو و همکاران، ۲۰۱۰). یکی دیگر از مسائل در زمینه تنوع، توانایی دسترسی به این دستگاه گران بها است. به عنوان مثال، در نروژ، اولین سیستم da Vinci توسط بیلیونرهای نروژی اهدا شد، زمانی که یکی از آنها سرطان پروستات داشت و به آن نیاز داشت.

**مشکلات سلامت روان**

برای بسیاری از مردم، سلامت، مترادف با سلامت جسمانی است، اما سلامت روانی نیز مهم است. مطالعات نشان می دهد که 20 درصد از بزرگسالان آمریکا با بیماری روانی زندگی می کنند و تخمینی نیز وجود دارد که 6 درصد دیگر با بیماری روانی تشخیص داده نشده مبتلا هستند. (NAMI، 2020). مطالعات اروپایی نشان می دهند که یکی از هر 6 اروپایی گزارش می دهد دچار بیماری روانی شده است. بسیاری از کشورهای آسیای شرقی نیز نشان داده‌اند که نرخ خودکشی آنها بالاست (مانند کره جنوبی با نرخ 28.6 خودکشی در هر ۱۰۰،۰۰۰ نفر، که آن را چهارمین نرخ بالاتر در جهان می‌سازد) (World-Population-Review، 2022). با این حال، در مقایسه با سلامت جسمانی، مسئله سلامت روانی به شدت از روی نتایج اولویت بندی خارج شده است. آیا هوش مصنوعی می تواند کمک کند؟ چندین مطالعه نشان می دهد که این امکان وجود دارد، از شناسایی اضطراب (Shen & Rudzicz، ۲۰۱۷)، ارزیابی ریسک خودکشی (Just et al.، 2017؛ Morales et al.، 2019)، و ارزیابی افسردگی مردم از طریق رفتارهای رسانه اجتماعی (De Choudhury et al.، ۲۰۱۳؛ Moreno et al.، ۲۰۱۱؛ Schwartz et al.، ۲۰۱۴). D’Alfonso (2020) ، توضیح می دهد که

درواقع، روش‌های هوش مصنوعی مبتنی بر داده می توانند برای توسعه مدل‌های پیش‌بینی / شناسایی برای شرایط سلامت روانی استفاده شوند. به ویژه، "پسماند دیجیتال" فرد، یعنی داده‌های جمع‌آوری شده از دستگاه‌های دیجیتال و تعاملات رسانه‌ای اجتماعی شخص، می‌تواند به عنوان همواره موجود در دسترس بودن فرد برای منبع استخراج اطلاعات مورد نیاز برای درک رفتار یا سلامت روان خود او، به کار گرفته شود.

وی در سه روش اصلی تمرکز می کند که موقعیت استفاده از هوش مصنوعی در حوزه سلامت روانی را پیش‌بینی می کند: (1) شناسایی شخصی یا فنوتایپ دیجیتال، (2) پردازش طبیعی زبان متن های بالینی و محتواهای رسانه های اجتماعی و (3) چت‌بازها.

1. شناسایی شخصی و فنوتایپ دیجیتال (ویژگی های فیزیکی قابل مشاهدۀ افراد از طریق روش های دیجیتال)، که D’Alfonso (2020) به آن اشاره می کند: "برای استنباط اطلاعات زمینه‌ای و رفتاری درباره فرد که می توان از آن به عنوان ورودی برای روش‌های یادگیری ماشین استفاده کرد تا نتایج روانشناسی / روان‌سنجی و شرایط سلامت روانی پیش‌بینی شود".این قطعات اطلاعات می توانند به عنوان مثال، توسط ساعت های هوشمند و دیگر ابزارهای نظارت بر الگوی خواب (Difrancesco و همکاران، ۲۰۱۹؛ Scott و همکاران، ۲۰۱۹)، یا الگوی تایپ در تلفن‌های هوشمند برای تشخیص افسردگی (Mastoras و همکاران، ۲۰۱۹) بازیابی شوند.یک مطالعه دیگر، توسط Buck و همکاران (۲۰۱۹)، نشان می دهد که کاهش تعداد و طول تماس های تلفنی خروجی ممکن است با بازگشت به مرحله بهبودی در اسکیزوفرنی مرتبط باشد. به گفتۀ D’Alfonso (2020)، یک گام دیگر "مداخلات الکولوژیکی لحظه‌ای (EMI) ارائه مداخلات روان‌شناختی لحظه‌ای یا نوعی تحریکات رفتاری از طریق دستگاه‌های شخصی تحویل داده شده در طول زندگی روزمره فرد است، که اغلب با پاسخ‌های آن اطلاعات دریافتی موضوعی تشکیل شده است."
2. پردازش زبان طبیعی متون بالینی و محتوای رسانه های اجتماعی که تحلیلی از نحو و معنی حرف زدن شخص دربارۀ موضوعی مشخص را دارند، می‌تواند به استنباط وضعیت روانی آن فرد کمک کند (کالو و همکاران، ۲۰۱۷). یکی از مکان‌های جمع آوری چنین اطلاعات می‌تواند رسانه‌های اجتماعی باشد، همانطور که D’Alfonso (2020) توضیح داده است: "با تحلیل ویژگی‌های زبانی در محتوای رسانه‌های اجتماعی، ممکن است مدل‌های یادگیری ماشین تولید شوند که بتوانند قبل از روش‌های سنتی وضعیت سلامت روانی فرد را استنباط کنند" (همچنین برای مطالعه بیشتر Briand و همکاران، ۲۰۱۸؛ Eichstaedt و همکاران، ۲۰۱۸؛ Ziwei و Chua، ۲۰۱۹ را ببینید).
3. یک حوزۀ دیگر که مسائل سلامت روان به عنوان یک بخش از هوش مصنوعی مطرح شده است، بحث در مورد چت‌بات‌ها است. زیرا تاریخچه چت بات‌ها به روانشناسی مربوط است و اولین چت بات ELIZA برای شبیه‌سازی یک روان‌درمانگر روجری کار ساخته شده است (Weizenbaum، ۱۹۶۶). بررسی های اولیه نتایج خوبی را نشان می دهد (Abd-Alrazaq و همکاران، ۲۰۱۹) و با توجه به اینکه چندین چت بات برای اهداف روان‌پزشکی مانند Replika, Tess, Woebot و Wysa توسعه داده شده‌اند، تعداد کاربران در میلیون‌ها است، با این حال مسائل اخلاقی مورد بحث است (Fiske و همکاران، ۲۰۱۹؛ Kretzschmar و همکاران، ۲۰۱۹). درباره تمامی کارهای مربوط به سلامت روان مبتنی بر هوش مصنوعی، D'Alfonso (2020) هشدار می‌دهد "اگر بازویی را در برابر هوش مصنوعی قرار ندهیم، این می‌تواند مشکلات اخلاقی سنتی در مراقبت های بهداشتی روانی را شدت بخشد."

در مثال زندگی، چن و همکارانشان در مقاله خود به نام "Can AI Help Reduce Disparities in General Medical and Mental Health Care?" به این نتیجه رسیدند که درمان‌های بهداشتی برای برخی افراد بهبود می‌یابد، اما شاید برای همه بهبود نیابد، مراقب تبعیض در داده‌ها و انتخاب مدل های محاسباتی باشید. در مورد تنوع میان کاربران، پرسیدن اینکه "کدام کاربران" درگیر هستند و چگونه درگیر می‌شوند، می‌تواند مفید باشد. اگر تا یک چهارم از جمعیت بیماری‌های روانی را تجربه می‌کنند، در حدود نصف جمعیت آمریکا (به طور میانگین) در طول عمر با بیماری روانی مواجه خواهند شد (Reeves et al.، 2011)، این مشکل سلامتی بی‌نظیری است که باید با بیشترین توجه و جدیت برخورد شود. با این حال، سلامت روانی هنوز هم در برخی از جمعیت‌ها تبعیض می‌کند. مردان به طور قطعیت کمتر احتمال دارند به دنبال کمک‌های درمانی در زمینه سلامت روانی بگردند و در مورد مشکلات سلامت روانی صحبت کنند (Rickwood et al.، 2007؛ Vogel et al.، 2014). یکی از این مسائل، به عنوان مثال، درگیری ممکن است مربوط به این باشد که "مشکلات سلامت روانی گروه‌های نژادی به شدت به زمینه اجتماعی بزرگتری وابسته هستند"، همانطور که چن و همکارانش در سال 2019 به آن اشاره کردند، در حالی که مشکلات اشتراکی با مردان جوان BIPOC (سیاهپوستان، بومیان و افراد رنگین پوست) همچنان باعث کمتر به دنبال درمان سلامت روانی بودن آنها می‌شوند (Rickwood et al.، 2007).

در زمان توسعه چت بات ها و دیگر ابزارهای هوش مصنوعی پردازش زبان، شناخت گروه های کاربری که به فراموشی سپرده شده‌اند بسیار حائز اهمیت است. به ویژه در مورد مردانی که مستضعف اقتصادی هستند ـ یک گروه گاهی در موضوعات تنوع و استثناها به فراموشی سپرده می‌شود ـ هنوز بسیاری از کارها در زمینه توانمندسازی این گروه ها برای انجام باید باشد. با این حال، در زمینه تنوع و هوش مصنوعی نگران کننده است که گروه هایی که بیشترین منفعت را از کمک‌ها به دست می‌آورند، سخت است تا به عنوان کاربران شامل شوند؛ این نیز تنها به دلیلی است که باورهای مردانگی سمی مربوط به مشکلات روانی، در گفتمان عمومی به عنوان زنانه و ضعیف کننده ایفای نقش می کنند.

**کووید ، هوش مصنوعی و تنوع مراقبت های بهداشتی**

همه گیری کووید-۱۹ در سراسر جهان، بزرگترین و تأثیرگذارترین رویداد بعد از جنگ جهانی ها بوده است. با مرگ میلیون‌ها نفر و درگیری صدها میلیون نفر، اثرات آن بر سلامت، اقتصاد و نگرش آینده جهانی ویرانگر بوده است. از دست دادن جامعه به شدت ویرانگر بوده و شاهد از دست دادن زیاد از سالخوردگان و گروه‌های قومی محروم هستیم. در میان این از دست دادن های تلخ، هوش مصنوعی هم به عنوان قهرمان و هم به عنوان شیطان در این شرایط عمل کرده است. یکی از نقش های مثبت هوش مصنوعی در تشخیص سریع منفی‌های این ویروس است که کمک به شناسایی و هشدار دهی سریع به مأمورین بهداشت در زمان شیوع جدید بوده است؛ به عنوان مثال در مورد استفاده کره جنوبی از فناوری تلفن همراه برای طرح ریزی الگوی حرکت و جایگاه. ربات ها نیز در بیمارستان های دانمارک برای شستشو و ضدعفونی کردن سطوح به‌کار گرفته شدند؛ همچنین برای شستشوی خیابان ها و تحویل دادن کالاها، مانند تحویل غذا در چین و آمریکا، استفاده و بکارگیری شدند. مخزن اطلاعاتی باز گوگل مربوط به کووید-۱۹ به نام “مخزن اطلاعاتی باز اپیدمیولوژی کووید-۱۹” که "یک منبع جامع و منبع باز منابع داده‌های اپیدمیولوژی کووید-۱۹ و متغیرهای مرتبط و اقتصادی و آماری جمعیتی از بیش از ۵۰ کشور" را فراهم کرده، یکی از مثال های وجود پتانسیل هوش مصنوعی برای یادگیری از این ویروس و برای آمادگی برای درمان آینده است.

از سوی دیگر، هوش مصنوعی نیز با عنوان ایمنی ویروس توسط برخی کشورها برای افزایش نظارت‌ و پایش به‌کار گرفته شده است. برخی از کشورها با استفاده از قوانین مبارزه با شیوع اخبار جعلی، برای بازداشت تظاهرکنندگان نه تنها افرادی که در اعتراض به استفاده از ماسک هستند بلکه افرادی که در اعتراض به سیاست‌های ستمکارانه دولت هستند، استفاده کرده‌اند. واضح است که در حال حاضر در مورد استفاده هوش مصنوعی نیاز به قوانین و مقررات وجود دارد تا بتوانیم از دستاوردهای بیشمار این فناوری‌ها بهره‌برداری داشته باشیم، به ویژه برای اطمینان از اینکه تدابیر شامل تنوع و شامل کردن گروه‌های مختلف گرفته شود. راه حل های فناوری خود به تنهایی کافی نیست؛ به عنوان مثال در نظر بگیرید که چگونه سوئد ـ یکی از ثروتمندترین و متعلم ترین کشورهای جهان ـ توانست با انتخابات نادرست و تصمیمات غیر عادلانه، در نتیجه شاهد نرخ مرگ و میر ده برابر علیه همسایگان نروژ خود از این ویروس بوده‌ است.

نوشتن درباره رویدادهای طولانی مدت مانند این باید موازنه‌ای مناسب باشد؛ هنوز نمی‌دانیم چگونه جامعه در برابر "وضعیت عادی جدید" زندگی در جامعه کوویدی پاسخ دادند و تغییر کردند، با این حال، در قضاوت در مورد استفاده از هوش مصنوعی ، نشان می دهد که فناوری می تواند هم به نفع و هم به ضرر مورد استفاده قرار گیرد. در زمینه تنوع، بدون شک کووید-۱۹ تبعیض کار می کند و در گفتمان این بیماری، بزرگسالان بدون ملاک بیشتری تحت فشار قرار گرفته و ارزش زندگی آنها به دلیل برداشتن توجه از سمت جامعه و نارضایتی از وضعیت اقتصادی کاهش یافته است. با این حال، تنها سالمندانی که تحت تأثیر این بیماری قرار میگیرند نیستند. جوانان به خصوص آسیب پذیرند و بسیاری از آنها روی دانشگاه یا دبیرستان خود تجربه مجازی به عنوان تجربه مشکل مثبت تأثیر میگذارد. مناطق فقیر همچنین از این بیماری بی‌توجهی+جهان ثروتمند سود جسته‌اند که پس از تزریق واکسن بهشته در کشورهای خود، آنها را عرضه کرده‌اند. این نشان می دهد که سلامتی عادلانه نیست، درمان ناسزگار است و هر فناوری از ابتدا با مسائل تنوع و استثنا رو‌برو می‌شود.

**هوش مصنوعی و معلولیت**

فناوری های مراقبتی نه تنها برای بزرگسالان، بلکه برای یک گروه کاربری مهم دیگر، یعنی افراد معلول، بدون قید و شرط سن خود قابل استفاده هستند. معلولیت تعریف شده است به عنوان "شرایط فیزیکی یا روانی که حرکت ها، حس ها یا فعالیت های یک شخص را محدود می کند."(تعریف استفاده شده توسط فرهنگ لغت لکزیکو اکسفورد). این می تواند فیزیکی مثل نداشتن قابلیت حرکت، فقدان حس بویایی و یا روانی مانند عدم توانایی در درک نشانه های اجتماعی باشد. گروه کاری که شیوه برخورد با آنها با مشکل روبرو است، باید در ارتباطی با این ارتباط مورد بررسی قرار گیرد. Whittaker و همکاران. (2019، ص 2-3) در مقاله جامع خود در مورد "معلولیت، تبعیض و هوش مصنوعی" ویژگی های چند سوال مهم برای مطرح کردن این ارتباط را فراهم کرده اند که به صورت خلاصه عبارتند از:

۱- چگونه می‌توانیم از فعالیت‌ها و تحصیلات در حوزه معلولیت برای حفاظت از افرادی که خارج از "اصول" بازتاب داده شده و ساختار دهی شده توسط سیستم های هوش مصنوعی ایجاد شده اند، استفاده کنیم؟ (۲) شناخت کردن این مسئله که هم‌آمیزی نژاد، جنسیت، جنسیت و معیارهای طبقاتی غالباً نحوه تعریف و درک "معلولیت" را تحت تأثیر خود قرار می دهند؟ (۳) چه استانداردهای "طبیعی" و "توانمندی" توسط سیستم های هوش مصنوعی خاص تولید و اعمال می شود؟ (۴) چگونه می توانیم پیامدهای بیشتری از تشخیص بیماری و تبدیل شدن به بیمار در زمانی که اجازه "بی‌مشارکتی" داده شده، برجسته کنیم؟ (۵) چگونه چهارچوب های حقوقی مخالف تبعیض، ما را در جستجوی روش هایی برای دفاع از حقوق معلولان در هوش مصنوعی جدید بهره مند کنند؟ (۶) چه راهکارهایی می توانیم از جنبش حقوق معلولان یادگیری کنیم برای ایجاد مسئولیت پذیری بیشتر در هوش مصنوعی؟ (۷) چه نوع طرح طراحی، تحقیق و مهندسی هوش مصنوعی ممکن است آینده های مطلوب تری را برای افراد معلول فراهم کنند؟

این راهنمایی بسیار عملی می‌تواند در بهبود شناخت و تشخیص مسیرها برای درک بهتر در جایگاه معلولان و تنوع کمک کند. همچنین به مرور مقالاتی مانند Hamraie و Fritsch (2019) که در مورد نیاز به بررسی و بازسازی دنیای مادی-گفتاری ما به منظور رسیدن به عدالت برای افراد معلولان در ساخت فناوری سخن می‌گویند، توجه کنید. افراد معلول روزانه با انواع تبعیضات مواجه هستند و با هوش مصنوعی، خطر این است که این مکانیزم‌های انحصاری تکرار شوند و یا مکان‌ها و سیستم‌های انحصاری جدید ایجاد شوند. به عنوان مثال، از نداشتن رامپ برای صندلی‌ها در محیط‌های ساختمانی و چگونگی سختی حرکت در شهرهایی که پله‌ها به مانعی برای افراد کم‌تحرک تبدیل می‌شوند، خودداری می‌کنید. یا به یک کنفرانس کاری که فقط از وقایع نوشیدن و آشنایی با نوشیدنی خاصی به عنوان رویداد اجتماعی استفاده می‌شود، نگاه کنید که افرادی با کمبود ذایقه و بویایی (و همچنین کسانی که نوشیدن نمی‌کنند) را از رویداد حذف می‌کند. یا به مقیاس بزرگتری فکر کنید، یک سیستم آموزشی که بر مبنای تعامل اجتماعی با همکلاسی‌ها بنا شده است.

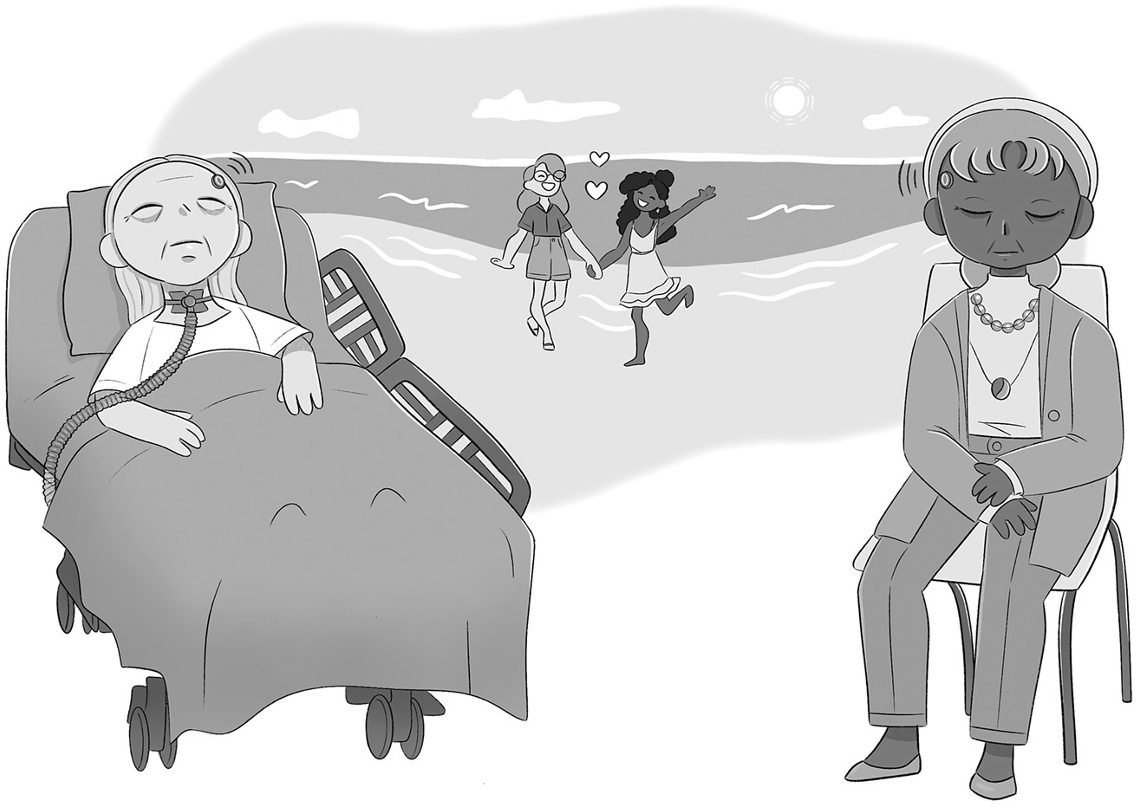
برای افراد معلول، می‌توان به وضوح دید که چگونه هوش مصنوعی در کاربردهای مرتبط می‌تواند برای بهبود زندگی آنها استفاده شود. به عنوان مثال، به مورد افراد کور فکر کنید. این موضوع می‌تواند به طور عادی مرتبط با افراد با ناتوانی در شناخت رنگ باشد. برای مثال، 8٪ از مردان با این مشکل مواجه هستند و با استفاده از رنگ‌ها برای تمایز دادن موارد در وب سایت ها سختی می کشند. برای افراد کور یا به صورت جزئی کور هستند، روش‌های بسیاری وجود دارد که می‌توانند از هوش مصنوعی بهره ببرند. یک مثال را می‌توان به خواندن صوتی متن (مانند "Seeing AI" شرکت مایکروسافت) یا اسکن عکس‌ها در مورد رنگ‌ها برای کمک به انتخاب لباس‌ها برای دختر خود توسط "Akiko Ishii"، مادر کور دختر کوچک خود، در این کاربرد پیدا کرد. او می‌تواند توضیحات زنده در مورد فعالیت‌های دخترش (مانند "خواندن کتاب، با بازی‌ها بازی کردن") را از طریق ویدیو دریافت کند و با کمک برنامه می‌تواند به شنیدن متن برای خواندن همراه با دخترش بپردازد. یک مثال دیگر این است که چگونه هوش مصنوعی به کیفیت زبان علامتی کمک می‌کند (Jiang et al. ، 2020؛ Michaud et al. ، 2000؛ Parton ، 2006) ، مانند ساخت هوش مصنوعی که حرکات دستی را به عبارت‌های صوتی برای شریک صحبتی‌ها ترجمه کند که زبان علامت را نمی‌فهمند.

برای افراد نورودیورجنت، همچنین، هوش مصنوعی پتانسیل بزرگی دارد، به عنوان مثال، برای پشتیبانی تصمیم گیری بالینی (Pedersen و همکاران، 2020)، برای کمک به ارزیابی و یادگیری افراد مبتلا به اوتیسم برای برقراری تعامل بهتر با دیگر افراد (Al Banna و همکاران، 2020؛ Anagnostopoulou و همکاران، 2020؛ Wall و همکاران، 2012)، اما ما نیز باید از کورسورهای اخلاقی در زمان «اتوماسیون اوتیسم» (Keyes، 2020) آگاه باشیم – ترجمه واقعیت های اجتماعی به مدل های محاسباتی می تواند بعضی جنبه ها را از دست بدهد. به همین ترتیب، برای اطمینان از آنکه فناوری اکسواسکلتون برای همه کار می کند، به طور بحرانی بررسی شود، به عنوان مثال، همچنین برای کاربرانی که برای فناوری موجودی بسیار سنگین یا بلند قرار گرفتند (Lajeunesse و همکاران، 2016؛ Søraa و Fosch-Villaronga، 2020). برای تامل در اینکه دنیا به گونه ای طراحی شده است که فرض می کند همه بدنی های «کاملاً عالی» دارند، به صفحات وب فکر کنید، و چگونه بسیاری از آنها به گونه ای طراحی می شوند که افراد با نقص بینایی یا ناتوانی رنگی ممکن است با مشکلاتی در دریافت اطلاعات دقیق مواجه شوند.

بدن با چندین سیستم هوش مصنوعی روبروست – چگونگی پیری، فرض بر بدن سالم و قابل انجام، و سنجش سلامتی ما توسط فناوری. و آیا این پیری در واسطه با هوش مصنوعی، صرفاً برای ثروتمندان انحصاری خواهد بود، تضمین کننده وجود جاودان برای کسی با منابع پرداخت آن باشد؟ همانطور که در فصل بعد خواهیم دید، هوش مصنوعی از پیش از زندگی، بر اینکه کاربران ثروتمند چقدر ثروتمندند، تبعیض می کند و شخصی را که به سرمایه دسترسی ندارد، از خارج می کند. همچنین در زمینه سلامتی: "برنامه های کاربردی هوش مصنوعی بسیاری که در زمینه های کشورهای پردرآمد استفاده شده است، استفاده آن در زمینه های با منابع کمتر هنوز نسبتاً نوپاست" (Wahl و همکاران، 2018). با این حال، نمونه هایی از هوش مصنوعی نیز وجود دارد که بیشتر به کشورها و افراد با درآمد کم، به عنوان مثال GraceHealth، یک برنامه بهداشتی باروری است که از هوش مصنوعی، فناوری هوشمند و پیش بینی های مبتنی بر علم استفاده می کند (Grace.health، بی تاریخ).

ناتوانی شکل دهنده دنیای ما است، به ویژه چگونه در مورد آن صحبت می شود، چگونه قانون‌گذاری می شود و از قدرت محروم می شود. تبعیض نسبت به افراد ناتوان - قابلیت بینی - "تصویرسازی های دیجیتال و فن آوری ما را شکل می دهد - مفاهیم از کسانی که از توسعه [هوش مصنوعی که به طور] طراحی و پیاده سازی شده است بهره خواهند برد ، ... چگونه عملکرد 'مناسب' بدن و ذهن را تصور می کنیم "(شیو، 2020، ص 41). قابلیت تکنولوژیک یک فرد ناتوان نزدیک تغییر رفتار فناوری است که در جامعه فنی-اجتماعی منتشر شده است را دارد.

**مرگ با هوش مصنوعی- بهشت جایی روی زمین است ؟**

در این فصل، در نهایت، می‌خواهم شما را با یک آزمایش فکری به قباله‌ی "آینه سیاه" (فصل 3، قسمت 4) مجموعه تلویزیونی انگلیسی "بلک میرور" توجه دهم. این مجموعه به بررسی مشکلات فناوری می پردازد – چه اتفاقی می افتد وقتی به یک صفحه سیاه نگریست و از طریق یک "آینه سیاه"، کاستی ها و ترس های خود را بازمی بینید؟ بیشتر از قسمت ها تاریک هستند، اما در حالی که این قسمت بیشترین تمرکز خود را بر مرگ و میر قرار داده است، به یاد ماندنی و باعث شادی است. در این قسمت، یورکی، جوانی خجالتی، عاشق و شهر ساحلی بسیار زیبای دهه 80 می‌شود که با کلی، زنی جذاب و روشنگری آشنا می‌شود و آن را بررسی می‌کند. با این حال، تلف شدن و ارتباط با دنیای شاد مجازی آی‌پی "سان جونیپرو" را از طریق سیستم واقعیت مجازی مبتنی بر هوش مصنوعی، دو زن بالغ و میانسال را به هم وصل می‌کند. این قسمت به چندین بحث دانشگاهی منجر شده است (کوک، 2019؛ دراگ، 2018؛ گالاتی، 2021) و در شکل 5.2 نمایش داده شده است.

شکل 5.2 دیدگاه سان جونیپرو در مورد مرگ و هوش مصنوعی.

نتخاب برای مردن، و آپلود ضمیر خودتان برای زندگی ادامه‌یافته در "محل خوشبختی"، به شدت موازی با مفهوم " بهشت " و " نیروانا" است. همانند بیشتر پرسشهای فکری سیاه آینه، داستان نزدیکتر از آنچه فکر می‌کنیم به واقعیت است. در ژاپن، افرادی که به دلیل انزوای اجتماعی شدید، که به آنها "هیکیکوموری" گفته می‌شود، همراه با افراد انزوای اجتماعی در سایر کشورها، به دنیای مجازی اینترنت برای بازی و ارتباطات اجتماعی خود در بازی‌های آنلاین چند نفره نقش‌آفرینی (MMORPGS) می‌روند، به جای دنیای دیجیتالی. برخی میلیاردرها به کریوژنیک کردن بدن‌های خود می‌پردازند تا بخواهند از مرگ دوری کنند و درصدد بیدار شدن و زندگی دائمی با تکنولوژیی پیشرفته‌ای که امکان زندگی جاودانه برای آنها ممکن شود، بیدار شوند. هوش مصنوعی هم به مرگخواب پایدار، بازی‌های آنلاین چند نفره نقش‌آفرینی، و پایداری دیجیتالی در سان خوزه‌پیرو کمک می‌کند. اجتماعی، حتی مرگ، دیجیتالی شده‌اند.

نمایش داستانی از مراقبت برای اتصال به هوش مصنوعی ما را با چندین سوال رها می کند: چه کسی انتخاب می کند که به چه نوع جاودانگی بروید؟ آیا سیستم های هوش مصنوعی می‌توانند اشتباهاتی را انتخاب کنند، به طور مثال، فکر کنند که شما از سگ‌ها خوشتان می‌آید، اما به صورت مخفیانه از آنها ترس دارید؟ به شما اجازه داده می‌شود که در صورت خستگی از وجود دیجیتالی، آن را خاموش کنید؟

**خلاصه فصل**

در این فصل، دیدیم که چگونه سیستم‌های هوش مصنوعی با بدن‌های ما برخورد می‌کنند و به سه موضوع مهم سلامتی، سن و اعاقه تمرکز می‌کنند. دیدیم که چگونه بزرگسالان می‌توانند از فناوری پایدار سالمندی برای بهتر کردن زندگی خود استفاده کنند. علاوه بر این، تأثیرات ویروس کووید-19 و تأثیرات آن بر هوش مصنوعی معرفی شد و دیدیم که زمینه سلامتی بازار بزرگی برای هوش مصنوعی است، اما هنوز مشکلات چندگانه‌ای برای پیاده‌سازی هوش مصنوعی سلامتی وجود دارد که برای تمام افراد خوب کار کند. برای افراد دارای معلولیت، دیدیم که هوش مصنوعی می‌تواند بسیار مفید باشد، به عنوان مثال هنگام کمک به افراد نابینا یا ناشنوا برای ارتباط، اما هوش مصنوعی اغلب فرض می‌کند که افراد دارای توانایی بالا هستند، و آگاهی افراد با معلولیت لازم است. تجربه زندگی افراد معلول بسیار مهم است، اما بسیار کم ارزش‌سنجی می‌شود. با این حال، این فرصت را از دست می‌دهیم که افراد تحت تأثیر را به عنوان متخصصان زندگی و تجربیات خودشان درگیر کنیم. در مطابق با فکر هاراوایی، در حالی که بسیاری از پروژه‌های هوش مصنوعی به دنبال کارهای خوب هستند، با شنیدن و همکاری با افراد تحت تأثیر، می‌توانیم جنبشی برای همکاری با افراد تحت تأثیر و تجربیات آنها در زندگی شکل دهیم.

**یادداشت**

فراموشی عمومی برای ناتوانی در به یادآوری، تفکر یا تصمیم‌گیری‌هایی است که با انجام فعالیت‌های روزمره مزاحم می‌شود. بیماری آلزایمر نوع شایعی از فراموشی عمومی است. اگرچه فراموشی عمومی اغلب بر سر بزرگسالان تأثیر می‌گذارد، اما بخشی از پیری طبیعی نیست. (مرکز کنترل و پیشگیری از بیماری‌ها (۲۰۱۹).)

**مرجع ها**

Abd-Alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, *132*, 103978.

Al Banna, M., Ghosh, T., Taher, K. A., Kaiser, M. S., & Mahmud, M. (2020). A monitoring system for patients of autism spectrum disorder using artifi- cial intelligence. *International Conference on Brain Informatics*. 19 September 2020, Padua, Italy.

Anagnostopoulou, P., Alexandropoulou, V., Lorentzou, G., Lykothanasi, A., Ntaountaki, P., & Drigas, A. (2020). Artificial intelligence in autism assess- ment. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(6), 95–107.

Benjamin, R. (2019). Assessing risk, automating racism. *Science*, *366*(6464), 421–422.

Bourla, D. H., Hubschman, J. P., Culjat, M., Tsirbas, A., Gupta, A., & Schwartz, S.

D. (2008). Feasibility study of intraocular robotic surgery with the da Vinci surgical system. *Retina*, *28*(1), 154–158.

Briand, A., Almeida, H., & Meurs, M.-J. (2018). Analysis of social media posts for early detection of mental health conditions. *The 31st Canadian Conference on Artificial Intelligence took place in Toronto*. 8–11 May 2018,York University, Ontario.

Brusselaers, N., Steadson, D., Bjorklund, K., Breland, S., Stilhoff Sörensen, J., Ewing, A., Bergmann, S., & Steineck, G. (2022). Evaluation of science advice during the COVID-19 pandemic in Sweden. *Humanities and Social Sciences Communications*, *9*(1), 1–17.

Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A., Choudhury, T., Hauser, M., Kane, J. M., & Ben-Zeev, D. (2019). Relationships between smart- phone social behavior and relapse in schizophrenia: A preliminary report. *Schizophrenia Research*, *208*, 167–172.

Calvo, R. A., Milne, D. N., Hussain, M. S., & Christensen, H. (2017). Natural lan- guage processing in mental health applications using non-clinical texts. *Natural Language Engineering*, *23*(5), 649–685.

CDC. (2019). *What is dementia?* Centers for Disease Control and Prevention. Retrieved May 28, 2022 from [https://www.cdc.gov/aging/dementia/index.](https://www.cdc.gov/) [html#:~:text=Dementia%20is%20not%20a%20specific,a%20part%20](https://www.cdc.gov/) [of%20normal%20aging](https://www.cdc.gov/).

Chen, I.Y., Szolovits, P., & Ghassemi, M. (2019). Can AI help reduce disparities in general medical and mental health care? *AMA Journal of Ethics*, *21*(2), 167–179.

Chou, K. (2022). *Making data useful for public health*. Retrieved March 19, 2022 from [https://blog.google/technology/health/making-data-useful-public-health/](https://blog.google/)

Cook, J. (2019). San Junipero and the digital afterlife: Could heaven be a place on Earth? In David Kyle Johnson (Ed.), *Black Mirror and Philosophy: Dark Reflections*, 109–117. John Wiley & Sons, Inc.

Creswell, C., Shum, A., Pearcey, S., Skripkauskaite, S., Patalay, P., & Waite, P. (2021). Young people’s mental health during the COVID-19 pandemic. *The Lancet Child & Adolescent Health*, *5*(8), 535–537.

D’Alfonso, S. (2020).AI in mental health. *Current Opinion in Psychology*, *36*, 112–117. Data-Core-Healthcare. (2022). *Success stories of using artificial intelligence in healthcare*. Retrieved May 28,2022 from [https://datacorehealthcare.com/success-stories-](https://datacorehealthcare.com/)

[of-using-artificial-intelligence-in-healthcare/](https://datacorehealthcare.com/)

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. *Proceedings of the SIGCHI con- ference on human factors in computing systems*. 27 April–2 May 2013, Paris, France.

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., Penke, L., Rafnsson, S. B., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135–152.

Difrancesco, S., Lamers, F., Riese, H., Merikangas, K. R., Beekman, A. T., van Hemert, A. M., Schoevers, R. A., & Penninx, B. W. (2019). Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disor- ders: A 2-week ambulatory assessment study. *Depression and Anxiety*, *36*(10), 975–986.

DiMaio, S., Hanuschik, M., & Kreaden, U. (2011). The da Vinci surgical system. In Jacob Rosen, Blake Hannaford, & Richard M. Satava (Eds.), *Surgical robotics* (pp. 199–217). Springer.

Drage, E. (2018). Utopia, race, and gender in Black Mirror’s “San Junipero”. In Barry Vacker, & Angela M. Cirucci (Eds.), *Black mirror and critical media theory*, (p. 27). Lexington Books.

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoti̧uc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, *115*(44), 11203–11208.

Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research*, *21*(5), e13216.

Galati, G. (2021). San Junipero. In German A. Duarte, & Justin Michael Battin (Eds.), *Reading Black Mirror* (pp. 291–308). Transcript-Verlag.

Grace.health. n.d. Reshaping health for the next billion women online. *The first women’s health platform designed for emerging markets*. <https://www.grace.health/>

Granquist, C., Sun, S. Y., Montezuma, S. R., Tran, T. M., Gage, R., & Legge, G. E. (2021). Evaluation and comparison of artificial intelligence vision aids: Orcam MyEye 1 and seeing AI. *Journal of Visual Impairment & Blindness*, *115*(4), 277–285.

Hamraie, A., & Fritsch, K. (2019). Crip technoscience manifesto. *Catalyst: Feminism, Theory,Technoscience*, *5*(1), 1–33.

Jiang, X., Satapathy, S. C., Yang, L., Wang, S.-H., & Zhang, Y.-D. (2020). A survey on artificial intelligence in Chinese sign language recognition. *Arabian Journal for Science and Engineering*, *45*(12), 9859–9894.

Jiang,Y., Edwards, A. V., & Newstead, G. M. (2021). Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology*, *298*(1), 38–46.

Just, M. A., Pan, L., Cherkassky, V. L., McMakin, D. L., Cha, C., Nock, M. K., & Brent, D. (2017). Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nature Human Behaviour*, *1*(12), 911–919.

Keyes, O. (2020). Automating autism: Disability, discourse, and artificial intel- ligence. *The Journal of Sociotechnical Critique*, *1*(1), 8.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., & Group, N.Y. P. S.

A. (2019). Can your phone be your therapist? Young people’s ethical perspec- tives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical Informatics Insights*, *11*, 1178222619829083.

Lajeunesse, V., Vincent, C., Routhier, F., Careau, E., & Michaud, F. (2016). Exoskeletons’ design and usefulness evidence according to a systematic review of lower limb exoskeletons used for functional mobility by people with spinal cord injury. *Disability and Rehabilitation: Assistive Technology*, *11*(7), 535–547.

Maeso, S., Reza, M., Mayol, J. A., Blasco, J. A., Guerra, M., Andradas, E., & Plana,

M. N. (2010). Efficacy of the Da Vinci surgical system in abdominal surgery compared with that of laparoscopy: a systematic review and meta-analysis. *Annals of Surgery*, *252*(2), 254–262.

Mastoras, R.-E., Iakovakis, D., Hadjidimitriou, S., Charisis, V., Kassie, S., Alsaadi, T., Khandoker, A., & Hadjileontiadis, L. J. (2019).Touchscreen typing pattern anal- ysis for remote detection of the depressive tendency. *Scientific Reports*, *9*(1), 1–12. Michaud, L. N., McCoy, K. F., & Pennington, C. A. (2000). An intelligent tutor- ing system for deaf learners of written English. *The 4th ACM SIGCAPH Conference*

*on Assistive Technologies on assistive technologies*. 13–15 November 2000, Arlington Virginia USA.

Morales, M., Dey, P., Theisen, T., Belitz, D., & Chernova, N. (2019). An investiga- tion of deep learning systems for suicide risk assessment. *Proceedings of the sixth workshop on computational linguistics and clinical psychology*. Minneapolis, Minnesota.

Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E., & Becker, T. (2011). Feeling bad on Facebook: Depression disclosures by col- lege students on a social networking site. *Depression and Anxiety*, *28*(6), 447–455.

NAMI. (2020). *Mental health by the numbers*. Retrieved May 28, 2022 from [https://](https://www.nami.org/) [www.nami.org/mhstats](https://www.nami.org/)

Östlund, B., Olander, E., Jonsson, O., & Frennert, S. (2015). STS-inspired design to meet the challenges of modern aging. Welfare technology as a tool to pro- mote user driven innovations or another way to keep older users hostage? *Technological Forecasting and Social Change*, *93*, 82–90.

Our-world-in-data. (2019). *Life expectancy, 2019*. Retrieved March 19, 2022 from [https://ourworldindata.org/grapher/life-expectancy-at-birth-total-years](https://ourworldindata.org/)

Parton, B. S. (2006). Sign language recognition and translation: A multidisci- plined approach from the field of artificial intelligence. *Journal of Deaf Studies and Deaf Education*, *11*(1), 94–101.

Pedersen, M., Verspoor, K., Jenkinson, M., Law, M., Abbott, D. F., & Jackson, G. D. (2020). Artificial intelligence for clinical decision support in neurology. *Brain Communications*, *2*(2), fcaa096.

Reeves, W. C., Pratt, L. A., Thompson, W., Ahluwalia, I. B., Dhingra, S. S., McKnight-Eily, L. R., Harrison, L., D’Angelo, D. V., Williams, L., & Morrow, B. (2011). Mental illness surveillance among adults in the United States. *MMWR. Morbidity and Mortality Weekly Report*. [https://stacks.cdc.gov/view/cdc/31627](https://stacks.cdc.gov/)

Rickwood, D. J., Deane, F. P., & Wilson, C. J. (2007). When and how do young people seek professional help for mental health problems? *Medical Journal of Australia*, *187*(S7), S35–S39.

Ross, C., & Ike, S. (2018). *IBM’sWatson supercomputer recommended‘unsafe and incorrect’cancer treatments, internal documents show*. Statnews.com. Retrieved January 27, 2022 from [https://www.statnews.com/2018/07/25/ibm-watson-recommended-](https://www.statnews.com/) [unsafe-incorrect-treatments/](https://www.statnews.com/)

Salthouse, T. A. (2009). When does age-related cognitive decline begin?

*Neurobiology of Aging*, *30*(4), 507–514.

Schwartz, H. A., Eichstaedt, J., Kern, M., Park, G., Sap, M., Stillwell, D., Kosinski, M., & Ungar, L. (2014). Towards assessing changes in degree of depression through Facebook. *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality* (pp. 118–125). June, Baltimore, Maryland, USA.

Scott, J., Grierson, A., Gehue, L., Kallestad, H., MacMillan, I., & Hickie, I. (2019). Can consumer grade activity devices replace research grade acti- watches in youth mental health settings? *Sleep and Biological Rhythms*, *17*(2), 223–232.

Shen, J. H., & Rudzicz, F. (2017). Detecting anxiety through reddit. *Proceedings of the fourth workshop on computational linguistics and clinical psychology—From linguistic signal to clinical reality*.

Shew, A. (2020). Ableism, technoableism, and future AI. *IEEE Technology and Society Magazine*, *39*(1), 40–85.

Sommers, J. (2021). *The pastry A.I. that learned to fight Cancer*. Retrieved January 25, 2022 from [https://www.newyorker.com/tech/annals-of-technology/](https://www.newyorker.com/) [the-pastry-ai-that-learned-to-fight-cancer](https://www.newyorker.com/)

Søraa, R. A., & Fosch-Villaronga, E. (2020). Exoskeletons for all: The interplay between exoskeletons, inclusion, gender, and intersectionality. *Paladyn, Journal of Behavioral Robotics*, *11*(1), 217–227.

Søraa, R. A., Manzi, F., Kharas, M. W., Marchetti, A., Massaro, D., Riva, G., & Serrano, J. A. (2020). Othering and deprioritizing older adults’ lives: Ageist discourses during the COVID-19 pandemic. *Europe’s Journal of Psychology*, *16*(4), 532.

Søraa, R. A., Nyvoll, P., Tøndel, G., Fosch-Villaronga, E., & Serrano, A. (2021). The social dimension of domesticating technology: Interactions between older adults, caregivers, and robots in the home [Vitenskapelig artikkel]. *Technological Forecasting and Social Change*, 167, 1–13. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.techfore.2021.120678) [techfore.2021.120678](https://doi.org/10.1016/j.techfore.2021.120678)

Strickland, E. (2019). IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care. *IEEE Spectrum*, *56*(4), 24–31.

Tezuka, K. (2020). *Seeing AI empowers people who are blind or with low vision for everyday life*. Retrieved April 16, 2022 from [https://news.microsoft.com/apac/2020/](https://news.microsoft.com/) [12/03/seeing-ai-empowers-people-who-are-blind-or-with-low-vision-](https://news.microsoft.com/) [for-everyday-life/#:~:text=Power%20of%20artificial%20intelligence](https://news.microsoft.com/), nearby%20people%2C%20text%20and%20objects.

Vogel, D. L., Wester, S. R., Hammer, J. H., & Downing-Matibag, T. M. (2014). Referring men to seek help: The influence of gender role conflict and stigma. *Psychology of Men & Masculinity*, *15*(1), 60.

Wachter, R. (2015).The digital doctor. *Hope, Hype and at the Dawn of Medicines Computer Age*, 2015.

Wahl, B., Cossy-Gantner, A., Germann, S., & Schwalbe, N. R. (2018). Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Global Health*, *3*(4), e000798.

Wall, D. P., Dally, R., Luyster, R., Jung, J.-Y., & DeLuca, T. F. (2012). Use of arti- ficial intelligence to shorten the behavioral diagnosis of autism. *PLoS ONE*, 7, e43855.

Weizenbaum, J. (1966). ELIZA—A computer program for the study of natu- ral language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., Morris, M. R., Rankin, J., Rogers, E., & Salas, M. (2019). *Disability, bias, and AI*. AI Now Institute. [https://www.ahrq.gov/downloads/pub/evidence/pdf/](https://www.ahrq.gov/) [alzheimers/alzcog.pdf](https://www.ahrq.gov/)

Williams, J. W., Plassman, B. L., Burke, J., & Benjamin, S. (2010). Preventing Alzheimer’s disease and cognitive decline. *Evidence Report/Technology Assessment*, *193*, 1–727.

World-Population-Review. (2022). *Suicide rate by country 2022*. Retrieved May 28 2022 from [https://worldpopulationreview.com/country-rankings/suicide-](https://worldpopulationreview.com/) [rate-by-country](https://worldpopulationreview.com/)

Yu, M., Cai, T., Huang, X., Wong, K., Volpi, J., Wang, J. Z., & Wong, S. T. (2020). Toward rapid stroke diagnosis with multimodal deep learning. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 4–8 October 2020, Lima, Peru.

Ziwei, B. Y., & Chua, H. N. (2019). An application for classifying depression in tweets. *Proceedings of the 2nd international conference on computing and big data*. 18–20 October 2019, Taiwan.

**6**

هوش مصنوعی و کلاس

**کار ، آموزش و پایداری**

منابع نامتناسب بین گروه‌های اجتماعی توزیع می‌شوند. خطرات هوش مصنوعی برای بهره‌برداری از ثروتمندان و اعضای ضعیف و کم‌درآمد جامعه درگیر است. در عین حال، سامانه هوش مصنوعی به مواد خام برای استخراج و منطق کاری نیاز دارد که گروه‌های فقیر معمولاً در آن فعالیت می‌کنند (Crawford، ۲۰۲۱). در حالی که ثروتمندان به نوعی فناوری دسترسی دارند، فقیران به احتمال بیشتری در زندگی خود از فناوری استفاده می‌کنند، بدون این که بتوانند در نوع و چگونگی پیاده‌سازی آن دخالت داشته باشند. بخصوص در سامانه‌های هوش مصنوعی که با نظارت سر و کار دارند، عدم برابری می‌تواند ثروتمندان را به عنوان نظاره‌گر با فناوری و فقیران را به عنوان افراد زیر نظر و نظاره شده قرار دهد. در این فصل، "طبقه" را به عنوان پدیده اجتماعی-فرهنگی مورد بررسی قرار می‌دهیم (که شامل میلهای گوناگون، از جمله هم‌افزایی با جنسیت، نژاد، سنّ و غیره است)، اما به بخش سوسیو اقتصادی این گروه در محدوده‌های فردی و جهانی تمرکز خواهم کرد. ثروت، فقر و تقسیم‌بندی درآمد بزرگترین مکانیزم‌های برجای گذاشتن غیر مطلوب در جهان هستند. برای همین، باید در نظر داشت که وقتی هوش مصنوعی وارد بخش اقتصادی وجود بشریت می‌شود، چه چیزی به دست می‌آید و کی در بدهی باقی می‌ماند.

**کامپیوتر می گوید "بدون پول، نمی‌شود رفت."**

سامانه‌های هوش مصنوعی در شناسایی و بازتولید الگوها و ساختارها و دسته‌بندی جهان از طریق این فریم‌ها خوب عمل می‌کنند، اما در تصمیم‌گیری‌های اخلاقی (اگر باشد که یک موضوعی است که فلاسفه با آن مبارزه می‌کنند) کمتر مهارت دارند. اگر مجموعه دموگرافیک ۱۰۰ نفری وجود داشته باشد و ۱۰ نفر آنان ثروتمند باشند و ۵۰ نفر فقیر باشند، این فقط داده و اطلاعاتی است، اما چگونگی برخورد الگوریتم‌ها با این داده‌ها، جایی است که مشکلات شروع می‌شود. برای روشن شدن، یک مثال می‌تواند سامانه‌های هوش مصنوعی باشد که برای پیدا کردن افرادی که حق گرفتن تولیدات مالیاتی (قرض الحسنه يا بيمه) را دارند، استفاده می‌شوند و تنها معیار شرایط آن فرد، سطح درآمدش است و برنامه آن‌ها به آن‌ها اجازه تحلیل کامل مشکلات شخص نمی‌دهد. سامانه‌های هوش مصنوعی می‌تواند فقر را تقویت کند، باعث ورشکستگی و دائمی بودن اجاره مسکن شود یا به دلایل عدم تأیید قرض الحسنه خانه باش تبدیل به بی‌خانمان شود. بودن بی‌خانمان نیز مجموعه‌ای از تبعیضات تکنیکی و اجتماعی برای آنها دارد، به عنوان مثال، طراحی "اشیاء بی‌رحم" برای محیط‌های شهری می‌تواند وجود داشته باشد که شامل جاهای نشستنی غیرمساعد برای خوابیدن بی‌خانمان‌ها باشد (Rosenberger، ۲۰۱۷).

چه کارهایی می‌توان انجام داد؟ یک مثال از مداخله در سامانه در بانک‌های نروژ قابل مشاهده است، جایی که محدودیت‌های دولتی ۱۵٪ از ارزش مالکیت خود در قبال قیمت یک ملک را برای اعطای وام مسکن شروع می‌شود. با این حال، بانک‌ها مجاز هستند که از این قانون سخت در ۱۰٪ از موارد انحراف کنند و تصمیمات براساس مواردی مانند عدم توافق با سیستم سخت (مانند اربابان درآمد بالا که هنوز به اندازه کافی پس‌انداز ندارند اما احتمالاً به زودی براي پس‌انداز کافی خواهند داشت) داده می‌شود. در حالی که سیستم سخت با یک میزان بالایی مشخص و سخت است، عامل انسانی به کارمندان، امکان می‌دهد که با در نظر گرفتن موارد بشری، تصمیمات انسانی بگیرند و مثلاً بر این عقیده باشند که شانس اینکه یک فقیر بتواند جزئیات پرداخت وام را در نظر بگیرد، از پیش‌ پرداخت آن بالاتر است. برخی جنبه‌های اقتصادی به طور قابل توجهی به طور خودکاراتیزه شده‌اند. به عنوان مثال، برای محاسبه مالیات، احتمالاً تصوّر کردید چندین حسابدار انسانی را با کارت‌ها و اسناد مالی مشغول به محاسبه میزان مالیات در دفاتر خود می‌بینید، اما این ضروری نیست. در کشورهای اسکندیناوی به همراه بسیاری دیگر، مالیات‌ها به طور عمدی با استفاده از روش‌های دیجیتالی کنترل می‌شوند.

بزرگ‌ترین مشکل با ارزیابی بر اساس ارزش، این است که ارزش فردی بر اساس ارزش پولی یا در برخی مناطق، امتیازات اعتباری اجتماعی آنها محاسبه می‌شود که میزان عملکرد آنها به عنوان شهروندان بر اساس هر معیاری که دولت معین می‌کند، را اندازه‌گیری می‌کند. برخی جنبه‌های زندگی‌مان، مانند درآمد، از دیگر عوامل بیشتری به عنوان متغیرهای داده به کار می‌رود. برخی از آنها ممکن است غیرقانونی باشد که به کامپیوترها تغذیه شوند، در حالی که برخی دیگر به طور داوطلبانه ارائه می‌شوند، به عنوان مثال تصاویری که از طریق شبکه‌های اجتماعی بارگذاری می‌کنیم. برای افرادی که قابل قبول نیستند، تهدیدی به حقوق و امتیازات مانند مسکن، وام و مزایا وجود دارد. در صورت کاهش ارزش شما به عنوان مشتری یا شهروند بر اساس دین، جنسیت، جنسیت، نژاد یا سن، به عنوان نمونه‌ای، بسیار مشکل است و ما باید اقداماتی را انجام دهیم تا جلوی تولید و اختراع چنین تعصباتی در هوش مصنوعی را بگیریم. اگر هوش مصنوعی یک هم‌بستری منفی بین افراد از نوع خاص و ویژگی های نامطلوب "کشف" کند، می‌توانیم ببینیم که به عنوان مثال سیستم حتی قصد ارزیابی بر اساس جنسیت را نداشته اما در نهایت این کار را می‌کند. با این حال، راه کارهایی وجود دارد، مانند پروژه ریکر 2، یک پروژه سوئدی به منظور اطمینان از آنکه الگوریتم های یادگیری جنسیتی مساوی در تصمیمات مالی استفاده شود و یا پلتفرم هوش مصنوعی "الیس" که کارآفرینان زن و اقلیت‌ها را به هم متصل می‌کند تا کسب و کار را به بالا ببرند (Ramboll، 2022، ص 15).

**کنترل از بالا به پایین در محل کار**

هوش مصنوعی به طور فزاینده ای در بخش استخدام استفاده می‌شود تا کارگران فرد را مدیریت و کنترل کند (Guenole & Feinzig، 2018؛ Kang et al.، 2020). برای مثال، People (2020) متوجه شدند که 24٪ از شرکت‌ها برای انجام رویدادهای استخدام از هوش مصنوعی استفاده می‌کنند. این می‌تواند به عنوان مثال از طریق تحلیل متن برنامه‌های استخدام (Dastin، 2018؛ Nimbekar et al.، 2019)، آموزش کارکنان (Ku، 2021) و نظارت بر رفتارهای نامطلوب کارگران (Eisenstadt، 2021) انجام شود. مطالعهیی توسط van den Broek و همکاران (2021) روی سازندگان هوش مصنوعی برای استفاده در استخدام، نشان داد که فناوران، اگرچه مراقب تعصب بودند، آن را به عنوان یک پدیده انسانی که وقتی تصمیمات براساس فرضیات ذهنی یا واکنش های عاطفی انسانها اتخاذ می‌شود، حساب می‌کردند و حضر فرایند استخدام را "مسموم" می‌کند. بررسیهای اجتماعی- فرهنگی نحوه تجربه فناوری و هوش مصنوعی توسط کارکنان (Bhargava et al، 2021؛ Li et al، 2019؛ Bhargava et al، 2021؛ Li et al، 2019؛ Smith، 2018) و با سازندگان این فناوری (Tomé et al، 2020؛ van den Broek et al، 2021) انجام شده است، با اکثریتی به سمت تعصب اشاره داده شده است. اگر این تعصبات غیرمجاز باقی بمانند، می‌توانند برای افراد محروم در روابط نامتقارن قدرت با محیط کارشان اثرات عمده‌ای داشته باشند.

چندین شرکت بزرگ اکنون از ابزارهای نظارتی هوش مصنوعی برای نظارت و "اصلاح" رفتار کارکنان خود استفاده می‌کنند. به عنوان مثال، برای مقابله با کمپین‌های اتحادیه‌ای که به عنوان یک راه برای چالش دادن کارکردن دقیق کارگران در Amazon شروع شده، به طوری که هر ثانیه کاری انجام شده توسط یک کارگر را ردیابی می‌کنند، این شرکت در حال حاضر در حال جنگیدن است. استراحت دست به فهرست زمانی که کارکران نیاز به استراحت دارند، و همچنین جلسات اجباری با ناظران از فهرست زمان‌های مشمول "زمان کار" خارج می‌شوند. (Delfanti، 2021؛ Kantor et al.، 2021). این مدیریت دقیق کارگران با درآمد کم، نوآور نیست، اما با یک سری روش های کاری تیلوری که توسط یک مهندس رفتاری آمریکایی به نام فردریک وینسلو تیلور در اوایل قرن بیستم معروف شده است، پیگیری می‌شود. تیلور ، که به فعالیت در برابر اتحادیه ها مشغول بود، باور داشت که کارگران به طبیعت خود تنبل هستند، اما با طراحی نظارت دقیق بر هر حرکت در روز کاری شان و با بازگشت مناسب، آنها را به بهبود کارکرد تشویق می‌کرد(Littler، 1978).

در انبارهایی که توسط هوش مصنوعی مدیریت می‌شوند، ما با خاستگاهی از تیلوریسم روبرو می‌شویم تنها با این تفاوت که فقط نظارت دقیق تر و رتبه بندی بیشتر کارکنان با استفاده از فن آوری های هوش مصنوعی امکان پذیر شده است. این نوع مدیریت نشان می‌دهد که به مدت طولانی، کارگران به عنوان بخشی از ماشین ها، موجودیت‌هایی که می‌توانند بهینه شوند، تلقی شده‌اند. با این حال، نظارت در محیط کار مختص به فعالیت‌های جسمانی نیست، به عنوان مثال "کارفرمایان می‌توانند هر حرکت کارکنان خود را ضبط کنند، در مورد گفتگوهایشان گوش دهند، جزئیات کوچک عملکرد را اندازه بگیرند و فعالیت های سازماندهی اعتراضی را شناسایی کنند"(Bales & Stone، 2020، صفحه ۱). بنابراین، بهینه سازی کارگران از طریق فناوری و با استفاده از آن می‌تواند به هر جنبه از رفتار کارگر امتداد پیدا کند، با این حال فراموشی کردن این نکته که زندگی کارگران به صورت اجتماعی قرار دارد و به هیچ وجه با ماشین ها مقایسه نمی‌شود، به دنبال خواهد داشت.(Delfanti، 2021؛ Gutelius & Theodore، 2019؛ Moradi & Levy، 2020).

به طریق مشابهی، کارکنان حمل و نقل مانند رانندگان تاکسی، رانندگان اتوبوس و کامیون، نیز با نظارت فعال روز کاری خود روبرو می‌شوند (Levy، 2016؛ Zhang et al.، 2019). با اینکه این شرکت‌ها همیشه می‌دانند که وسایل حمل و نقل شرکت در کجا هستند، رانندگان این وسایل حمل و نقل نیز به طور دقیق پیگیری می‌شوند. این می‌تواند جنبه مفیدی داشته باشد و به اطمینان از این که رانندگان مراحل استراحت اجباری خود را برای طول رانندگی کافی حاظر می‌کنند، کمک کند، اما به سامانه‌های رانندگی متمرکز بر هوش مصنوعی که در جهت پخش وظایف نهایی برای رانندگان بکار می‌رود و به طور متقارن به مفاهیم درونی و انسانی روزانه، فشار می‌آورند، نیز کمک شده است. با این حال، نباید فراموش کرد که سامانه های لجستیکی بر اساس هوش مصنوعی به برنامه‌ریزی و بهینه‌سازی انتقال افراد و کالا در شبکه‌های مادی-اجتماعی پیچیده کمک کرده است. نقل وانتقال بسته ها از کارخانه ها، جلوگیری از شلوغی ترافیک با برنامه ریزی بهینه مسیر و به همین ترتیب، مطمئن شدن از اینکه بسته ها و شما به موقع مشخص شده به مقصد خود برسید.

یکی دیگر از نمونه‌های فناوری ناتوان، خودروهای خودران هستند که اغلب به همراه هوش مصنوعی یاد می‌شوند. خودروهای خودران، به عنوان مثال، توسط کرافورد (2021) به مشکل برخوردار هستند که هنوز برای پیاده سازی منصفانه و ایمن در جامعه راه طولانی را پیش رو دارند (بدون در نظر گرفتن، اینکه آیا این آینده مورد نظر است و بحث عمومی آن برگزار شده است یا خیر). در ارتباط با تنوع درآمد، دسترسی به سیستم‌های امنیت حمل و نقل پیشرفته مختلف اهمیت دارد. سیستم‌های هوش مصنوعی که می‌توانند در مواقع اضطراری خودکاراً ترمز کنند، در صورت وقوع حوادث، مسیر را تغییر دهند یا هنگام تشخیص حیوانات در منطقه، کند شوند، در بسیاری از موارد برای ثروتمندانی که می‌توانند خودروهایی با این سیستم‌ها را تهیه کنند، محفوظ می‌مانند. مسأله کلی این است که خودروهای هوشمند باید کی را نجات دهند که در آزمایش moral machine دانشگاه MIT مشکل برخوردار شده است؛ به عنوان مثال ، پیدا شد که بیشتر افراد تمایل دارند زنان را نجات دهند تا مردان و مسافران ارزش بیشتری را به نسبت عابرین پیاده دارند. همچنین، در فرهنگ‌های مختلف تفاوت‌ها وجود دارد؛ به عنوان مثال، مردم در کشورهای شرق آسیا ترجیح می‌دهند که بزرگسالان را نجات دهند، در حالی که اروپایی‌ها و آمریکایی‌ها می‌خواهند خودروهای خودکار باید جوانان را نجات دهند. اگر خودروها باید خودکار باشند، پرسش بعدی این است که چه کسی مسئول است و آیا خودرو باید اولویت نجات مسافران یا عابرین پیاده را داشته باشد؟ این یک معضل اخلاقی است که همچنان مورد بحث است. در کل، تعهدات اخلاقی در زمینه ایمنی جاده است، اما در مورد دسترسی هم هستند: وقتی ایمنی یک عامل خودرویی امتیاز ثروتمندانه‌ای می‌شود، چگونه محرومیت افراد فقیر بر جامعه تأثیر می‌گذارد؟ و چگونه ارزش‌ها و اخلاقیات جامعه در تکنولوژی‌هایی که ما ایجاد می‌کنیم، تعبیه می‌شوند؟

این موارد کنترل محیط کار با کمک هوش مصنوعی دارای تأثیرات اجتماعی و اقتصادی هستند؛ زیرا بسیاری از شغل‌هایی که بیشترین قابلیت برای کنترل بیشتر دارند، شغل‌هایی هستند که به‌طور سنتی به عنوان "کم مهارت" و با دریافت حقوق کم توصیف می‌شوند، در حالی که افرادی که سیستم‌های فناوری کنترل را طراحی می‌کنند و تصمیم می‌گیرند چگونه اجرا کنند، دارای درآمدهای بالا هستند. از این رو، این یکی دیگر از نمونه‌های کنترل افراد با وضعیت اقتصادی بالا بر روی افراد با وضعیت اقتصادی پایین است.

**دسترسی به آموزش مبتنی بر کلاس**

هوش مصنوعی و سیستم‌های کامپیوتری توانایی بزرگی در توسعه دانش از طریق کلاس‌های دیجیتالی دارند. دوره‌های آنلاین باز و بزرگ از پیش در حال رشد بودند، اما در طول ویروس کرونا، آنها با افزایش قابل توجه مواجه شدند زیرا انتقال به کلاس‌های دیجیتالی باعث شد آموزش آنلاین به عنوان طبیعی جدیدی برای بسیاری از دانش آموزان در طول قرنطینه، ادارات خانه و موارد مشابه شود. برای برخی از مربیان، یکی از مسائل می‌تواند این باشد که دانشجویان هیچگاه دوربین‌های خود را در طول سخنرانی‌ها روشن نمی‌کنند و احساس عجیبی از صحبت کردن با صفحات سیاه دارند، در حالی که برای مناطق با پهنای باند کم، مسئله معکوس می‌تواند مشکل باشد اگر دانشجویان دوربین‌هایشان را خاموش نکنند، که موجب مشکلات اتصال پهنای باند می‌شود. اینترنت منبع بی‌نهایتی نیست و مشکلات در اشکال مختلف اجتماعی-مادی ممکن است پیش آید. اوهارا و هال (2021) می‌گویند که اینترنت می‌توان آن را به عنوان چهار اینترنت مختلف در نظر گرفت که ارزش‌های رقیب برای حکومت و پایداری دارند:

چهار اینترنت، مدعی است که به نظر می‌رسد "اینترنت"ی که به نظر یکپارچه است، در واقع توسط چهار سیستم ارزش مجزا یا چهار "اینترنت مجزا"، یعنی سیلیکون ولی اپن اینترنت، اینترنت بورژواز بروکسل، اینترنت تجاری DC و اینترنت پدرانه بیجینگ حفظ می‌شود و این سیستم‌های رقابتی برای تعیین جهت‌های آینده فراهمی اینترنت آزادی، نوآوری، امنیت و حقوق بشر هستند.

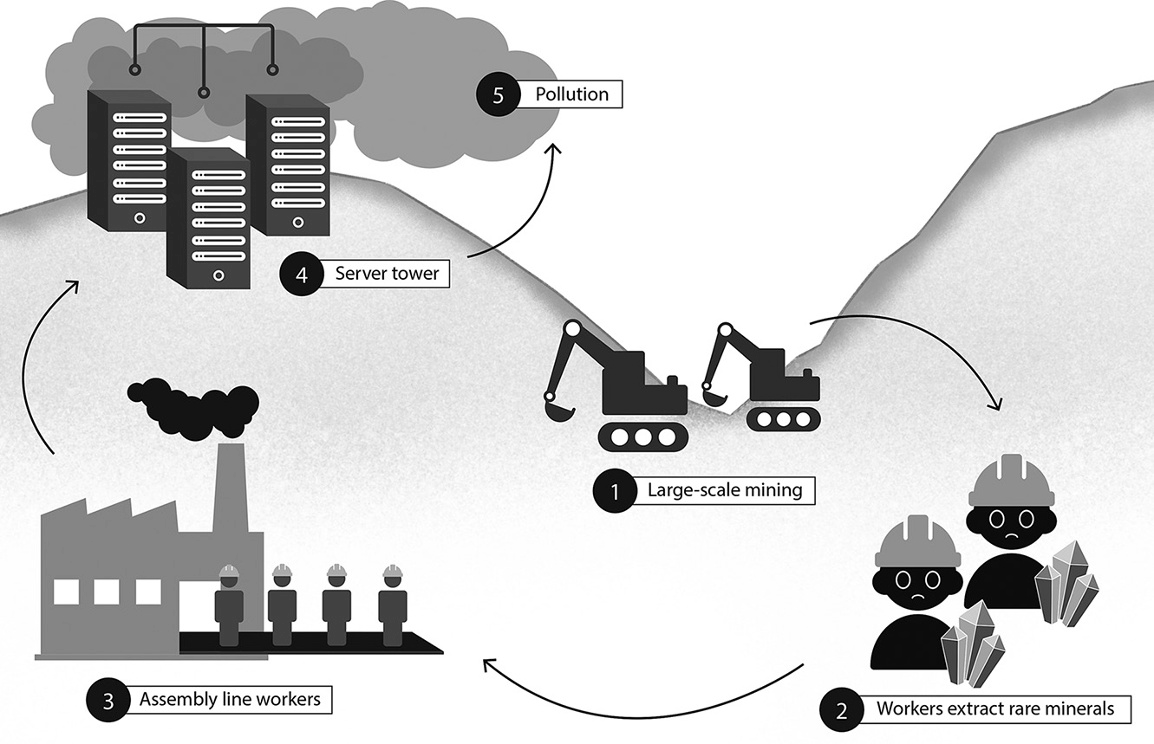
در این مورد ، آسان است که ببینیم آنچه که ما اغلب آن را به عنوان ثابت های بیشتر یا کمتر در واقعیت دیجیتال مشاهده می کنیم ، در مذاکرات بین چندین بازیگر اجتماعی دستخوش تغییر و ادماگوجا ایجاد شده است.

در این طرح فناوریک ، مهم است که به دسترسی و کیفیت آموزش نیز توجه کنیم. هوش مصنوعی های تبعیض پذیر ممکن است در سیستم آموزشی باشند ، مانند زمانی که مدارس عمومی بوستون سعی کردند از ADS برای مبارزه با تقسیمات سوسیواقتصادی براساس نژاد در سیستم های مدرسه عمومی خود استفاده کنند ، اما در واقع ، این تقسیمات را در مدارس عمومی شهر تشدید کردند (Aloisi & Gramano، 2019؛ Bales & Stone، 2020؛ De Stefano، 2019). خوشبختانه ، پروژه برای یافتن اختلاف ارزیابی شد؛ با این حال ، بیشتر بدنه های عمومی که از ADS استفاده می کنند ، منابع و / یا اولویت های لازم برای ارزیابی ندارند (Richardson، 2021). یکی از راه حل های ممکن یا تلاش برای کاهش تاثیرات مثبت و منفی هوش مصنوعی بر جامعه ، ارزیابی اجباری اثرات آنها است ، به خصوص در ابتدایی ترین دوره فناوری که پیش بینی نتیجه آن ساده نیست.

نه تنها دانش آموزان ممکن است تحت تاثیر هوش مصنوعی در آموزش قرار بگیرند بلکه سیستم های ارزیابی معلمان نیز برای بررسی کیفیت معلمان بر اساس عملکرد دانش آموزان در آزمون های استاندارد بزرگ مستقر شده اند (Richardson، 2021، صفحه 11). این گروه بندی معلمان در مدارس یا ناحیه های خاص می تواند دسته بندی نادرستی از کارایی آموزش ایجاد کند. این مشکل است ، زیرا تحقیقات نشان می دهد که یک سری تبعیض در برابر استادان زن وجود دارد (Amrein-Beardsley، 2014؛ Richardson، 2021) ، همچنین تبعیض نژادی نیز وجود دارد (Fisher et al.، 2019؛ Kaschak، 1978). اما اگر جامعه تبعیض درعلیه یک قشر داشته باشد و دستگاه در معرض خطر تکرار این تبعیض قرار داشته باشد (یا به واقع، ایجاد ایجاد مناطق ناشناخته نوع تبعیض جدید برای بشر)، چه کار می توان انجام داد؟ این یک موضوعی است که در فصل بعد بررسی خواهد شد ، اما قبل از آن، بیایید به سایر تبعیضات سوسیواقتصادی که هوش مصنوعی ممکن است به همراه داشته باشد ، بپردازیم.

**هوش مصنوعی (نا)پایدار**

زمینه تبعیض طرف AI نه تنها یک مساله "حصری" برای افراد نیست ، بلکه یک مساله برای پایداری جهان اجتماعی-مادی است. در Atlas of AI، Fisher و همکاران (2019، صفحه 322) برای درک جامع دیدگاه در مورد اینکه چگونه هوش مصنوعی روی زنجیره های ارزش جهانی مواد خام تأثیر می گذارد، فراخوانی می کنند، به عنوان مثال ، کارگرانی که مواد معدنی نادر را استخراج می کنند تا باتری های لیتیوم را تامین کنند ، شبکه های لجستیک حمل و نقل این مواد، که به عنوان مثال با کارگران ضعیف حقوق و حقوق کم انجام می شود و زنجیره های بیشتری از کارگران عامل با حقوق پایین با برنامه Mechanical Turk Amazon در حلقه ای کار می کنند که کار مشروط بر نظارت و کنترل دقیق است ، برای مثال فرمان کار، که در غیر این صورت از دستمزد ها کاسته می شود. به عبارت دیگر ، کرافورد و سایر تحقیق کنندگان سیستم AI جامع برای اینکه کی باید توقف داده شود، به شدت مطرح می کنند. آیا روش هایی وجود دارد که بتوان بدون بارگیری اکوسیستم هایمان برای ادامه تولید کامپیوتر های بیشتر ، بهتر ، سریعتر و قدرتمندتر و سخت‌افزارهای مورد نیاز آنها؛ به جلو حرکت کرد؟ چگونه می توان چرخه های استثمار تکنولوژیکی که در شکل 6.1 مثال آمده، پایدار باشد؟

تکنولوژی هرگز بی طرف نیست؛ اغلب شتاب دهنده اختلافات اجتماعی است. در حالی که فقیرها فقیرتر و ثروتمندان ثروتمند تر می شوند ، ما باید با هوش مصنوعی به عنوان ستم سیستمی برخورد کنیم - بپرسیم کدامان به تکنولوژی دسترسی دارند ، کدامان از آن به دور هستند ، کدامان نظارت و کنترل می شوند و توسط چه جهاتی - و قدرتی که با لمس تکنولوژیکی بر مشتریان ، ساکنین یا دیگر افراد دارند را سوال کنیم. همه بازماندگان وب مساوی نیستند و عناصر فقیر را همه شبکه های وبی را کشف و اندازه گیری می کنند ، بدین شکل شده که توسط سیستم های هوش مصنوعی قابل انجام باشد.

شکل 6.1 چرخه اجتماعی-مادی هوش مصنوعی.

همچنین، به چندین مسئله اخلاقی هم اشاره شده است، همانطور که اواد و همکارانش (2018) و نیهولم و اسمیدز (2016) در کتاب AI & Ethics تذکر داده اند. در حالی که فناوری AI در حال توسعه است، یکی از مهمترین سوالات این است که مالکیت AI به کی تعلق دارد و چه کسانی از آن بهره می برند؟ آیا به عنوان مثال باید ربات مالیاتی برای جایگزینی موقعیت شغلی انسان داشته باشد؟ درآمد حاصل از استفاده از ربات ها به کجا جریان پیدا می کند و چه چارچوب های نظارتی برای اطمینان از دوگانه جمله "در حالی که ثروتمندان داراتر می شوند، فقیران به فقر دست می یابند" در نظر گرفته شده است؟ برخی از محققان، سیاستمداران و اقتصاددانان درباره درآمد پایه جهانی به عنوان راه مناسب برای اطمینان از عدم ورود به طوفان کامل تخصیص ناپایدار منابع به دلیل انقلاب رباتیکی و آی پی سخن می گویند.

تفاوت های عمده دیگر در AI، تفاوت های بین توسعه، استفاده و تنظیم AI در دو حوزه عمومی و خصوصی است. در حالی که کشورها و شرکت های عمومی تا حدی بیشتر به سیاست های دموکراتیک (یا استبدادی) خود در مورد نحوه استفاده از فناوری برای شهروندان خود، و همچنین غیرشهروندان، به عنوان مثال در زمینه مهاجرت، پایبندند، شرکت های خصوصی نیز مسئولیتی نسبت به مشتریان خود، و همچنین کسانی که تحت تأثیر فناوری هستند، اما هنوز از آن استفاده نمی کنند، دارند. طراحی از بالا به پایین به ندرت برای شامل بودن بهتر است و مشارکت کاربران از طریق طراحی مشارکتی (جایی که کسانی که تحت تأثیر فناوری قرار دارند، در طراحی فناوری دخیل هستند) می تواند کمک کند تا سیستم ها در خصوص قومیت های اقلیت کمتر انحصاری شوند. در نهایت، بخش های بعدی این فصل هم از درک گسترده تر AI به عنوان چیزی که نه تنها "غربی" است صحبت می کنند، با برجسته کردن توسعه های مهم AI در آسیا و آفریقا.

**هوش مصنوعی شرق آسیا**

به طور خاص، آسیای شرقی یک نقطه داغ برای فناوری های مبتنی بر AI است، به طوری که ژاپن، چین و کره جنوبی را در بسیاری از توسعه های AI رهبری می کنند. ژاپن در حال حاضر رهبر جهان در زمینه رباتیک سلامت AI است (توشکوف، 2018) (همچنین بخش 5 این کتاب در مورد سلامت صحبت می کند). چین در یک مجموعه گسترده از سیستم های مبتنی بر AI برای زیرساخت، پشتیبانی، بازاریابی و... پیشرفت چشمگیری کرده است و در حال حاضر بیشترین مقالات تحقیقاتی در زمینه AI را منتشر می کند (27%) و به همین زودی، بیشترین پتانسیل را برای ثبت پتنت AI در دنیا دارد. چین در تشخیص سخن و بینایی تخصص دارد و با داشتن یک سیاست حمایتی از AI که با فرصت های تأمین منابع فراوان پشتیبانی می شود، این کشور در حال به جای گذاشتن علامت خود در زمینه AI جهان است (Ima et al.، 2016؛ Lundin & Eriksson، 2016؛ Okuda et al.، 2013؛ Sato et al.، 2020)، با این حال در مقابل مقررات حریم شخصی متفاوتی از کشورهای غربی دارد (Li et al.، 2019). با توجه به اینکه چین در دهه بعد بیش از 150 میلیارد دلار در پروژه های AI سرمایه گذاری می کند (Hao، 2019؛ Larson، 2018)، مشخص است که آینده AI به شدت تحت تأثیر تصمیماتی است که چین تصمیم می گیرد.

کره جنوبی دارای تمرکز خاصی بر روی فناوری های صنعتی AI نظیر خودروها، تلویزیون ها و گوشی های هوشمند است که شرکت های هیوندا، ال جی و سامسونگ رهبری می کنند. این کشور تمرکز خاصی بر روی قدرت نرم و سرگرمی دارد (مشابه استراتژی "کول جاپان" ژاپن)، به عنوان مثال با بازگشایی دیجیتال ستارگان K-pop درگذشته (Dräger & Müller-Eiselt، 2020) و به ویژه با استفاده نوآورانه از فناوری موقعیت جغرافیایی در طی همه گیری کووید-19 (Choi، 2020).

**هوش مصنوعی در آفریقا**

داده‌های بزرگ و مداخلات مبتنی بر هوش مصنوعی معمولا به عنوان نوآوری‌هایی در نظر گرفته می‌شوند که از شمال جهان برخاسته و به سایر مناطق جهان تحویل داده می‌شوند. کوکلبرگ (2020) در خواست خود برای "امثال AI از جنوب جهان / از جنوب جهان" این افکار را انتقاد کرد. اصطلاحات مانند "جنوب جهان" برای متمرکز کردن بر این نکته مفید هستند که 75٪ جمعیت جهان فقط به 20٪ درآمد جهانی دسترسی دارند و بیشتر گفتمان‌های علمی ، سیاسی و فناورانه آن‌ها را شامل نمی‌شوند. البته ، هوش مصنوعی به طور ذاتی یک چیز از شمال جهان نیست، نه در ساخت و نه در استفاده خود، بلکه فناوری دیجیتال پتانسیل بیشتری از پیشینیان آنالوگ خود دارد. اما با عدم تنوع و موقعیت‌بندی مطالعات را به ریسک قرار می‌دهیم تا فرضیاتی را بازتولید کنیم. این کار را با عدم در نظر گرفتن جنوب جهان نمی‌توانیم انجام دهیم. اگرچه بسیاری از فناوری ها و مطالعات توسط غربی ها توسعه داده شد و مورد انتقاد قرار گرفت؛ اما ما بسیاری از نمونه های مثال برای آنها از جنوب جهان داریم. تأیید می‌کنم که اصطلاحات "جنوب جهان" ، "غرب" و "کشورهای در حال توسعه" همگی ملغمات خیلی مورد بحثی هستند و این اصطلاحات بی‌طرف نیستند. بنابراین تلاش می‌کنم در انتخاب مناطق خاص به دقت بیشتری رسید.

"جنوب جهان" می تواند به هم زمان چندین چیز مختلف را نشان دهد، همانطور که سینگ و گوزمان (2021) آن را تعریف می کنند:

[1] یک محل مناسب برای مطالعه توسعه های فناورانه واقعی، [2] روشی برای درک، تجزیه و تحلیل و ساخت شیوه های محاسباتی توسعه ای، پسااستعماری و دکولونیال، [3] یک مجاز برای انواع مختلف رنج و رنج شده توسط سرمایه داری و استعمار به مقیاس جهانی و [4] تلاش برای مقاومت خلاقانه و مخالفت متمرکز با چنین رنج و رنجیدگی.

رای مطالعه‌ بیشتر درباره اصطلاحات، مطالعه کتاب سینگ و گوزمان (2021) و مقاله توشکوف (2018) به نام "جنوب جهان" اصطلاح بدی است، از آن استفاده نکنید! " را پیشنهاد می‌شود.

بعد از انتقال به شهر آکرا در غنا برای شروع اولین آزمایشگاه هوش مصنوعی گوگل در آفریقا، موستفا سیسی توصیف می‌کند که "عدم تنوع می‌تواند باعث تثبیت تبعیض های روباتیکی غیر قصدی شده و سایر بخش های AI شود." سیسه که تحصیلات خود را در آمریکا دریافت کرده‌ است، تصمیم گرفت تا سازمان "Black in AI" را تشکیل داده و پس از تجربه اینکه اغلب تنها فرد سیاه پوست حاضر در کنفرانس های AI در اروپا با هزاران شرکت کننده اساسا سفیدپوست بود. این یکی دیگر از مثال‌هایی است که فرضیات ساختاری، به عنوان مثال انکار ویزای حضور در کنفرانس های خارجی، علیه فناوران آفریقایی کار می‌کنند. سیسه (2018) می‌گوید که کلمات کلیدی مربوط به AI از شمال جهان، مانند تبعیض و عدالت، معمولا به عنوان ابزارهایی مورد نیاز برای اختراع انسانی شناخته شده‌اند. پربحثترین اولویت‌های جنوب جهان، مانند محاسبات پساستعماری (سینگ و گوزمان، 2021؛ علی، 2016) و محاسبات دکولونیال (ایرانی و همکاران، 2010)، نشان می‌دهد که درک محاسبات به عنوان چیزی محلی و چیزی که ساختار اجتماعی را ایجاد و حفظ می کند، مهم است. در عین حال، محاسبات همچنین امکان تغییرات رادیکال، استخراج داده و مشاهده عزت نفس انسان به عنوان پلی بین AI و حقوق بشر، استعمار داده و حاکمیت داده بومی را دارد (علی، 2016). سایر بخش‌های مهم تمرکز، همبستگی فمینیستی در شیوه‌های طراحی (لویتت و همکاران، 2019)، و عدالت داده بر تجربیات با فناوری‌های مبتنی بر داده به عنوان روش جدید زندگی تمرکز دارد. و به عنوان کومار و همکاران (2019) می بیند، به دنبال شیوه های خوب داده نویسی برای حاکمیت داده و حاکمیت مردمی بومی هستیم.

**خلاصه فصل**

این فصل به مسائل اقتصادی طبقه‌ای نگاه کرده است که سیستم‌های AI ممکن است دائمی کنند. از این رو، AI می تواند به عنوان یک عاملی عمل کند که وضعیت را تثبیت می کند و مگر اینکه برای آن برنامه‌ریزی شده باشد، نتایج جدیدی را تولید نمی کند، بلکه نتایج موجود را تکرار می‌کند. سیستم‌های AI که برای بررسی اینکه آیا افراد باید وام یا مسکن دریافت کنند استفاده می شوند، مگر اینکه برای آنها برنامه‌ریزی شده باشند تا بتوانند ریسک‌هایی را بپذیرند و افرادی که پارامترهای سیستم را برآورده نمی کنند مورد تحریم قرار خواهند گرفت. ما همچنین دیدیم که محل کارها به دلیل کنترل تاب‌دستان AI، با مشکلات تنوع روبرو هستند، جایی که کارگران با کمترین قدرت نه تنها توسط رؤسای خود و مدیران کنترل می شوند، بلکه در مورد کارگرانی که دارای سطح اجتماعی کمتر هستند، سیستم‌های AI که کنترل کارگران را برعهده دارند، نیز کنترل خواهند کرد. این تبعیض طبقاتی را هم می‌توان در سیستم آموزشی مشاهده کرد، جایی که برنامه های اصلاحات هم ممکن است تبعات غیر منتظره داشته باشند و پیچیدگی پیاده سازی AI برای تغییر نشان می‌دهد. همچنین دیدیم که AI نه تنها یک مسئله فردی، بلکه یک مسئله پایدار در کل سیاره است که با استخراج خاک منابع مورد استفاده برای زیرساختی که AI بر آن تکیه دارد، اختلافات طبقاتی را تثبیت می کند و مشاهده شود که AI یک پدیده جهانی است. بنابراین، AI می تواند برای ثروتمندان به عنوان یک منبع برای به دست آوردن ثروت بیشتر عمل کند، اما برای آدمهای فقیر، یک مانع جدی خواهد بود.

**مرجع ها**

Ali, S. M. (2016). A brief introduction to decolonial computing. *XRDS: Crossroads*, *The ACM Magazine for Students*, *22*(4), 16–21.

Aloisi, A., & Gramano, E. (2019). Artificial intelligence is watching you at work: Digital surveillance, employee monitoring, and regulatory issues in the EU context. *Comp. Lab. L. & Pol’y J*., *41*, 95.

Amrein-Beardsley, A. (2014). *Rethinking value-added models in education: Critical perspectives on tests and assessment-based accountability*. Routledge.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64.

Bales, R. A., & Stone, K. V. (2020). The invisible web at work: Artificial intelli- gence and electronic surveillance in the workplace. *Berkeley Journal of. Employment. & Labour. Law*, *41*, 1.

Bhargava, A., Bester, M., & Bolton, L. (2021). Employees’ perceptions of the implementation of robotics, artificial intelligence, and automation (RAIA) on job satisfaction, job security, and employability. *Journal of Technology in Behavioral Science*, *6*(1), 106–113.

Choi, J. (2020). Why Covid-19 only accelerates South Korea’s AI ambitions. *The Interpreter*. Retrieved March 1, 2022 from <https://www.lowyinstitute.org/> [the-interpreter/why-covid-19-only-accelerates-south-korea-s-ai-ambitions](https://www.lowyinstitute.org/)

Cisse, M. (2018). Look to Africa to advance artificial intelligence. *Nature*, *562*(7728), 461–462.

Coeckelbergh, M. (2020). *AI ethics*. MIT Press. Crawford, K. (2021). *The atlas of AI*.Yale University Press.

Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. [Reuters.com](http://reuters.com/). Retrieved December 6, 2021 from <https://www.reuters.com/> [article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G](https://www.reuters.com/)

De Stefano, V. (2019). “Negotiating the algorithm”: Automation, artificial intel- ligence, and labor protection. *Comparative. Labour. Law. & Policy Journal*, *41*, 15.

Delfanti, A. (2021). Machinic dispossession and augmented despotism: Digital work in an Amazon warehouse. *New Media & Society*, *23*(1), 39–55.

Dräger, J., & Müller-Eiselt, R. (2020). *We humans and the intelligent machines: How algo- rithms shape our lives and how we can make good use of them*.Verlag Bertelsmann Stiftung. Eisenstadt, L. F. (2021). # MeTooBots and the AI workplace. *University of Pennsylvania Journal of Business Law, Forthcoming*. Available at SSRN: <https://ssrn.com/>

[abstract=3921186](https://ssrn.com/)

Ema, A., Akiya, N., Osawa, H., Hattori, H., Oie, S., Ichise, R., Kanzaki, N., Kukita, M., Saijo, R., & Takushi, O. (2016). Future relations between humans and artificial intelligence: A stakeholder opinion survey in Japan. *IEEE Technology and Society Magazine*, *35*(4), 68–75.

Evans, K., de Moura, N., Chauvier, S., Chatila, R., & Dogan, E. (2020). Ethical decision making in autonomous vehicles: The AV ethics project. *Science and Engineering Ethics*, *26*(6), 3285–3312.

Fisher, A. N., Stinson, D. A., & Kalajdzic, A. (2019). Unpacking backlash: Individual and contextual moderators of bias against female professors. *Basic and Applied Social Psychology*, *41*(5), 305–325.

Guenole, N., & Feinzig, S. (2018). The business case for ai in HR. Retrieved November 4, 2020, from [https://www.ibm.com/downloads/cas/AGKXJX6M](https://www.ibm.com/).

Gutelius, B., & Theodore, N. (2019).The future of warehouse work: technological change in the US logistics industry. *UC Berkeley Labor Center*.

Hao, K. (2019). China has started a grand experiment in AI education. It could reshape how the world learns. *MIT Technology Review*, *123*(1). [https://www.](https://www.technologyreview.com/) [technologyreview.com/2019/08/02/131198/china-squirrel-has-started-a-](https://www.technologyreview.com/) [grand-experiment-in-ai-education-it-could-reshape-how-the/](https://www.technologyreview.com/)

Irani, L., Vertesi, J., Dourish, P., Philip, K., & Grinter, R. E. (2010). Postcolonial computing: a lens on design and development. *Proceedings of the SIGCHI con- ference on human factors in computing systems*. 10–15 April 2010, Atlanta Georgia USA.

Kang,Y., Cai, Z., Tan, C.-W., Huang, Q., & Liu, H. (2020). Natural language pro- cessing (NLP) in management research: A literature review. *Journal of Management Analytics*, *7*(2), 139–172.

Kantor, J., Weise, K., & Ashford, G. (2021).The Amazon that customers don’t see.

*The New York Times*.

Kaschak, E. (1978). Sex bias in student evaluations of college professors.

*Psychology of Women Quarterly*, *2*(3), 235–243.

Ku, L. (2021). *An introduction of NLP and how it’s changing the future of HR*. Plug and Play. [https://www.plugandplaytechcenter.com/resources/introduction-nlp-and-](https://www.plugandplaytechcenter.com/) [how-its-changing-future-hr/](https://www.plugandplaytechcenter.com/)

Kumar, N., Karusala, N., Ismail, A., Wong-Villacres, M., & Vishwanath, A. (2019). Engaging feminist solidarity for comparative research, design, and practice. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1–24.

Larson, C. (2018). China’s AI imperative. *Science*, 359, 628–630.

Levy, K.E.C. (2016). Digital surveillance in the hypermasculine workplace.

*Feminist Media Studies*, *16*(2), 361–365.

Li, J. J., Bonn, M. A., & Ye, B. H. (2019). Hotel employee’s artificial intelligence and robotics awareness and its impact on turnover intention: The moderat- ing roles of perceived organizational support and competitive psychological climate. *Tourism Management*, *73*, 172–181.

Littler, C. R. (1978). Understanding taylorism. *British Journal of Sociology*, 29, 185–202.

Lovett, R., Lee,V., Kukutai, T., Cormack, D., Rainie, S. C., & Walker, J. (2019). Good data practices for Indigenous data sovereignty and governance. In Angela Daly, Monique Mann, & S. Kate Devitt (Eds.), *Good Data*, 26–36.

Lundin, M., & Eriksson, S. (2016). Artificial intelligence in Japan (R&D, market and industry analysis). *EU-JAPAN Centre For Industrial Cooperation*, *39*, 26–36.

Moradi, P., & Levy, K. (2020). *The future of work in the age of AI: Displacement or risk- Shifting?* Frank Pasquale, Sunit Das, Markus Dirk Dubber (Eds.). Oxford University Press.

Nimbekar, R., Patil, Y., Prabhu, R., & Mulla, S. (2019). Automated resume eval- uation system using NLP. *2019 International Conference on Advances in Computing, Communication and Control (ICAC3)*. 20-21 December 2019, Mumbai, India.

Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self- driving cars: An applied trolley problem? *Ethical Theory and Moral Practice*, *19*(5), 1275–1289.

O’Hara, K., & Hall, W. (2021). *Four internets: data, geopolitics, and the governance of cyberspace*. Oxford University Press.

Okuda, T., Shiotani, S., Sakamoto, N., & Kobayashi, T. (2013). Background and current status of postmortem imaging in Japan: short history of “Autopsy imaging (Ai)”. *Forensic Science International*, *225*(1-3), 3–8.

People, S. (2020). *The changing face of HR: HR and people leaders’ report*. [Sage.com](http://Sage.com/). [https://www.sage.com/en-gb/blog/the-changing-face-of-hr/](https://www.sage.com/)

Ramboll. (2022). *AI for gender equality—Addressing inequality through AI*. [https://www.](https://www.vinnova.se/) [vinnova.se/globalassets/mikrosajter/ai-for-jamstalldhet-starker-tillvaxten-](https://www.vinnova.se/) [samhallsekonomin-och-arbetsmarknaden/ramboll---ai-for-gender-equality-](https://www.vinnova.se/) [2020-11-19.pdf](https://www.vinnova.se/)

Richardson, R. (2021). Defining and demystifying automated decision systems.

*Maryland Law Review, Forthcoming*, 81, 785–840.

Rosenberger, R. (2017). *Callous objects: Designs against the homeless*. University of Minnesota Press.

Sato, M., Yasuhara, Y., Osaka, K., Ito, H., Dino, M. J. S., Ong, I. L., Zhao, Y., & Tanioka, T. (2020). Rehabilitation care with Pepper humanoid robot: A quali- tative case study of older patients with schizophrenia and/or dementia in Japan. *Enfermeria Clinica*, *30*, 32–36.

Singh, R., & Guzmán, R. L. (2021). *Parables of AI in/from the Global South*. Retrieved December 7, 2021 from [https://www.4sonline.org/parables-of-ai-in-from-](https://www.4sonline.org/) [the-global-south/](https://www.4sonline.org/)

Smith, C. (2018). An employee’s best friend? How AI can boost employee engagement and performance. *Strategic HR Review*.

Tomé, E., Rivera, O., Lopez, D., & de Mello, P. S. (2020). Robotics and artificial intelligence (R&Ai) perceptions of consumers and producers: An interna- tional comparison among Portugal and Spain. *Proceedings of the European conference on the impact of artificial intelligence and robotics, ECIAIR*.

Toshkov, D. (2018). The ‘global south’ is a terrible term. Don’t use it! *Research Design Matters*. [http://re-design.dimiter.eu/?p=969](http://re-design.dimiter.eu/)

van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the machine meets the expert: An Ethnography of developing ai for hiring. *MIS Quarterly*, *45*(3), 1557–1580.

Zhang, C., Lu, Y., Feng, M., & Wu, M. (2019). Trucker behavior security surveil- lance pbased on human parsing. *IEEE Access*, *7*, 97526–97535.

7

**تلاقی‌گرایی و هوش مصنوعی پاسخ‌گو**

هوش مصنوعی به شکلی بی طرف نیست و با ارزش‌ها و فرهنگ‌های اجتماعی بارگذاری شده است که برخی از آن‌ها ممکن است منجر به تبعیض و نتایج ناتوان کننده شود. فصل‌های قبلی نشان می‌دهد که مسائل گوناگونی از تنوع، استثناء و تبعیض در حال بازی هستند هنگامی که هوش مصنوعی در محیط‌های اجتماعی پیچیده پیاده سازی می‌شود. در عین حال، بحث در مورد هوش مصنوعی با احساسات (انسانی) پر است. از یک طرف، شما به بهینه‌گرایی تکنولوژیکی از سیلیکون ولی و فرهنگ‌های برنامه نویسی برمی‌خورید که قسمت اجتماعی اتصال فناوری را فراموش می‌کنند. از سوی دیگر، صداهای بحرانی با کمترین درک از آنچه هوش مصنوعی را معنی می‌دهد و چه می‌تواند بکند و نتیجه گیری لودیت به دنبال هر تغییر فناوری هستند. و سپس گویی صداهای و نظرات گوناگون وجود دارد. برای چنین فناوری‌های خرابکارانه‌ای مانند هوش مصنوعی، نیازمند دانش از کارشناسان و همچنین کسانی هستیم که در این حوزه آگاهی به دست آورده‌اند. هدف من با این کتاب داوطلبانه است: برداشت اسرار از بسیاری از آنچه هوش مصنوعی است و بررسی اینکه چه کسانی در سیستم‌های مختلف هوش مصنوعی شامل شده یا مستثنی شده‌اند و چه مسائل تنوعی پدیدار می‌شود. در این فصل نهایی، قصد دارم یک قدم عقب برداشته و به خواننده راه‌های مختلفی را نشان دهم که چگونه می‌توان تأثیر هوش مصنوعی و تنوع بر زندگی، سازمان و یا موضوع پژوهشی شما را تجزیه و تحلیل کرد و راهکارهایی که می‌توانند ما را به سوی جوامع بیشتر دربرگیرنده و پاسخگو هدایت کنند با هوش مصنوعی، به اشتراک بگذارم.

**مشکلات جهانی با هوش مصنوعی**

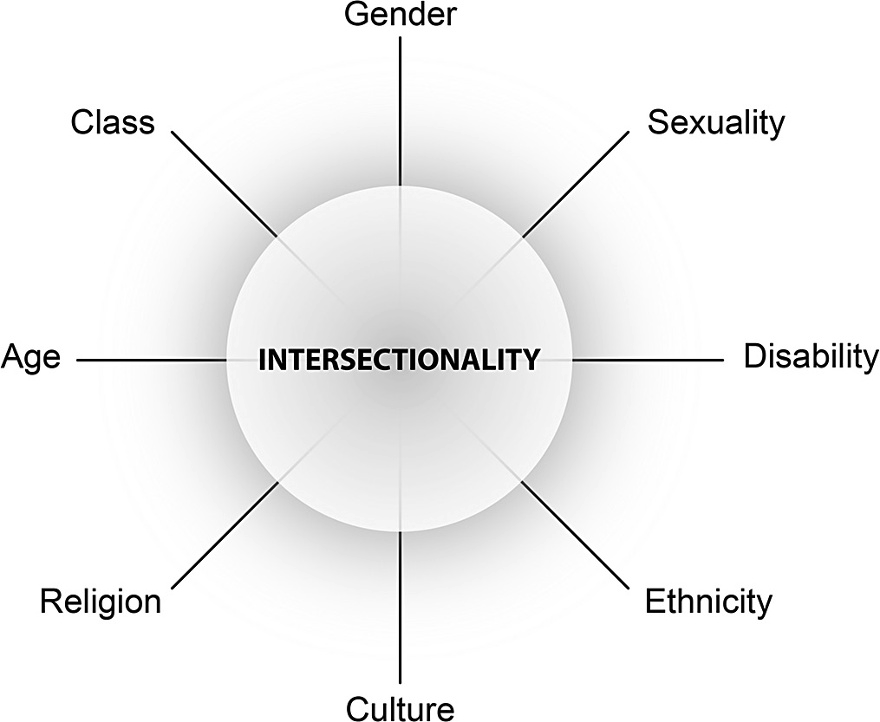
کاربران فناوری همیشه راه‌هایی را برای استفاده متفاوت از آن پیدا خواهند کرد و کسی که به عنوان کاربر استاندارد تصور می‌شود، در واقع به عنوان کاربر واقعی نمی‌شود. در این رابطه، منتقد شدن اسطوره‌ای که یک کاربر "فقط" یک ویژگی شخصی است بسیار مهم است. اگرچه این کتاب برای موضوعات جنسیت، همجنس‌گرایی، نژاد، بدن و طبقه‌بندی به فصل‌های مختلف تقسیم شده است، این به معنی این نیست که این ویژگی‌ها واحدهای مجزا هستند که بدون تأثیر یکدیگر هستند. علاوه بر این، مهم است بر تأکید شود که چنین تقسیمی نباید تسهیل کننده ذات گرایی باشد.

نیاز است که به بررسی اینکه چگونه این ویژگی‌ها به طور ساختاری، عملی و معنایی بر یکدیگر تأثیر می‌گذارند، توجه شود. تلاقی‌گرایی می‌تواند یک ابزار روش‌شناسی و نظری مهم برای این منظور باشد (Lovett et al.، ۲۰۱۹؛ Crenshaw، ۱۹۹۰؛ Mahoney et al.، ۲۰۱۹؛ May، ۲۰۱۵). تلاقی‌گرایی می‌تواند "هم تلاقی نظام‌ها، سیستم‌ها و دسته‌بندی‌هایی که سرکوب را تولید می‌کنند و هم تلاقی دسته‌بندی‌های هویت در فردان و گروه‌ها را برملا کند" (Runyan، ۲۰۱۸). مسائل تلاقی‌گرایی می‌تواند منجر به تبعیض دوتایی شود، به عنوان مثال در صورتی که برنامه‌های جایگزینی شامل مردان سیاه‌پوست و زنان سفیدپوست باشد، اما زنان سیاه‌پوست را شامل نشود. در این صورت تبعیض تلاقی‌گرایی به سختی قابل دفاع است، اگر که شرکت‌ها مردان سیاه‌پوست و زنان را شامل شده باشند. چندین مثال از اینکه چگونه استثنای تلاقی‌گرایی می‌تواند برای افراد با دو یا بیشتر ویژگی‌های شخصی یا هویتی کمتری مضر باشد، وجود دارد، به عنوان مثال، در مورد افراد سالخور و معلول همجنس‌گرای، زنان مبتلا به ایدز (Dressel et al.، ۱۹۹۷)، و تداخل نژاد، طبقه و جنسیت (Margulies، ۱۹۹۴) (شکل 7.1).

در طول این کتاب، ما شاهد بسیاری از مثال‌هایی بودیم که یک فرد فقط یک دسته نیست، بلکه باید شخصیت را به عنوان یک مفهوم پیچیده و وابسته به زمینه در نظر گرفت. استدلال من درباره فرد نیست، بلکه مهم است که ساختاردهی اجتماعی-فناورانه، تبعیض و محرومیتی که تکنولوژی رها می‌کند، چه معناهایی را می‌تواند به همراه داشته باشد. استفاده از چارچوب تلاقی‌گرایی در توسعه و پیاده‌سازی فناوری می‌تواند به شناسایی و کاهش این مسائل کمک کند. با در نظر گرفتن ابعاد چندگانه هویت و تلاقی آن‌ها، می‌توانیم بهتر درک کنیم که گروه‌های مختلف به چه صورتی می‌توانند به یک فناوری خاص تحت تأثیر قرار گیرند.

به عنوان مثال، فناوری شناسایی چهره ممکن است برای افراد با پوست سفید سبک کار کند، اما برای افراد با پوست سیاه به دلیل تبعیضات در مجموعه‌های داده‌ای که برای آموزش الگوریتم‌ها استفاده می‌شود، کار نکند. این می‌تواند منجر به تبعیض علیه افرادی از رنگ پوست متنوع در حوزه‌هایی مانند پلیس و فرآیندهای استخدام شود. با دربرداشتن تلاقی‌گرایی در فرآیند توسعه، تنوع و همگرایی مورد اولیه قرار می‌گیرد و فناوری به گونه‌ای طراحی می‌شود که برای تمامی گروه‌ها، بدون توجه به ویژگی‌های شخصی و هویتی‌شان، کار کند.

به طور کلی، باید تشخیص داد که توسعه فناوری در یک جامعه بدون تأثیرات بر روی برابری و سلسله مراتب اجتماعی به وجود نمی‌آید. با به کارگیری عینک تلاقی‌گرایی، می‌توانیم به سوی سامان‌دهی به سیستم‌های فناورانه مردم‌محور و شامل‌کننده حرکت کنیم.



دسته‌بندی‌های استفاده شده در نوشتارهای دانشگاهی در مورد جنسیت، طبقه، نژاد و غیره، به دلیل ضرورت استفاده از آن‌ها در تحلیل هویت‌ها برای ارائه گفتمان تغییر ساختاری است. پنج سال پیش، توجه به اینکه حسگرهای ضدعفونی کننده دست برای افراد رنگی نسبت به افراد سفیدپوست بدتر عمل می‌کنند، نمونه‌ای از تبعیض فناورانه می‌توانست باشد، اما به عنوان یک مورد جزئی تلقی می‌شد. با این حال، در شرایطی که دنیا درگیر یک وبا شد، عملکرد حسگرهای ضدعفونی کننده دست به اهمیت بالایی دست یافت، و بنابراین تباهی‌های نژادی می‌توانستند تأثیر بزرگی داشته باشند. همچنین، سخت‌افزارهایی مانند تبلت و تلفن‌های هوشمند که برای افراد سالخور با چین و چروک بیشتر در دست‌ها و صورت نسبت به دیگران بدتر عمل می‌کنند، می‌توانند تجربه کاربری بدتر را شتاب داده و به ایجاد عدم اعتماد در فناوری‌ها منجر شوند. در مثال دیگری که در کتاب دیدیم، شناسایی دست و صورت برای افراد رنگی نیز خوب عمل نمی‌کرد.

رویکرد تلاقی‌گرایی می‌تواند با پیش‌بینی بهتر نیازهای کاربران، به راه‌حل‌های تحقیقاتی و مهندسی شامل‌تری منجر شود (Cuadraz & Uttal، 1999). تیم متخصص دانشگاه استنفورد و کمیسیون اروپایی تحقیقات با جنبه جندری - که توسط استاد لوندا شیبینگر هدایت شده بود و من عضو آن بودم - لیستی از توصیه‌ها را برای اعمال تحلیل تلاقی‌گرایی به پروژه‌های تحقیقاتی و صنعتی ارائه دادند. این رویکرد شامل پنج مرحله است (Weber & Fore، 2007) و شامل نکاتی برای داده‌های کیفی و کمی در منبع کامل است. خلاصه آن به شرح زیر است:

مشکلات را "به منظور شناسایی عوامل و دسته‌بندی‌هایی از احتمالات مرتبط [مانند] ویژگی‌های زیستی، اجتماعی-فرهنگی و روانشناختی کاربران، مشتریان و شرکت‌کنندگان شناسایی کنید. [...] همچنین مهم است که شرایط اجتماعی، شامل شرایط جامعه، سطح سازمانی و اجتماعی [...] که تجربیات زندگی، فرصت‌ها و انتخاب‌های مردم را با تأثیرات مختلفی براساس جنسیت، نژاد، وضعیت اقتصادی-اجتماعی، جنسیت و موقعیت جغرافیایی تعیین می‌کند، در نظر گرفته شود. [...] درگیر کردن گروه‌های مختلف از موضوعات تحقیقاتی یا کاربران پایانی بالقوه در فرآیند تحقیق [...] می‌تواند تحلیل تلاقی‌گرایی را تیزتر کند و منجر به راه حل‌های شامل‌تری شود."

طراحی تحقیق "برای روشن شدن اثرات چندگانه دسته‌بندی‌ها و عوامل متفاوت ولی هم وابسته به هم، انجام می‌شود. روش‌های مناسبی را برای بررسی متغیرهای تلاقی‌گرایانه [...] تعیین کنید. فاکتورها ممکن است در زمینه‌های اجتماعی متفاوت باشند و ممکن است با گذر زمان تغییر کنند. [قبل از شروع جمع‌آوری داده‌ها] سوالات و دسته‌بندی‌ها را برای فرضیات نادرست یا استرئوتیپ‌های نادرست بررسی کنید."

جمع‌آوری داده‌ "اعمال رویکرد تلاقی‌گرایی، شامل جمع‌آوری داده‌های مرتبط با عواملی که با جنسیت و جنسیت (مانند نژاد، دین، رفتار جنسی، سبک زندگی، وضعیت اقتصادی، ناتوانی، دسته‌بندی‌های جنسی مرتبط با افراد ترنس‌ژندر، و غیره) تلاقی می‌کنند، را شامل می‌شود. به توجه به دسته‌بندی‌ها و ویژگی‌های فردی باید با تمرکز بر فاکتورهای سطح گروه (مانند در خانه، محله، سطح موسسات، منطقه، استان یا سطح ملی [...] در تحقیقات کیفی، نمونه باید به اندازه کافی گوناگون باشد تا موقعیت‌های مختلف تلاقی‌گرایانه مرتبط با مسئله تحقیق را دربرگیرد."

تحلیل "برای روشن شدن اثرات چندگانه دسته‌بندی‌ها و عوامل متفاوت ولی هم وابسته به هم، انجام می‌شود. در تحقیقات کمّی، نباید فقط به تأثیرهای اصلی (مانند برآورد اثرات جداگانه برای جنسیت، نژاد و جهت جنسی) توجه کرد، بلکه باید بررسی کنید که چگونه متغیرهای در تلاقی با یکدیگر هستند [...]. تحلیل‌های کیفی در بیشتر موارد به صورت بررسی اکتشافی انجام می‌شوند و باید شرح جامعی از دسته‌بندی‌ها، عوامل و فرآیندهای مختلفی که با هم تلاقی کرده و هویت، فرصت‌ها و شیوه‌های افراد را در یک زمینه خاص شکل می‌دهند، ارائه دهند."

انتشار "گزارش‌دهی باید مشخص کند که چگونه اطلاعات برای هر متغیر به دست آمده است. برای ترویج شفافیت، پژوهشگران باید همه نتایج مرتبط با تحلیل تلاقی‌گرایی از جمله نتایج ناقص و نامعین را گزارش دهند."

این لیست تنها یکی از بسیاری از ابزارهای نوآوری‌های جنسیتی است؛ برای بررسی بیشتر، به وب سایت genderedinnovations.stanford.edu مراجعه کنید. در خصوص پژوهشگران و دانشجویان هوش مصنوعی، مورد مطالعاتی بر روی مهندسی فناوری های کمکی برای سالمندان، واقعیت مجازی گسترده، تشخیص چهره، ربات اجتماعی با جنسیت، یادگیری ماشین، ترجمه ماشینی و دستیارهای مجازی، برای مطالعه موردی در تحلیل جنسیت و تلاقی‌گرایی به آنها مراجعه کنید.

در رویکرد تلاقی‌گرایی هوش مصنوعی، چند نکته وجود دارد که باید به آنها به طور ویژه توجه شود: (1) آیا در داده‌ها یا تفسیر داده‌ها به دلیل "خوشه های نمادین" (به عنوان مثال، داشتن مجموعه داده‌هایی با تعداد زیادی از افراد PoC و زنان، اما تعداد کمی از زنان سیاه پوست؟) به نقاط کور برخورد می‌شود؟ و یا داشتن مجموعه داده‌هایی با تعداد زیادی از افراد از جامعه LGBTQ + (اگر در مرحله مربوط به پروژه مورد استفاده باشند) و سالمندان، اما بدون وجود مردان سالمند همجنس‌گرا؟ افراد ممکن است در گروه‌های محروم قرار داده شوند، اما به دلیل اینکه افراد بسیار گوناگون هستند و تنها جنسیت، سن، جنبه‌های جنسیتی و غیره به آنها کمک نمی‌کنند، ما با خطر نقاط کور روبرو هستیم.(2) آیا توکنیسم برای توجیه شامل شدن گروه‌های محروم استفاده می‌شود، یا به دلیل احترام و ارزش عمیقی که برای چرا نیاز به دسته‌بندی‌های متنوع در ابتدا وجود دارد؟ داده و ابزارهای دیجیتال برای افراد است، نه برعکس، بنابراین داشتن پیشنهادات قوی، راهنماها و در برخی موارد، مقرراتی که شرکت‌های فناوری می‌توانند به آنها هدایت شوند، می‌تواند به هوشمندی بیشتر AI کمک کند. (3) آیا تلاقی‌گرایی به عنوان "رفع خطای یک بار" یا اقدامی مداوم برای کار، جایگزینی های مداوم، و مداخلات لازم برای اطمینان از عدم بروز شکل‌های جدید از بیرون کردن دیده می‌شود؟ همه این سه مسأله در مورد مسائل تنوع مرتبط با هوش مصنوعی وجود دارد، و رویکرد تلاقی‌گرایی می‌تواند به کاهش برخی از مسائل بیرون کردن و تبعیض که پیش می‌آیند، کمک کند.

تلاقی‌گرایی به عنوان یک اصطلاح، بدون جدال نیست. به عنوان مثال، برخی آن را مبهم و فاقد نتایج واضح می‌دانند. یک مطالعه در مورد راهبرد سیاسی "تلاقی‌گرایی" برای ایجاد برنامه کلی برای "همه" زنان (Gendered Innovations، 2020)، نشان داد چگونه تلاش برای استفاده از تلاقی‌گرایی، با هدف ایجاد دستور کار سیاسی مناسب برای "تمام" زنان، انجام شده است. در حالی که واردمین و سبستا (Vardeman and Sebesta، 2020)، در تحلیل خود پیرامون بحث‌های مربوط به تلاقی‌گرایی، از ما خواسته‌اند که به فرضیات مسأله برای ما اعتراض کنیم، زیرا این مسأله به ما کمک می‌کند تا به خوداندیشی، موقعیت گیری و انتقاد بپردازیم، اما در عین حال ریسک دارد که به ژست‌های روتین شود در فعالان و دانشمندان. ایجاد و پیاده‌سازی چارچوب‌های خوب برای تلاقی‌گرایی، یک چالشی است که معمولاً با مقاومت، سوءتفاهم و نادرستی در استفاده از آن هم از سوی منتقدان و هم از سوی حامیان مواجه می‌شود (Aldrin Salskov، 2020).

**به سوی تنوع فن‌آوری‌های اجتماعی در هوش مصنوعی؟**

این کتاب می‌توانست با عناوین مختلفی نامگذاری شود. "هوش مصنوعی و تبعیض" یا "هوش مصنوعی و برون‌کردن" تمرکز روی جنبه‌های منفی آنچه هوش مصنوعی می‌تواند باشد را فراهم می‌کرد، اما به نوعی یک نظارت منفی بر همه چیز است که اشتباهات هوش مصنوعی را به عنوان پیش‌فرض قرار می‌دهد. از طرف دیگر، صدا زدن به کتاب "هوش مصنوعی و شامل شدن" یا "هوش مصنوعی و تغییرات اجتماعی" در واقع یک پیش‌فرض نادرست در مورد هوش مصنوعی به عنوان محبوب، و آن مسائل به راحتی قابل حل خواهند بود. به همین دلیل، من می‌خواستم در میانه‌راهی قرار بگیرم، به طوری که از جنبه‌های منفی هوش مصنوعی و نحوه تاثیرگذار بودن آن به صورت تعصب‌آمیز، آگاه باشم، اما به زمینه‌هایی نیز بپردازم که احتمالات روییدن را دارند. روش راجریایی که در پی برقراری یک میانه‌راه بین دو نقطه نظر مخالف است، به نوعی دشوار است، زیرا هر طرف بحث لزوماً با برخی بخش‌های کتاب مخالف خواهد بود، زیرا هیچ کار خوبی بدون مجازات در مواجهه با مسائل شرور - یعنی مسائل اجتماعی که به گونه‌ای پیچیده و همبسته هستند که تقریباً امکان حل آنها وجود ندارد - انجام نمی‌شود. آسان است که خودمان را در منفیت فرو بریم. با این حال، نیاز به یک رویکرد مسأله‌محوری و خودآگاه که می‌تواند به سوی تنوع در هوش مصنوعی پیش برود، وجود دارد.

به عنوان یک دانشمند از مطالعات علوم و فناوری (STS) و تفکر پسااینسانی تکنوساینسی، چهار نکته اصلی را می‌خواهم بیان کنم. همانطور که در طول این کتاب دیدیم، فناوری‌هایی که در زندگی روزمره خود استفاده می‌کنیم توسط (1) افرادی که آنها را تولید می‌کنند شکل گرفته‌اند، که باعث می‌شود نمایندگی متنوع در پشت ساخت آنها حیاتی باشد زمانی که یک اسکریپت فناوری را تحلیل می‌کنیم (Akrich et al.، 2002؛ Fallan، 2008؛ Verbeek، 2005). به همین ترتیب، (2) دیدگاه کاربر و عضویت در هم‌توسعه مهم است، تا به صورت مستقیم از گروه‌هایی که برای آنها طراحی می‌کنید، سود ببرید، اما همچنین به کاربران کم‌تر اولویت بخشید و به گروه‌های محروم توجه کنید، به طور خاص زمانی که نگاه می‌کنید که چگونه کاربران فناوری را در خانه‌های خود به کار می‌برند (Berker et al.، 2005؛ Søraa et al.، 2021؛ Sørensen، 1994).

سومین نکته این است که (3) ما باید درکی داشته باشیم از شبکه سوسیوماتریال بازیگران، هم انسان و هم غیرانسان، که در ساختار آنچه هوش مصنوعی به معنای واقعی را دارد و چگونه در جوامع استفاده می‌شود، شریک هستند. با دیدن (1) تولیدکنندگان، (2) دیدگاه‌های کاربران و (3) شبکه‌های گسترده ای که فناوری در آنها مسیری را پیموده است، برخی از نخستین پیچیدگی‌های "درد و سود" تنوع ظاهر می‌شوند که می‌تواند به درک عمیق‌تری از چگونگی قرارگیری فناوری هوش مصنوعی منجر شود. این قرارگیری پیچیده است، زیرا داده‌های محلی برای مخاطبان و هدف‌های مختلف ساخته، قابل‌فریم‌کردن و پردازش می‌شوند (Rettberg، 2020) و دانش محلی همیشه در محیط اجتماعی ساخته می‌شود (Haraway، 1988) که همچنین بر روی آنچه فناوری است یا باید باشد، تأکید دارد. بنابراین، تنوع سوسیوتکنیکال به معنای تفکر بحرانی به عنوان سنگ‌بنای تجزیه و تحلیل نحوه تأثیر هوش مصنوعی بر گروه‌های اجتماعی مختلف و چگونگی ایجاد تعصب و برون‌کردن در زمینه‌های خاص است.

تنوع در هوش مصنوعی می‌تواند قبل، در طول و پس از پردازش داده پیاده شود. با نگاهی به مثال Gender Shades که Buolamwini (2017) در آن توضیح می‌دهد چگونه نرم افزار تشخیص چهره تبعیض بین زنان و افراد رنگین پوست ایجاد می‌کند - و بنابراین به صورت اینترسکتوری، به خصوص برای زنان رنگی - می‌توانیم ببینیم که فناوری غربی مستقر، تعصبات داخلی را در برابر افرادی که به عنوان "کاربر استاندارد" دسته‌بندی نمی‌شوند، در خود جای داده است. در اینجا، زمانی که مجموعه داده‌های پیش از پردازش تعصب دارند، اگر اقداماتی برای حذف تعصب از داده‌ها انجام نشود، داده خروجی نیز به ضرورت تعصب دار خواهد شد. در اینجا می‌بینیم که این یک شمشیر دو لبه است. اگر کاربر با "روش من" (مردان سفیدپوست تصور می‌کنند که کاربرانشان همچنین مردان سفیدپوستی هستند) پیکربندی شده باشد، یا اگر سعی کند کاربر را به عنوان "همه" تصور کند، هنوز "بعضی" را بر "هر کسی" قرار می‌دهد (Buolamwini، 2017) - بنابراین مرد را به عنوان "پیش‌فرض" کدگذاری می‌کند (Oudshoorn et al.، 2004). با این حال، اگر تمرکز بر روی همه، بعضی و یا هر کسی باشد - بدن‌هایی که مهم هستند (Perez، 2019) همچنان در زمینه‌های اجتماعی قرار دارند و تحت تأثیر هوش مصنوعی قرار می‌گیرند. در عین حال، باید به این نکته دقت کنیم که "تنوع" یک اصطلاح منازعه‌برانگیز است که با مشکلات و انتقادات خودش همراه است. همانطور که در فصل 1 بیان شد، تنوع می‌تواند به عنوان شرایط داشتن یا ترکیب شده از عناصر مختلف، به ویژه شامل افراد مختلف در یک گروه یا سازمان تعریف شود. بنابراین، استفاده از رویکرد تنوع، به معنای در نظر گرفتن یک مجموعه گسترده از ویژگی‌های شخصی مانند جنسیت، جهت‌گیری جنسی، سن، نژاد / قومیت، توانایی‌های فیزیکی / روانی و سلامت، زمینه اجتماعی-اقتصادی و دین / مذهب است. اما این "تقسیم بندی" بدون مشکل نیست، زیرا می‌تواند درک‌های فرهنگی-اجتماعی جهان را از انسان‌های متعصب بازتولید کند، و در نتیجه هوش مصنوعی متعصب ایجاد کند. زمانی که داده‌ها به یک برنامه کامپیوتری وارد می‌شود و فرد X را به عنوان "یک مهاجر آفریقایی قوم آلترناتیو با پست سطح پایین" شناسایی می‌کند، برنامه خود به طور ضروری هیچ درک عمیق‌تری از آن پارامترها ندارد یا این که چگونه آنها در جامعه ساختاریافته می‌شوند. به عنوان مثال، برای بودن یک مهاجر، برخی دانش پیشین لازم است: از کجا به کجا مهاجرت کرده است؟ چرا این شخص در حال حرکت است؟ آیا بخاطر قحطی یا فرصت شغلی مناسب است؟ با تکرار زمان و موقعیت‌های مختلف، می‌بینیم که هوش‌های مصنوعی اغلب نتایج نژادپرستانه‌ای را ارائه می‌دهند، زمانی که اولویت برای دسته‌بندی، به عنوان مثال، مهاجران به عنوان چیزی منفی قرار می‌گیرد. به همین ترتیب، یک هوش مصنوعی واقعا نمی‌داند که "آلترناتیو" در اینجا چه معنایی دارد، زیرا درک جنسیت در سیستم‌های هوش مصنوعی، اغلب با "جنسیت بیولوژیکی" گیج می‌شود، که باید زمانی که وب سایت‌های ابتدایی از شما می‌پرسند "آیا شما مرد یا زن هستید؟" مورد مشکل قرار گیرد. برای افراد جنسیت-مایع و غیردودویی، این دسته‌بندی دقیق "جنسیت = جنسیت بیولوژیکی" معنی نمی‌دهد. به همین ترتیب، برای پارامتر "مهاجر"، آیا یک هوش مصنوعی می‌داند که آیا شخص دیگر مهاجر است یا نه؟ آیا این زمانی است که او مرز را عبور می‌کند، یا به عنوان مثال زمانی که شهروندی جدیدی دریافت می‌کند؟ از نظر نسل بعدی او، آیا هنوز هم در فضای اجتماعی به عنوان "متفاوت" شناخته می‌شود، زمانی که با سؤالات نژادپرستانه مانند "اما شما واقعا اهل کجایید؟" روبرو می‌شوند؟

دیدگاه من با بیانی Crawford (2021) همخوانی دارد که می‌گوید هوش مصنوعی نه مصنوعی است و نه هوشمند. شفافیت اجتماعی-مادی این فرایند آشکار است، زیرا منابع از زمین برای ساخت ماشین‌ها، نرم‌افزارها و سیستم‌هایی که هوش مصنوعی به آنها وابسته است، استخراج می‌شود و توسط دستان و کار انسانی به هم می‌چسبد. همچنین، به عنوان هوشمند، اگر براساس آنچه ما به عنوان هوش انسانی درک می‌کنیم، بررسی شود، هیچ هوشمندی ندارد. همانطور که در فصل ۱ دیدیم، تعریفی که هوش مصنوعی را به دو بخش هوش مصنوعی محدود که در حال حاضر برای حل مشکلات خاص استفاده می‌شود و هوش مصنوعی عمومی که می‌تواند به عنوان انسان فکر و عمل کند، تقسیم می‌کند. اما آیا هوش مصنوعی عمومی در آینده وجود خواهد داشت، چیزی است که نمی‌دانیم؟ این یک تخیل اجتماعی-فنی با رشته‌های پیچیده‌ای از داستان، رویاهای آینده‌نگر و تصاویر است که در آزمایشگاه‌ها و شرکت‌ها در سراسر جهان به عنوان رویای یک موجود مصنوعی با هوش برافروخته می‌شود. همچنین، تاکنون ساخته نشده است و در آینده ممکن است هرگز ساخته نشود. وقتی این اتفاق می‌افتد، باید ذهن تحلیلی خود را به سمت آن برگردانیم، اما قبل از آن، باید در نظر داشت که هوش مصنوعی فعلی (محدود) چگونه بر جامعه معاصر و آینده نزدیک تأثیر می‌گذارد. همانطور که کتاب من سعی دارد نشان دهد، تنوع یکی از تأثیرات کلیدی است که باید در زمینه هوش مصنوعی با آن مقایسه شود.

اما، همانطور که در فصل ۱ دیدیم، تنوع به دلیل عدم شامل بودن برای همه افراد (دعوت به پارتی ولی نرقصیدن)، مورد انتقاد قرار گرفته است. بنابراین، هرگاه شامل بودن ذکر شود، باید بازتاب دهنده آن باشیم که منظورمان از کیست و چگونه شامل کردن گروه خاص مهم است. روش‌های تنوع و شامل بودن باید به طور مداوم استفاده، چالش داده و بازتعریف شوند و با هوش مصنوعی، هیچ "راه حل سریع" برای جادویی کردن داده‌ها، الگوریتم‌ها و خروجی "تنوع بخشیده شده برای همیشه" وجود ندارد. جامعه به طور مداوم در حال تغییر است، تعصبات قدیمی آشکار شده و بررسی می‌شوند، ارزش‌های جدید ساخته و اولویت داده می‌شوند، دانش، داده‌ها، اقدامات و تخیلات قرار گرفته‌اند، و اگر بتوانیم به سمت "حل مسئله‌گری" ماشین‌ها که Crawford (2021) آن را توصیف می‌کند، برای بازتنظیم روابط انسانی خود با ماشین‌ها حرکت کنیم، آنگاه به چیز جدیدی می‌رسیم - اما چه گام‌هایی برای تبدیل کردن هوش مصنوعی به چیزی عادلانه و پاسخگو برای همه می‌توان گرفت؟

**راهی به سوی هوش مصنوعی مسئولیت پذیر**

قبل از خاتمه دادن به این کتاب، قصد دارم در مورد مسئولیت‌های تنوع در هوش مصنوعی فکر کنم و پس از آن، برای شما به عنوان یک خواننده انتقادی، منابع و موضوعاتی را که در سفر آموزشی شما در مورد تنوع و هوش مصنوعی قرار دارد، ارائه دهم. بسیاری از کتاب‌های هوش مصنوعی، به ویژه در رشته‌های علوم اجتماعی و انسانی، بسیار ناامید کننده هستند. اگرچه، همانطور که در این کتاب نشان داده شده است، وقتی سیستم‌های هوش مصنوعی با آگاهی یا بدون آگاهی از نتایج شرارت آمیز طراحی یا استفاده می‌شوند، مسائل جامعه‌ای عمده‌ای با آنها مرتبط هستند. اما بگذارید فراموش نکنیم که هوش مصنوعی نیز تأثیرات مثبتی بر جامعه دارد، با تشخیص و توصیه‌های شخصی در حوزه بهداشت به عنوان مثال، از طریق ابزارها و سیستم‌های هوش مصنوعی و جراحی رباتیک. سیستم‌های مسیریابی و لجستیک هر روز به شکل مجدد افراد و کالاها را از طریق شبکه‌های پیچیده هدایت، برنامه‌ریزی و بهینه‌سازی می‌کنند تا از ترافیک دور بمانند و اطمینان حاصل کنند که شما و بسته‌های شما به موقع می‌رسند. سیستم‌های هوش مصنوعی برای ویرایش متن مطمئن می‌شوند که ارتباطات ما خنک باقی می‌ماند و ابزارها برای اطمینان از دستور زبان خوب به ویژه مفید هستند.

اگر قبل از هزاره سوم بزرگ شده باشید، ممکن است به دنیایی که بدون گوگل و الگوریتم‌های جستجو- اطلاعات بسیار دشوارتر بود و زمان‌بر بیشتر بود، یاد می‌کنید. همه نظر خود را درباره رسانه‌های اجتماعی دارند، اما دنیایی بدون آن، کمتر از نظر برقراری ارتباط با دیگران و گسترش گروه اصلی دوستان و خانواده شما اجتماعی است. مثال‌های دیگر شامل تجارت الکترونیک، بهداشت الکترونیک و آموزش الکترونیک است. من درباره هوش مصنوعی به صورت انتقادی نمی‌نویسم، زیرا برای آن بی‌علاقه نیستم. برعکس، من معتقدم که رویکردی انتقادی و سازنده می‌تواند به ما در بهبود و ساخت سیستم‌های هوش مصنوعی بهتر- برای سود بشریت- کمک کند.

برای اینکه هوش مصنوعی پاسخگو باشد، نیاز است که تمام عوامل در طول دوره عمر آن نیز پاسخگو باشند. این به این معنی است که طراحان، طراحان، تولید کنندگان و سازندگان هوش مصنوعی باید با دقت بسیاری درباره آنچه ایجاد می‌کنند، چرا، به چه کسانی کمک می‌کند و کی را از بین می‌برد فکر کنند. این بدان معنی است که کاربران، هر دو کاربر نهایی، کاربران ثانویه، کاربران تصادفی، اولویت‌های قبلی، دیررس‌ها، غیر کاربران و همه کسانی که به هر شکل دیگر از هوش مصنوعی استفاده می‌کنند یا تحت تأثیر استفاده آن قرار می‌گیرند، قادر به بیان نظر خود درباره اینکه چه معنی این برای زندگی‌شان دارد و به جدی گرفته شوند. این بدان معنی است که سیاست‌گذاران، نیروهای حاکم، رهبران و مدیران به جدی‌گرفتن تأثیر هوش مصنوعی بر سیستم‌هایی که آنها مدیریت یا کنترل می‌کنند، به خصوص تأثیرات بر تجارت، شغل، سطح قدرت و سیاست‌های حکومتی پرداخته شود. این تأثیرات بر روی کارگران در بخش‌های مختلفی که به طور مستقیم یا غیرمستقیم با هوش مصنوعی سر و کار دارند، بسیار مهم است. در کل، ما همه به عنوان انسان‌ها و غیرانسان‌ها در شبکه اجتماعی-مادی هوش مصنوعی، مسئولیت پاسخگویی آن را داریم. مسئولیت پذیری هوش مصنوعی به معنایی، در برخی موارد به این معناست که هوش مصنوعی می‌تواند به طور خودبخودی مسئول باشد. این یک سؤال فلسفی جذاب است که اخلاقی‌گرایان هوش مصنوعی، فیلسوفان، خداشناسان و همه افراد روزمره در اینترنت دارای نظر خود هستند. به عنوان مثال، به مطالعه Broussard (2018، ص 194) برای مرور کلی از بحث حقوق روبات بپردازید، به مطالعه Gunkel (2018) در مورد هوش مصنوعی و حقوق برای «حقوق بلافاصله روبات»، Darling (2016) که به دلیل تأثیرات آنتروپومورفیسم، برای گسترش حفاظت قانونی به ربات‌های اجتماعی دفاع می‌کند، به Gellers (2020) برای دیدگاه هوش مصنوعی در مورد حیوانات و قوانین محیط زیست، و Smith (2021، 2022) برای یک گفتمان الهی در این مورد مراجعه کنید. برای راه‌های پذیرش مسئولیت هوش مصنوعی، من می‌توانم بگویم که چند نکته باید در نظر گرفته شود:

1. باید از حل مسائل با رویکرد فناورانه، یعنی فکر کردن به اینکه برای هر مسئله اجتماعی، یک راه حل فناورانه سریع وجود دارد، خودداری کرد. گاهی برای حل مسایل تنوع، اصلاح‌طلبی یا تعصبات، نیازی به توسعه فناوری بیشتر نیست، بلکه باید یک قدم عقب برداشته شده و این سؤال پرسیده شود که "آیا در ابتدا واقعاً این مفید است، یا راه‌های دیگری برای تصور مسئله در دست است؟" در حالی که برخی مسائل بهترین‌شان با سیاست‌ها، بازیگران درگیر، یا تغییرات در نگرش‌ها حل می‌شود، برخی مسائل به سادگی قابل حل نیستند.
2. حق عدم شمارش نیز مهم است. برای اینکه هوش مصنوعی به یک دامنه کامل از پارامترها توجه کند، ممکن است از داده‌ها برای اهداف پلید استفاده شود. این به خصوص در مواردی مهم است که افراد آسیب پذیر ممکن است در سیستم‌هایی که موافقت یا علاقه‌ای به وجود آنها ندارند، لیست شوند. به عنوان مثال، پایگاه داده‌هایی که به تعداد افراد با جنسیت جذابیت شمول یافته‌اند، در دستور کشورهایی با قوانین LGBTQ+ وحشیانه، خطرناک است. برای نیروهای حاکمیتی قدرتمند که به دنبال شناسایی اعضای خاص گروه‌ها هستند، قطع یا عدم توسعه سیستم‌هایی که برای تولید نتایجی که به گروه‌های آسیب پذیر ضرر می‌زنند، طراحی شده‌اند، اهمیت دارد.
3. هوش مصنوعی را به عنوان چیزی جادویی بدون پیامدها تلقی نکنید، به عبارت Suchman و Suchman (2007) "ما باید از تکنولوژی‌ها خودداری کرد. باید الگوریتم‌ها را بررسی کرد، به دنبال عدم تساوی باشیم، و تعصبات موجود در سیستم‌های محاسباتی و صنعت فناوری را کاهش دهیم". به طور مشابه با دیگر تغییردهندگان قدرتمند ساختاری در جامعه، فناوری مانند هوش مصنوعی نیز باید مسئولیت‌پذیر باشد.

**منابع بیشتر**

هوش مصنوعی و تنوع یک موضوع با جذابیت‌های علمی و عامه‌پسند است. در این بخش، لیستی از منابع مرتبط برای مطالعات بیشتر در این زمینه، از جمله حوزه مرتبط هوش مصنوعی، روباتیک که نیز بر روی درک پرسش‌های تنوع برای هوش مصنوعی تأثیر می‌گذارد، فراهم شده است:

• "هوش مصنوعی برای خیر،" برجسته‌ترین پلتفرم عمل گرا، جهانی و شامل سازمان ملل متحد در زمینه هوش مصنوعی است. این پلتفرم با هدف شتاب دادن به کار برای دستیابی به اهداف توسعه پایدار، تشکیل شده است.

• AI4ALL یک سازمان غیرانتفاعی مستقر در ایالات متحده است که "به افزایش تنوع و شامل بودن در آموزش، پژوهش، توسعه و سیاست‌گذاری هوش مصنوعی می‌پردازد و از طریق آموزش و راهنمایی، به استعدادهایی که به صورت سنتی از آموزش هوش مصنوعی محروم شده‌اند، درب‌های هوش مصنوعی را باز می‌کند." این سازمان توسط اولگا روساکوفسکی، فی-فی لی و ریک سامر از دانشگاه استنفورد حمایت مالی می‌شود.

• "هوش مصنوعی در آفریقا" یک پلتفرم برای طرح‌های نوآورانه با تمرکز بر تغییر سیستماتیک برای جوامع آسیب‌پذیر در آفریقا است.

• "هوش مصنوعی در آسیا"، مجموعه‌ای از دیدگاه‌ها درباره بازتعریف حکومت هوش مصنوعی در آسیا، از برنامه قانون اساسی آسیا و کنراد-آدناور-اشتیفتون آلمان است که در سنگاپور فعالیت می‌کند.

• موسسه AI Now یک مرکز تحقیقات بین رشته‌ای در دانشگاه نیویورک در ایالات متحده است که با هدف تولید پژوهش‌های بین رشته‌ای و جذب علاقمندی عمومی، به کمک برآورده شدن این مطلب که سیستم‌های هوش مصنوعی شامل حسابرسی اجتماعی و محلی باشند، فعالیت می‌کند. این موسسه در سال ۲۰۱۷ توسط کیت کرافورد (قبلا در مایکروسافت) و مردیت ویتاکر (قبلا در گوگل) تأسیس شده است.

• AI4EU با هدف ارائه آخرین اخبار در مورد کاربردهای پیشرفته هوش مصنوعی، روندهای توسعه، پژوهش، اخلاق و تأثیرات اجتماعی، با گردآوری جامعه هوش مصنوعی و حمایت از ارزش‌های اروپایی، به دنبال تحقق این هدف است.

• "لیگ عدالت الگوریتمیک" یک سازمان فعال در حوزه دفاع دیجیتال است که در کمبریج، ماساچوست، ایالات متحده قرار دارد. هدف AJL افزایش آگاهی عمومی درباره تأثیرات هوش مصنوعی، تجهیز دفاع کنندگان با پژوهش‌های تجربی برای تقویت کمپین‌ها، ساخت صدا و انتخاب جوامعی که بیشترین تأثیر را متحمل شده‌اند، و تحریک پژوهشگران، سیاست‌گذاران و صنعت‌گران به منظور کاهش تعصبات و آسیب‌های هوش مصنوعی است. این سازمان در سال ۲۰۱۶ توسط جوئی بوآلاموینی تأسیس شده است.

• "آتوناتون" یک استودیو تحقیقاتی است که به اختراع روش‌های بهتر برای زندگی با ماشین‌ها می‌پردازد، از طریق ترکیب نوآوری آزمایشگاه تحقیقات و خلاقیت استودیو طراحی، پروتوتایپ‌هایی از آینده‌های جایگزین ایجاد می‌کند. این استودیو توسط مدلین گانون در منطقه بزرگ پیتسبرگ، ایالات متحده، هدایت می‌شود.

• DataKind با تمرکز بر به ارمغان آوردن علم داده [برای ساخت] یک سیاره پایدار که در آن همه به نیازهای اساسی انسانی خود دسترسی دارند، توسط جیک پوروی با مقر در نیویورک، تأسیس شده است.

• "علم داده برای خیر اجتماعی" با هدف "ایجاد و حفظ جوامع، برنامه‌ها و راه‌حل‌هایی که استفاده مسئولانه از علم داده و هوش مصنوعی را برای بهبود عدالت اجتماعی فراهم می‌کنند" در دانشگاه شیکاگو است.

• "Data & Society" یک سازمان مستقل غیرانتفاعی تحقیقاتی در نیویورک است که با مأموریت تولید "پژوهش اصلی در مورد موضوعات شامل هوش مصنوعی و خودکارسازی، تأثیر تکنولوژی بر کار و سلامتی و تبلیغات آنلاین" و "دعوت کردن پژوهشگران، سیاست‌گذاران، فناوران، خبرنگاران، کارآفرینان، هنرمندان و وکلای حقوقی به منظور چالش قدرت و هدف تکنولوژی در جامعه" فعالیت می‌کند. این سازمان توسط دانا بوید (پژوهشگر همکار در مایکروسافت و استاد مهمان در دانشگاه نیویورک) حمایت مالی می‌شود.

• "آزمایشگاه حریم خصوصی داده" در موسسه علوم اجتماعی کمیتاتیو در دانشگاه هاروارد "فناوری را ایجاد و استفاده می‌کند تا به بررسی و حل مسائل اجتماعی، سیاسی و حاکمیتی بپردازد و به دیگران یاد می‌دهد که چگونه همین کار را انجام دهند." این آزمایشگاه توسط لاتانیا سوینی رهبری می‌شود.

• "بنیاد رباتیک مسئول" یک سازمان غیرانتفاعی، غیردولتی با مقر در لاهه، هلند است که با مأموریت "شکل دادن به آینده‌ای از طراحی، توسعه، استفاده، تنظیم و پیاده‌سازی مسئولانه رباتیک و هوش مصنوعی" فعالیت می‌کند. این بنیاد توسط آیمی وینزبرگه و نوئل شارکی رهبری می‌شود.

• "رباتیک برای اتحادیه اروپا" یک پروژه هماهنگی و پشتیبانی است که توسط کمیسیون اروپایی تحقیقات تأسیس شده و به بررسی تأثیر ربات‌ها بر جامعه می‌پردازد. این بررسی به گستره‌ی دید اجتماعی و انسانی، با تمرکز بر روی این که چگونه ستادهای مختلف جامعه برای زندگی و بخش‌های خود تحت تأثیر رباتیک قرار می‌گیرند، انجام می‌شود. این پروژه توسط آنلی روز در CIVITTA استونی رهبری می‌شود.

• "آزمایشگاه تحقیقاتی رباتیک اجتماعی و هوشمند" در دانشگاه واترلو، به بررسی تعامل انسان-ربات، رباتیک شناختی و توسعه‌ای، و هوش مصنوعی بدنی می‌پردازد. این آزمایشگاه توسط کرستین داتنهان و کریستوفر نهانیو رهبری می‌شود.

**جمع بندی:**

در این کتاب، نگاهی به نحوه‌ای که سیستم‌های هوش مصنوعی با تنوع در بین کاربران و تولیدکنندگان آنها مرتبط هستند و نحوه‌ای که تفکر پویا بر روی سیستم‌های هوش مصنوعی تأثیر می‌گذارد، شده است. در این کتاب، بررسی شده است که چگونه این سیستم‌ها در شبکه‌ها و ساختارهای فنی-اجتماعی قرار دارند و از جهانی که در آن ساخته شده‌اند جدا نیستند. آگاهی از اینکه چه کسانی از هوش مصنوعی بهره‌مند هستند و چه کسانی مورد استثناء قرار می‌گیرند، برای حرکت به سمت استفاده عادلانه‌تر از هوش مصنوعی که به گروه‌های متنوعی از مردم بهره می‌دهد، بسیار مهم است. سیستم‌های هوش مصنوعی می‌توانند به عنوان مکانیزم‌های بازنشر، نظارت و کنترل بر "غیرقابل قبول" در جامعه و سیستم‌های مشابه استفاده شوند. با این حال، تضمین کامل عدالت، از بین بردن تعصب و عدم وجود استثناء غیرممکن است. یک کامپیوتر برای تفکیک دو متغیر 0 و 1 ساخته شده است و از آنجا، تصمیمات برای اولویت‌بندی‌ها گرفته می‌شود. تعصبی که در جهانی که منابع، زمان و داده ناکافی هستند و باید اولویت‌ها تعیین شود، همیشه وجود دارد. این به معنای پذیرش هرگونه تعصب، و به طور خاص تصمیمات ناعادلانه‌ی گرفته شده توسط انسان‌ها یا سیستم‌های کامپیوتری نیست!

با مسئولیت‌پذیری، آگاهی و برنامه‌های متمرکز بر تفکر پویا، هوش مصنوعی می‌تواند به عنوان یک ابزار تغییر، برای جلب شامل بودن و ساخت جهانی بهتر استفاده شود. با شکستن جعبه‌های سیاه و روشن کردن اینکه هوش مصنوعی چیزی است که توسط انسان‌ها ساخته شده و برای انسان‌ها است، می‌توانیم اطمینان حاصل کنیم که گروه‌هایی از مردم براساس ویژگی‌های فردی خود، مورد استثناء قرار نمی‌گیرند. اگرچه هوش مصنوعی محدود است، اما دسترسی انسان به آن باید عمومی باشد و برای همه‌ی افراد در دسترس باشد. امید من این است که با خواندن این کتاب، شما به سوی یک جهان سیاتکنیکال متنوع و شامل هوش مصنوعی کار کنید - با شروع از شیوه‌ها و سازمان خود، الگوریتم به الگوریتم.

**منابع:**

Akrich, M., Callon, M., Latour, B., & Monaghan, A. (2002). The key to success in innovation part I: the art of interessement. *International Journal of Innovation Management*, *6*(02), 187–206.

Aldrin Salskov, S. (2020). A critique of our own? On intersectionality and “Epistemic Habits” in a study of racialization and homonationalism in a nor- dic context. *NORA-Nordic Journal of Feminist and Gender Research*, *28*(3), 251–265.

Berker, T., Hartmann, M., & Punie, Y. (2005). *Domestication of media and technology*. McGraw-Hill Education.

Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.

Buolamwini, J. A. (2017). *Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*. MIT.

Crawford, K. (2021). *The atlas of AI*.Yale University Press.

Crenshaw, K. (1990). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, *43*, 1241.

Cuadraz, G. H., & Uttal, L. (1999). Intersectionality and in-depth interviews: Methodological strategies for analyzing race, class, and gender. *Race, Gender & Class*, 6, 156–186.

Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects. In Ryan Calo, A. Michael Froomkin, & Ian Kerr (Eds.), *Robot law* (pp. 213–232). Edward Elgar Publishing.

Dressel, P., Minkler, M., & Yen, I. (1997). Gender, race, class, and aging: advances and opportunities. *International Journal of Health Services*, *27*(4), 579–600.

Fallan, K. (2008). De-scribing design: Appropriating script analysis to design history. *Design Issues*, *24*(4), 61–75.

Gellers, J. C. (2020). *Rights for Robots: Artificial Intelligence, Animal and Environmental Law*

(Edition 1). Routledge.

Gendered Innovations (2020). *Intersectional approaches*. Retrieved February 2, 2022 from [http://genderedinnovations.stanford.edu/methods/intersect.html](http://genderedinnovations.stanford.edu/)

Gunkel, D. J. (2018). *Robot rights*. MIT Press.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575–599.

Lovett, R., Lee, V., Kukutai, T., Cormack, D., Rainie, S. C., & Walker, J. (2019). Good data practices for Indigenous data sovereignty and governance. In Angela Daly, Monique Mann, & S. Kate Devitt (Eds.), *Good Data*, 26–36. [https://](https://books.google.no/) [books.google.no/books?hl=en&lr=&id=Y0vUDwAAQBAJ&oi=fnd&pg=PA8](https://books.google.no/) [&dq=Good+data+(2019+devitt&ots=hraWipl8Jl&sig=lAr3kXxmsDbjKrb](https://books.google.no/) [wuFi17Jev5\_o&redir\_esc=y#v=onepage&q=Good%20data%20(2019%20](https://books.google.no/) [devitt&f=false](https://books.google.no/).

Mahoney, A. D., Cuellar, M. G., Johnson-Ahorlu, R. N., Jones, T.-A., Warnock, D. M., Clayton, K. A., Ray, V. E., Lee, E. M., Maynard, T., & Goerisch, D. (2019). *Intersectionality and higher education: Identity and inequality on college campuses*. Rutgers University Press.

Margulies, P. (1994). Asylum, intersectionality, and AIDS: Women with HIV as a persecuted social group. *Georgetown Immigration Law Journal*, *8*, 521.

May, V. M. (2015). *Pursuing intersectionality, unsettling dominant imaginaries*. Routledge.

Oudshoorn, N., Rommes, E., & Stienstra, M. (2004). Configuring the user as everybody: Gender and design cultures in information and communication technologies. *Science,Technology, & Human Values*, *29*(1), 30–63.

Perez, C. C. (2019). *Invisible women: Exposing data bias in a world designed for men*. Random House.

Rettberg, J. W. (2020). Situated data analysis: a new method for analysing encoded power relationships in social media platforms and apps. *Humanities and Social Sciences Communications*, *7*(1), 1–13.

Runyan, A. S. (2018). *What is intersectionality and why is it important?* American Association of University Professors. Retrieved February 5, 2022 from [https://www.aaup.org/article/what-intersectionality-and-why-it-important](https://www.aaup.org/)

Smith, J. K. (2021). *Robotic persons: Our future with social robots*. Westbow Press. Smith, J. K. (2022). *Robot Theology: Old Questions through New Media*. Wipf and Stock

Publishers.

Søraa, R. A., Nyvoll, P., Tøndel, G., Fosch-Villaronga, E., & Serrano, J. A. (2021). The social dimension of domesticating technology: Interactions between older adults, caregivers, and robots in the home. *Technological Forecasting and Social Change*, *167*, 120678

Sørensen, K. H. (1994). *Technology in use:Two essays in the domestication of artefacts* (Centre for Technology and Society Working Paper 2); p. 94). Centre for Technology and Society.

Suchman, L., & Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situ- ated actions*. Cambridge University Press.

Vardeman, J., & Sebesta, A. (2020). The problem of intersectionality as an approach to digital activism: the Women’s March on Washington’s attempt to unite all women. *Journal of Public Relations Research*, *32*(1-2), 7–29.

Verbeek, P.-P. (2005). Artifacts and attachment: A post-script philosophy of mediation. In Hans Harbers (Ed.), *Inside the politics of technology* (pp. 125–146). Amsterdam University Press.

Weber, L., & Fore, M. E. (2007). Race, ethnicity, and health: An intersectional approach. In *Handbooks of the sociology of racial and ethnic relations* (pp. 191–218). Springer