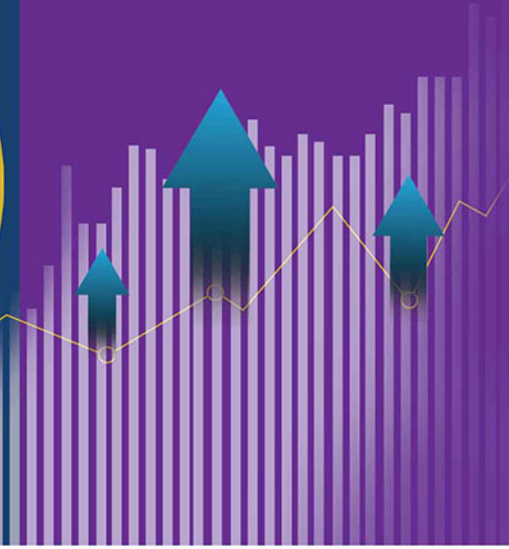




مرکز ملی فضای مجازی
پژوهشگاه فضای مجازی

گزارش
سند
چهار



طرح راهبردی ملے تحقیق و توسعه هوش مصنوعے
نسخه ۲۰۱۹

The National Artificial Intelligence
Research And Development Strategic
Plan : 2019 Update

سند

گزارش
سند

گزارش شماره ۴

اسفند ۱۴۰۰



مرکز ملی فضای مجازی
پژوهشگاه فضای مجازی

طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی: نسخه ۲۰۱۹

محتوای انتشار یافته در این گزارش
الزاماً بیانگر دیدگاه مرکز ملی فضای مجازی نیست

تهیه شده در پژوهشگاه فضای مجازی
(گروه مطالعات اخلاقی)

مترجم: دکتر جواد درویش آقاجانی (دکتری فلسفه
علم و فناوری دانشگاه شریف)

ناظر علمی: محمد سجاد زارع میرک آباد
(کارشناسی ارشد فلسفه دانشگاه تهران)

حقوق مادی و معنوی این اثر متعلق به مرکز ملی فضای
مجازی است و استفاده از آن با ذکر منبع مجاز می باشد.

نشانی: تهران، میدان آرژانتین، خیابان بیهقی، نبش
خیابان ۱۶ غربی، پلاک ۲۰
تلفن: ۰۲۱-۸۶۱۵۱۰۶۱
کد پستی: ۱۵۱۵۶۷۴۳۱۱

فهرست

۷ سخن نخست
۱۱ مقدمه مترجم
۱۵ همکاران گرامه
۲۱ چکیده
۲۷ مقدمه

بخش اول

- ۴۵ سرمایه‌گذاری بلندمدت در تحقیقات هوش مصنوعی
- ۴۷ حفظ سرمایه‌گذاری‌های بلندمدت در تحقیقات بنیادی هوش مصنوعی، نسخه ۲۰۱۹
- ۵۱ تحقیقات هوش مصنوعی بنیادی بلندمدت: برنامه‌های تحقیق و توسعه اخیر موسسه
- ۵۳ پیشبرد روش‌شناسی‌های داده‌محور برای تولید دانش
- ۵۴ ارتقا قابلیت‌های ادراکی سیستم‌های هوش مصنوعی
- ۵۵ شناخت قابلیت‌ها و محدودیت‌های هوش مصنوعی
- ۵۷ پیگیری تحقیقات در باره هوش مصنوعی همه‌منظوره
- ۵۸ توسعه سیستم‌های هوش مصنوعی مقیاس‌پذیر
- ۵۹ ترغیب تحقیقات به سوی هوش مصنوعی انسانگونه
- ۶۰ توسعه روبات‌هایی با توانایی و اعتبار بیشتر
- ۶۱ پیشرفت ساخت‌افزار برای تقویت هوش مصنوعی
- ۶۲ ایجاد هوش مصنوعی برای ساخت‌افزارهای پیشرفته

بخش دوم

- ۶۵ توسعه روش‌های مؤثر برای همکاری انسان و هوش مصنوعی
- توسعه سیستم‌های هوش مصنوعی که قابلیت‌های انسانی را تکمیل کرده و به آن‌ها
- ۶۷ می‌افزایند، با افزایش تمرکز بر آینده کار (به روز سازی ۲۰۱۹)
- ۶۹ همکاری انسان و هوش مصنوعی: برنامه‌های اخیر تحقیق و توسعه مؤسسات
- ۷۵ جستجوی الگوریتم‌های جدید برای سیستم‌های هوش مصنوعی آگاه-انسانی

- ۷۶ _____ توسعه فنون هوش مصنوعی برای افزایش توانایی انسان
- ۷۸ _____ توسعه فنونی برای مصورسازی و واسطه‌های انسان-هوش مصنوعی
- ۷۹ _____ توسعه سیستم‌های پردازش زبانی مؤثرتر

بخش سوم

- ۸۱ _____ بررسی و فهم پیامدهای اخلاقی، قانونی و اجتماعی هوش مصنوعی
- ۸۳ _____ بررسی ملاحظات اخلاقی، قانونی و اجتماعی هوش مصنوعی (نسخه ۲۰۱۹)
- ۸۴ _____ تبیین پذیری، انصاف و شفافیت: برنامه‌های اخیر مؤسسات تحقیق و توسعه
- ۸۷ _____ بهبود انصاف، شفافیت و پاسخگویی آگاهانه
- ۸۸ _____ ساخت هوش مصنوعی اخلاقی
- ۸۹ _____ طراحی معماری برای هوش مصنوعی اخلاقی

بخش چهارم

- ۹۱ _____ تضمین ایمنی و امنیت سیستم‌های هوش مصنوعی: نسخه ۲۰۱۹
- ۹۳ _____ ایجاد سیستم‌های قوی و قابل اعتماد هوش مصنوعی
- ۹۶ _____ ایمنی و امنیت هوش مصنوعی: برنامه‌های اخیر تحقیق و توسعه
- ۹۹ _____ بهبود قابلیت تبیین و شفافیت
- ۹۹ _____ ایجاد اعتماد
- ۱۰۰ _____ ارتقا تالیف و اعتبارسنجی
- ۱۰۱ _____ ایمن‌سازی در برابر حملات
- ۱۰۲ _____ دستیابی به ایمنی بلند مدت هوش مصنوعی و همترازی ارزش‌ها

بخش پنجم

- توسعه مجموعه داده‌ها و محیط‌های عمومی مشترک برای آموزش و آزمایش هوش مصنوعی: نسخه ۲۰۱۹
- ۱۰۳ _____ توسعه مجموعه داده‌ها و چالش‌های مربوطه
- ۱۰۵ _____ افزایش دسترسی به مجموعه داده‌ها و چالش‌های مربوطه

- مجموعه داده‌ها و محیط‌های عمومی مشترک برای آموزش و آزمایش هوش مصنوعی: برنامه‌های اخیر مؤسسات تحقیق و توسعه — ۱۰۸
- توسعه و در دسترس قرار دادن طیف گسترده‌ای از مجموعه داده‌ها برای برآورده کردن نیازهای طیف متنوعی از ذینفعان و کاربردهای هوش مصنوعی — ۱۱۱
- پاسخگو کردن منابع آموزش و آزمایش به ذینفعان تجاری و عمومی — ۱۱۳
- توسعه بانک‌ها و جعبه‌ابزار نرم‌افزارهای منبع باز — ۱۱۴

بخش ششم

- اندازه‌گیری و ارزیابی فن‌آوری‌های هوش مصنوعی از طریق استانداردها و ملاک‌ها: نسخه ۲۰۱۹ — ۱۱۷
- حمایت از توسعه استانداردهای فنی هوش مصنوعی و ابزارهای مربوطه — ۱۱۹
- استانداردها، ملاک‌ها و ابزارهای مرتبط: برنامه‌های اخیر مؤسسات تحقیق و توسعه — ۱۲۱
- توسعه طیف وسیعی از استانداردهای هوش مصنوعی — ۱۲۳
- ایجاد ملاک‌های فناوری هوش مصنوعی — ۱۲۴
- افزایش دسترسی به بسترهای آزمایش هوش مصنوعی — ۱۲۶
- مشارکت جامعه هوش مصنوعی در استانداردها و معیارها (الگوها) — ۱۲۷

بخش هفتم

- درک بهتر نیازهای نیروی کار ملی تحقیق و توسعه هوش مصنوعی: نسخه ۲۰۱۹ — ۱۳۰
- به جلو راندن نیروی کار تحقیق و توسعه هوش مصنوعی، شامل کسانی که بر روی سیستم‌های هوش مصنوعی کار می‌کنند و کسانی که در کنار آن‌ها کار می‌کنند تا رهبری آمریکا را حفظ کنند — ۱۳۳
- نیروی کار ملی تحقیق و توسعه هوش مصنوعی: فعالیت‌های اخیر مؤسسات — ۱۳۷

بخش هشتم

- گسترش مشارکت بخش‌های عمومی - خصوصی برای سرعت بخشیدن به پیشرفت‌ها در هوش مصنوعی — ۱۴۱

پیشبرد زیست‌بوم نوآوری هوش مصنوعی کشور، از طریق پل زدن میان دولت، دانشگاه‌ها

و صنعت _____ ۱۴۴

شرکت‌های دولتی - خصوصی: برنامه‌های اخیر مؤسسات تحقیق و توسعه _____ ۱۴۶

بخش نهم

در باره شورای ملی علوم و فناوری _____ ۱۵۳

در باره دفتر سیاست‌گذاری علم و فناوری _____ ۱۵۵

در باره کمیته منتخب هوش مصنوعی _____ ۱۵۶

در مورد کمیته فرعی یادگیری ماشینی و هوش مصنوعی _____ ۱۵۷

در باره کمیته فرعی تحقیق و توسعه شبکه‌سازی و فن‌آوری اطلاعات _____ ۱۵۷

در باره گروه کاری بین موسسه‌های تحقیق و توسعه هوش مصنوعی _____ ۱۵۸

در باره این نوشتار _____ ۱۵۸

اطلاعات کپی‌رایت _____ ۱۵۹

سخن نخست



فضای مجازی با شتاب شگرف و رو به تزایدی که در حال بسط و گسترش است تمام ساحات اجتماعی، اقتصادی، سیاسی و فرهنگی زندگی بشر را درنوردیده و هر روز بخش بزرگی از زندگی واقعی را در خود فرو برده و حیات متفاوت و جدیدی به آن می‌دهد. لذا به نظر می‌رسد دو نگاه کلان به فضای مجازی وجود دارد: نگاه اول که بالاخص در ابتدای رشد و تکوین فضای مجازی مسلط شده بود، آن را همچون ابزاری کنار سایر ابزارهای بشری تصویر می‌کرد که تنها طریقت داشت. اما نگاه دوم، در نتیجه رشد تحولات خیره‌کننده فضای مجازی و سایه گسترده آن در حوزه‌ها و شئون بشر در یک دهه اخیر آن را چون سکویی می‌داند که بسیار فراتر از شأن ابزاری حیات انسان‌ها را سامان جدیدی داده و ادعای تمدن نوینی را دارد. رویکردی که از قضا از چشمان بصیر رهبر انقلاب نیز دور نمانده و انتظاری تمدنی از فضای مجازی در ایران را مطالبه داشته‌اند.

در همین راستا گزارش‌های عصر فضای مجازی تلاش می‌کند تا فهم سازمان‌ها و دستگاه‌های مرتبط با حوزه فضای مجازی را ارتقاء بخشیده و آن‌ها را برای مواجهه فعال و خردمندانه با تحولات این عرصه مهیا سازد.

سید ابوالحسن فیروزآبادی
دبیر شورای عالی و رئیس مرکز ملی فضای مجازی

مقدمه مترجم



سند پیش رو بخشی از طرح کلان «پیشگامی آمریکا در هوش مصنوعی» است که با هدف استمرار رهبری آمریکا در زمینه هوش مصنوعی در سال ۲۰۱۹ به امضای رئیس جمهور این کشور رسیده است. این طرح در اصل یک راهبرد دولتی است که در همکاری و تعامل با بخش خصوصی، دانشگاه، عموم مردم و شرکای بین‌المللی اجرا خواهد شد. یکی از مهمترین بخش‌های این طرح کلان که سند پیش رو را تشکیل می‌دهد، بررسی اولویت‌های سرمایه‌گذاری در زمینه تحقیق و توسعه (R&D) هوش مصنوعی است. این موضوع پیش از این و در سال ۲۰۱۶ نیز مورد توجه سیاستگذاران آمریکایی قرار گرفته بود و اولین نسخه از «طرح راهبرد ملی تحقیق و توسعه هوش مصنوعی» که طرح حاضر اشتراکات زیادی با آن دارد، همان زمان منتشر شده بود.

سند حاضر از ۸ راهبرد اصلی تشکیل شده است که بطور خلاصه عبارتند از: ضرورت سرمایه‌گذاری بلند مدت در هوش مصنوعی، توسعه روشهای تعامل انسان و هوش مصنوعی، توجه به پیامدهای اخلاقی-اجتماعی، ارتقا امنیت سیستم‌های هوش مصنوعی، ایجاد مجموعه

داده‌ها برای آزمایش نتایج، ارزیابی فناوری‌های هوش مصنوعی از طریق استانداردهای مشخص، توجه بیشتر به نیروهای کار متخصص در زمینه تحقیق و توسعه، گسترش همکاری بین بخش‌های خصوصی و دولتی.

این سند بر تحقیق و توسعه هوش مصنوعی به عنوان یک امر دولتی، تاکید دارد. اگرچه آن را در انحصار دولت نمی‌داند ولی از وظایف دولتها به شمار می‌آورد. در عین حال نکته قابل توجه در این سند، نحوه تنظیم آن است. در گام اول، موضوع تحقیق و توسعه در هوش مصنوعی به عنوان یک مساله به فراخوان عمومی گذاشته شده است و از همه گروه‌های مرتبط اعم از شرکت‌ها، دانشگاه‌ها و مردم عادی، خواسته شده که در این زمینه طرح‌های خود را ارسال کنند. لذا به نظر می‌رسد، ضمن حفظ رویکرد دولتی در این سند، نظرات و انتظارات مجموعه ذینفعان در نظر گرفته شده است.

جواد درویش

پاییز ۹۹

ہمکاران گرامے



رئیس‌جمهور ترامپ در ۵ فوریه ۲۰۱۹ در سخنرانی سالانه خود در باره وضعیت کشور، بر اهمیت اطمینان از رهبری آمریکا در توسعه فن‌آوری‌های نوظهور از جمله هوش مصنوعی (AI) که «صنایع آینده» را تشکیل می‌دهند، تأکید کرد. رئیس‌جمهور ترامپ پس از انعکاس اهمیت این موضوع، در ۱۱ فوریه ۲۰۱۹ فرمان اجرایی ۱۳۸۵۹ را امضا کرد که طرح هوش مصنوعی آمریکا را بنا نهاد. این طرح، رویکردی دولتی برای حفظ رهبری آمریکا در هوش مصنوعی و حصول اطمینان از این است که هوش مصنوعی منافع مردم آمریکا را تأمین کرده و ارزش‌های کشور ما را منعکس می‌کند. اولین دستور در این فرمان اجرایی، مربوط به مؤسسات فدرال جهت اولویت‌بخشی به تحقیق و توسعه (R & D) هوش مصنوعی در بودجه‌بندی سالانه و فرآیند برنامه‌ریزی است. نسخه پیش‌رو «طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی: نسخه ۲۰۱۹»، اولویت‌های کلیدی برای سرمایه‌گذاری فدرال در بخش تحقیق و توسعه هوش مصنوعی را مورد تأکید قرار می‌دهد.

هوش مصنوعی فرصت‌های بی‌شماری را عرضه می‌دارد که منجر

به پیشرفت‌های ساختار شکن در مواردی چون مراقبت‌های بهداشتی بهتر، حمل‌ونقل امن‌تر و کارآمدتر، آموزش شخصی سازی شده، اکتشافات علمی مهم، ساخت و تولید بیشتر، افزایش عایدات تولید محصولات کشاورزی، پیش‌بینی بهتر آب‌وهوا و بسیاری موارد دیگر می‌شود. این مزایا عمدتاً ناشی از دهه‌ها سرمایه‌گذاری بلندمدت فدرال در مبانی هوش مصنوعی است که منجر به نظریه‌ها و رویکردهای جدیدی در سیستم‌های هوش مصنوعی و تحقیقات کاربردی شده و امکان استفاده از هوش مصنوعی را در کاربردهای عملی فراهم می‌کنند.

به واسطه سرمایه‌گذاری‌های قابل توجهی که توسط صنعت، دانشگاه و سازمان‌های غیرانتفاعی صورت می‌گیرند، چشم‌انداز تحقیق و توسعه هوش مصنوعی به طور فزاینده‌ای پیچیده می‌شود. به‌علاوه، پیشرفت‌های هوش مصنوعی به سرعت در حال ارتقا هستند. بنابراین، دولت فدرال باید به طور مداوم، اولویت‌های خود را برای سرمایه‌گذاری‌های تحقیق و توسعه هوش مصنوعی ارزیابی کند تا اطمینان حاصل کند که سرمایه‌گذاری‌ها به پیشبرد مرز دانش این حوزه ادامه داده و تکرار سرمایه‌گذاری‌های صنعتی غیرضروری نیستند.

در ماه اوت سال ۲۰۰۷، دولت به کمیته منتخب هوش مصنوعی دستور داد تا طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی را احیا کند. این فرآیند با صدور یک «درخواست اطلاعات»^۱ برای تقاضای اظهار نظر عموم مردم پیرامون روش‌های بازبینی یا بهبود در راهبرد آغاز شد. این نسخه به روز رسانی شده از طرح راهبردی،

1. Request for Information (RFI)

متأثر از پاسخ به این درخواست و همچنین بررسی موسسه مستقل است.

در این طرح، ۸ اولویت راهبردی شناسایی شده است. ۷ راهبرد اول، ادامه طرح ۲۰۱۶ و انعکاسی از تائید مجدد اهمیت این راهبردها و تاکید بر عدم نیاز به حذف هیچ‌یک از آنها از سوی پاسخ‌دهندگان، شامل عامه مردم و دولت، است. راهبرد ۸ جدید است و بر اهمیت روزافزون مشارکتهای مؤثر بین دولت فدرال و دانشگاهیان، صنعت، سایر نهادهای غیر فدرال و هم‌پیمانان بین‌المللی، در جهت ایجاد پیشرفت‌های ساختارشکن فناورانه در هوش مصنوعی و انتقال سریع آنها به قابلیت‌ها، تمرکز دارد.

در حالی که این طرح، دستور کار تحقیقاتی خاصی را برای سرمایه‌گذاری‌های آژانس فدرال تعریف نمی‌کند، توقعاتی را از سبد سرمایه‌گذاری کلی فدرال در تحقیق و توسعه به وجود می‌آورد. این راهبرد هماهنگ فدرال در تحقیق و توسعه هوش مصنوعی، به ایالات متحده کمک خواهد کرد تا آمریکا در پیشرفت‌های لبه دانش در زمینه هوش مصنوعی، پیشرو باشد. این امر موجب رشد اقتصاد ما خواهد شد و امنیت ملی و بهبود کیفیت زندگی را به همراه خواهد داشت.

با احترام،

مایکل کراتسیوس، معاون شرکت تکنالوجی پالیسی

ژوئن ۲۰۱۱

چکیده



هوش مصنوعی نوید فوق‌العاده‌ای است که تقریباً همه جوانب جامعه از جمله اقتصاد، مراقبت‌های بهداشتی، امنیت، قانون، حمل‌ونقل و حتی خود فناوری از آن منتفع خواهند شد. رئیس‌جمهور در ۱۱ فوریه ۲۰۱۹ دستور اجرایی ۱۳۸۵۹، یعنی «حفظ جایگاه رهبری آمریکا در هوش مصنوعی» را امضا کرد.^۱ این دستور، سرآغاز طرح هوش مصنوعی آمریکا بود، تلاشی مشترک در جهت ترویج و حفاظت از فناوری و نوآوری هوش مصنوعی در ایالات متحده آمریکا. به موجب این طرح، یک راهبرد کاملاً دولتی در همکاری و تعامل با بخش خصوصی، دانشگاهیان، عموم مردم و شرکای بین‌المللی همفکر به مرحله اجرا در می‌آید. از جمله اقدامات این طرح، دستورات کلیدی طرح به منظور فراخوان مؤسسات فدرال برای اولویت‌بندی سرمایه‌گذاری‌های تحقیق و توسعه در حوزه هوش مصنوعی، ارتقاء دسترسی به زیرساخت‌های سایبری با کیفیت بالا و داده‌ها، حصول اطمینان از پیشگامی کشور در تدوین استانداردهای فنی هوش مصنوعی و فراهم آوردن فرصت‌های تحصیلی و تربیتی به‌منظور آماده‌سازی نیروی کار آمریکا برای عرصه جدید هوش مصنوعی هستند.

1. <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>

در حمایت از طرح هوش مصنوعی آمریکا، طرح «راهبردی ملی تحقیق و توسعه هوش مصنوعی: نسخه ۲۰۱۹»، زمینه‌های دارای اولویت برای سرمایه‌گذاری فدرال در تحقیق و توسعه هوش مصنوعی را مشخص می‌کند. این نسخه جدید ۲۰۱۹ مبتنی بر اولین طرح راهبردی تحقیق و توسعه ملی هوش مصنوعی سال ۲۰۱۶ بوده و شامل پژوهش‌های جدید، نوآوری‌های فنی و سایر ملاحظات است که طی سه سال گذشته به منصفه ظهور رسیده‌اند. تهیه این نسخه جدید بر عهده محققان پیشرو و مدیران تحقیقاتی هوش مصنوعی دولت فدرال و نظرات جامعه مدنی گسترده‌تری همچون بسیاری از مؤسسات تحقیقاتی دانشگاهی پیشرو آمریکا، سازمان‌های غیرانتفاعی و شرکت‌های فناوری متعلق به بخش خصوصی بوده است. بازخوردهای این ذینفعان اصلی، مؤید ارتباط مستمر هر یک از بخش‌های طرح راهبردی سال ۲۰۱۶ است که در عین حال خواستار توجه هرچه بیشتر به قابل اعتماد بودن هوش مصنوعی برای مشارکت با بخش خصوصی و سایر موارد ضروری است.

طرح راهبردی تحقیق و توسعه ملی هوش مصنوعی: نسخه جدید ۲۰۱۹ مجموعه‌ای از اهداف تحقیقات هوش مصنوعی با بودجه فدرال را تثبیت کرده و هشت اولویت راهبردی زیر را در این زمینه مشخص می‌کند:

راهبرد ۱: بر روی تحقیقات هوش مصنوعی، سرمایه‌گذاری‌های بلندمدتی انجام دهید. سرمایه‌گذاری بر روی نسل بعدی هوش مصنوعی را که پیشران اکتشاف و بینش‌ساز است و ایالات متحده را قادر خواهد ساخت تا جایگاه خود را به عنوان رهبر جهان در عرصه هوش مصنوعی حفظ کند، در اولویت قرار دهید.

راهبرد ۲: روش‌های مؤثری برای همکاری بین انسان و هوش مصنوعی را توسعه دهید. درک نحوه ایجاد سیستم‌های هوش مصنوعی را که به‌طور مؤثری قابلیت‌های انسان را تکمیل و اضافه می‌کنند، افزایش دهید.

راهبرد ۳: پیامدهای اخلاقی، حقوقی و اجتماعی هوش مصنوعی را درک نموده و آن‌ها را بررسی کنید. در باره سیستم‌های هوش مصنوعی که دغدغه‌های اخلاقی، حقوقی و اجتماعی را از طریق سازوکارهای فنی یکپارچه می‌سازند، پژوهش نمایید.

راهبرد ۴: از ایمنی و امنیت سیستم‌های هوش مصنوعی اطمینان حاصل کنید. دانش مربوط به نحوه طراحی سیستم‌های هوش مصنوعی معتبر، قابل‌اتکا، ایمن و قابل‌اعتماد را توسعه دهید.

راهبرد ۵: مجموعه داده‌ها و محیط‌های عمومی مشترکی را برای آموزش و آزمون هوش مصنوعی ایجاد کنید. مجموعه داده‌ها و محیط‌های با کیفیت بالا و همچنین منابع آزمون و آموزش

را توسعه دهید و امکان دسترسی به آنها را فراهم کنید.

راهبرد ۶: فناوری‌های هوش مصنوعی را از طریق استانداردها و معیارهای مشخص، ارزیابی کنید. طیف گسترده‌ای از فنون ارزیابی هوش مصنوعی، از جمله استانداردها و ملاک‌های فنی را ایجاد کنید.

راهبرد ۷: به نیازهای نیروی کار ملی تحقیق و توسعه در زمینه هوش مصنوعی را بیشتر توجه کنید. فرصت‌های توسعه نیروی کار بخش تحقیق و توسعه را در جهت پرورش راهبردی نیروی کار مستعد برای هوش مصنوعی بهبود ببخشید.

راهبرد ۸: همکاری بین بخش‌های خصوصی و دولتی را به‌منظور تسریع پیشرفت در زمینه هوش مصنوعی گسترش دهید. با همکاری دانشگاهیان، صنایع، شرکای بین‌المللی و سایر نهادهای غیر فدرال، فرصت‌های سرمایه‌گذاری پایدار تحقیق و توسعه در زمینه هوش مصنوعی و تبدیل پیشرفت‌ها به قابلیت‌های عملی را ارتقا دهید.

مقدمه



مقدمه‌ای بر طرح راهبردی تحقیق و توسعه مایه هوش مصنوعی سال ۲۰۱۹

هوش مصنوعی، رایانه‌ها و سایر سیستم‌های خودکار را قادر می‌سازد تا وظایفی را انجام دهند که از لحاظ تاریخی نیازمند قوای شناختی انسان و چیزی است که معمولاً توانمندی‌های تصمیم‌گیری انسان تلقی می‌کنیم. در طول چند دهه گذشته، هوش مصنوعی شاهد پیشرفت‌های عظیمی بوده و امروزه نوید مراقبت‌های بهداشتی بهتر و دقیق‌تر، ارتقاء امنیت ملی، حمل‌ونقل بهبود یافته و آموزش مؤثرتر را می‌دهد؛ مواردی که تنها گوشه‌ای از مزایای آن هستند. افزایش توان محاسباتی، در دسترس بودن مجموعه داده‌های بزرگ و جریان داده‌ها و پیشرفت‌های الگوریتمی در یادگیری ماشین (ML) برای توسعه هوش مصنوعی، امکان ایجاد بخش‌های جدید اقتصاد و احیای صنایع را فراهم آورده‌اند. از آنجاکه صنایع بیشتری فناوری‌های بنیادین هوش مصنوعی را در اختیار می‌گیرند، این حوزه موجب تأثیر عمیق اقتصادی و بهبود کیفیت زندگی در سراسر جهان خواهد شد.

نسخه ۲۰۱۹	پاسخ‌های RFI موجب اطلاع‌رسانی به طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۹ می‌شود
<p>در سپتامبر ۲۰۱۸، دفتر هماهنگی ملی تحقیق و توسعه شبکه‌سازی و فناوری اطلاعات، به نمایندگی از سوی کمیته منتخب هوش مصنوعی، درخواستی برای اطلاعات (RFI) ارائه نمود^۱ و از همه طرف‌های علاقه‌مند به طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ تقاضای اطلاعات کرد. از سوی محققان، سازمان‌های تحقیقاتی، جوامع حرفه‌ای، سازمان‌های جامعه مدنی و افراد، نزدیک به ۵۰ پاسخ ارائه شد که به صورت آنلاین در دسترس هستند^۲.</p> <p>بسیاری از پاسخ‌های ارائه شده، تحلیل، سازمان‌دهی و رویکرد مطرح‌شده در طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ را تأیید کردند. تعداد قابل توجهی از پاسخ‌ها به اهمیت سرمایه‌گذاری در کاربرد هوش مصنوعی در حوزه‌هایی مانند تولید و زنجیره‌های تأمین؛ مراقبت‌های بهداشتی؛ تصویربرداری پزشکی؛ هواشناسی، آب‌شناسی، اقلیم‌شناسی و سایر حوزه‌های مرتبط؛ امنیت سایبری؛ آموزش؛ علوم فیزیکی داده-بنیان نظیر فیزیک انرژی‌بالا و حمل‌ونقل اشاره داشتند. به‌طور قطع، از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، این علاقه به کاربردهای تبدیلی فناوری‌های هوش مصنوعی افزایش یافته است. سایر مضامین مشترک منعکس‌شده در پاسخ‌های RFI عبارت بودند از: اهمیت توسعه سیستم‌های هوش مصنوعی قابل اعتماد (مثل انصاف، اخلاق، پاسخگویی) و شفافیت سیستم‌های هوش مصنوعی، مجموعه داده‌های گردآوری‌شده و در دسترس، ملاحظات مربوط به نیروی کار و مشارکت‌های عمومی-خصوصی در جهت راهبری تحقیق و توسعه هوش مصنوعی.</p>	

1. Request for Information
2. <https://www.nitrd.gov/news/RFI-National-AI-Strategic-Plan.aspx>
3. <https://www.nitrd.gov/nitrdgroups/index.php?title=AI-RFI-Responses2018->

اساساً موتور محرک این پیشرفت‌ها، سرمایه‌گذاری‌های دولت فدرال در تحقیق و توسعه هوش مصنوعی، متخصصان مؤسسات تحقیق و توسعه بی‌نظیر آمریکا و خلاقیت جمعی بسیاری از مبتکرترین شرکت‌های فناوری و کارآفرینان آمریکایی بوده است.

دولت فدرال در سال ۲۰۱۶ نخستین طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی را منتشر کرد که بر نویدبخشی عظیم هوش مصنوعی و نیاز به پیشرفت مداوم آن صحنه گذاشت. این طرح به چند دلیل تدوین شد، از جمله هدایت کشور در سرمایه‌گذاری‌های تحقیق و توسعه هوش مصنوعی، ارتقا قابلیت‌های هوش مصنوعی آمریکا و حصول اطمینان از اینکه این قابلیت‌ها موجب رفاه، امنیت و بهبود کیفیت زندگی مردم آمریکا در سال‌های آتی خواهد شد.

این طرح برای مؤسسات فدرالی که در هوش مصنوعی سرمایه‌گذاری می‌کنند، چندین حوزه کلیدی دارای اولویت را تعریف کرد. این حوزه‌های تمرکز یا راهبردها عبارت‌اند از: تداوم سرمایه‌گذاری‌های بلندمدت روی هوش مصنوعی، روش‌های مؤثر برای همکاری بین انسان و هوش مصنوعی، فهم و بررسی پیامدهای اخلاقی، قانونی و اجتماعی هوش مصنوعی، تضمین ایمنی و امنیت هوش مصنوعی، توسعه مجموعه داده‌ها و محیط‌های عمومی مشترک برای آموزش و آزمون هوش مصنوعی، اندازه‌گیری و ارزیابی فناوری‌های هوش مصنوعی از طریق استانداردها و ملاک‌ها و درک بهتر نیازهای نیروی کار تحقیق و توسعه هوش مصنوعی کشور. این کار آینده‌نگرانه بود: امروزه کشورهای جهان این طرح را دنبال کرده و نسخه‌های مخصوص به خودشان را منتشر کرده‌اند.

سه سال بعد از تدوین طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی، شاهد پیشرفت‌های سریعی در تحقیقات جدید، نوآوری‌های فنی و استقرار واقعی سیستم‌ها هستیم. دولت این نسخه ۲۰۱۹ از طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی را برای پرداختن به این پیشرفت‌ها، از جمله چشم‌انداز در حال تکامل هوش مصنوعی بین‌المللی، راه‌اندازی نمود.

شایان ذکر است که مطابق طراحی انجام شده، این نسخه ۲۰۱۹ از طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی، صرفاً به اولویت‌های تحقیق و توسعه مربوط به پیشرفت فناوری‌های هوش مصنوعی می‌پردازد. این طرح، اقدامات سیاست‌گذاری یا مقررات گذاری مربوط به اداره یا استقرار هوش مصنوعی را تشریح یا توصیه نمی‌کند، هرچند به‌طور حتم تحقیق و توسعه هوش مصنوعی، آگاهی‌بخش توسعه چارچوب‌های منطقی سیاستی و مقررات گذاری خواهد بود.

هوش مصنوعی به عنوان یک اولویت دولتی

از سال ۲۰۱۷، دولت با تأکید بر نقش خود در آینده آمریکا در چندین سند مهم سیاستی از جمله راهبرد امنیت ملی^۱، راهبرد دفاع ملی^۲ و یادداشت اولویت‌های بودجه‌ای تحقیق و توسعه FY ۲۰۲۰^۳، به اهمیت تحقیق و توسعه هوش مصنوعی اشاره کرده است.

در ماه مه سال ۲۰۱۸، دفتر سیاست‌گذاری علم و فناوری (OSTP^۴) میزبان نشست کاخ سفید در زمینه هوش مصنوعی برای صنعت آمریکا

1. <https://www.whitehouse.gov/wp-content/uploads/12/2017/NSS-Final0905-2017-18-12-.pdf>

2. <https://dod.defense.gov/Portals/1/Documents/pubs/-2018National-Defense-Strategy-Summary.pdf>

3. <https://www.whitehouse.gov/wp-content/uploads/07/2018/M22-18-.pdf>

4. Office of Science and Technology Policy

بود تا به بحث و تبادل نظر درباره هوش مصنوعی و سیاست‌های موردنیاز برای تحقق آن برای مردم آمریکا و حفظ رهبری آمریکا در عصر هوش مصنوعی پردازند. این نشست، بیش از ۱۰۰ مقام ارشد دولتی، کارشناسان فنی از مؤسسات دانشگاهی برتر، روسای آزمایشگاه‌های تحقیقاتی و رهبران تجاری آمریکا را گرد هم آورد. رئیس‌جمهور ترامپ در سخنرانی خود در پنجم فوریه ۲۰۱۹ بر اهمیت حصول اطمینان از رهبری آمریکا در توسعه فناوری‌های نوظهور، از جمله هوش مصنوعی تأکید کرد که صنایع آینده را تشکیل خواهد داد.

در یازدهم فوریه ۲۰۱۰، رئیس‌جمهور فرمان اجرایی ۱۳۸۵۹ را تحت عنوان «حفظ رهبری آمریکا در هوش مصنوعی»^۱ امضا کرد. این فرمان، شروعی بر طرح هوش مصنوعی آمریکا بود که در واقع تلاشی هماهنگ در جهت ترویج و حفاظت از فناوری هوش مصنوعی و نوآوری در ایالات متحده. این طرح با همکاری و تعامل با بخش خصوصی، دانشگاهیان، عموم مردم و شرکای بین‌المللی هم‌فکر، یک راهبرد کاملاً دولتی را به اجرا در می‌آورد. از جمله اقدامات این طرح، دستورهای کلیدی طرح به منظور فراخوان مؤسسات فدرال برای اولویت‌بندی سرمایه‌گذاری‌های تحقیق و توسعه هوش مصنوعی، ارتقا دسترسی به زیرساخت‌های سایبری و داده‌های با کیفیت بالا، حصول اطمینان از پیشگامی کشور در تدوین استانداردهای فنی هوش مصنوعی و فراهم کردن فرصت‌های تحصیلی و تربیتی، به منظور آماده‌سازی نیروی کار آمریکا برای عصر جدید هوش مصنوعی هستند.

ارتقا نسخه ۲۰۱۹ طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی در طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ توصیه شد که مؤسسات فدرال که عهده‌دار پیشبرد یا به‌کارگیری هوش مصنوعی هستند، برای شناسایی فرصت‌های تحقیق و توسعه و حمایت از هماهنگی مؤثر فعالیت‌های تحقیق و توسعه هوش مصنوعی فدرال، در تحقیقات درون دانشگاهی و برون دانشگاهی، با یکدیگر همکاری کنند. با توجه به اولویت‌بندی هوش مصنوعی از طرف دولت، شورای ملی علوم و فناوری (NSTC) چارچوب جدیدی را برای پیاده‌سازی این توصیه ایجاد کرده است که شامل سه زیرگروه خاص NSTC است که متشکل از اعضای از همه مؤسسات تحقیق و توسعه فدرال هستند: (۱) مدیریت ارشد و چشم‌انداز راهبردی، (۲) برنامه‌ریزی عملیاتی و پیاده‌سازی تاکتیکی و (۳) تخصص تحقیقاتی و فنی. این زیرگروه‌ها عبارت‌اند از:

- کمیته منتخب هوش مصنوعی^۲ که متشکل از روسای ادارات و مؤسساتی است که اصولاً مسئول بخش دولتی تحقیق و توسعه هوش مصنوعی هستند؛ در زمینه اولویت‌های بین بخشی تحقیق و توسعه هوش مصنوعی به دولت مشاوره می‌دهند؛ ایجاد مشارکت‌های فدرال با صنایع و دانشگاهیان را مدنظر قرار می‌دهد؛ ساختارهایی را در جهت بهبود برنامه‌ریزی دولت و هماهنگی تحقیق و توسعه در حوزه هوش مصنوعی ایجاد می‌کند؛ فرصت‌های ارتقا داده‌های فدرال و منابع محاسباتی جهت پشتیبانی از زیست‌بوم تحقیق و توسعه ملی هوش مصنوعی ما را شناسایی می‌کند و از رشد نیروی کار فنی و ملی هوش مصنوعی پشتیبانی می‌نماید.
- کمیته فرعی NSTC در مورد یادگیری ماشین و هوش مصنوعی

1. National Science and Technology Council
 2. <https://www.whitehouse.gov/wp-content/uploads/05/2018/Summary-Report-of-White-House-AI-Summit.pdf>

(MLAI)، متشکل از رهبران و مدیران هوش مصنوعی موسسه، به عنوان بازوی عملیاتی و اجرایی کمیته منتخب عمل کرده و مسئولیت‌های زیر را دارد: انجام وظایف محوله از سوی کمیته منتخب؛ ایجاد و حفظ «طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی»؛ شناسایی و بررسی مسائل سیاستی مهم که مرتبط با تحقیق، آزمایش، استانداردها، آموزش، پیاده سازی، کمک‌رسانی هوش مصنوعی و حوزه‌های مرتبط و فعالیت‌های مربوطه است.

• گروه کاری بین بخشی تحقیق و توسعه هوش مصنوعی که تحت نظارت کمیته فرعی تحقیق و توسعه شبکه‌سازی و فناوری اطلاعات (NITRD^۲) فعالیت می‌کند، متشکل از مدیران برنامه پژوهشی و متخصصان فنی دولت فدرال است. این گروه به کمیته فرعی MLAI گزارش می‌دهد؛ به هماهنگ‌سازی فعالیت‌های ذیل تحقیق و توسعه هوش مصنوعی کمک می‌کند؛ به‌عنوان محفل کاری^۳ تحقیق و توسعه هوش مصنوعی بین بخشی عمل می‌کند و از طریق ضمیمه سالانه کمیته فرعی NITRD به بودجه ریاست جمهوری، هزینه‌های دولتی تحقیق و توسعه هوش مصنوعی را گزارش می‌کند.

در سپتامبر ۲۰۱۸، کمیته منتخب، نسخه جدید برنامه راهبردی ۲۰۱۶ را راه‌اندازی کرد که با یک RFI شروع می‌شد. این RFI در پی گردآوری نظرات جمعی مبنی بر این بود که آیا هفت راهبرد طرح ۲۰۱۶ نیازمند اصلاح یا جایگزینی هستند یا خیر و این کار چگونه باید انجام شود؟ (یادداشت کناری را ببینید). ادارات و مؤسسات فدرال که تحقیق و توسعه هوش مصنوعی را انجام می‌دهند یا بودجه آن را تأمین می‌کنند، بصورت مستقل عهده‌دار ارزیابی فعالیت‌هایشان هستند.

1. Machine Learning and Artificial Intelligence (Subcommittee of the NSTC)
2. Networking and Information Technology R&D (NITRD)
3. community of practice

مروری بر موارد به روز شده نسخه ۲۰۱۹، نسبت به نسخه ۲۰۱۶ طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی

کمیته منتخب هوش مصنوعی، کمیته فرعی یادگیری ماشینی و هوش مصنوعی NTSC و گروه کاری NITRD، با همدیگر نظرات واسله در باره طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی را بررسی کردند. هر یک از هفت حوزه تمرکز اصلی یا راهبردهای طرح ۲۰۱۶ توسط چندین پاسخ‌دهنده عمومی و دولتی مورد تأیید قرار گرفت و هیچ دلیل معقولی برای حذف هیچ راهبردی اقامه نشد. این راهبردها که در نسخه ۲۰۱۹ طرح راهبردی به روز شدند تا انعکاسی از وضع موجود باشند، عبارت‌اند از:

راهبرد ۱: سرمایه‌گذاری بلندمدت در پژوهش هوش مصنوعی؛

راهبرد ۲: ارتقا روش‌های مؤثر برای همکاری انسان و هوش مصنوعی؛

راهبرد ۳: درک پیامدهای اخلاقی، قانونی و اجتماعی هوش مصنوعی و بررسی آن‌ها؛

راهبرد ۴: حصول اطمینان از ایمنی و امنیت سیستم‌های هوش مصنوعی؛

راهبرد ۵: توسعه مجموعه داده‌ها و محیط‌های عمومی برای آموزش و آزمون هوش مصنوعی؛

راهبرد ۶: اندازه‌گیری و ارزیابی فناوری‌های هوش مصنوعی از طریق استانداردها و ملاک‌ها؛ و

راهبرد ۷: درک بهتر نیازهای نیروی کار ملی تحقیق و توسعه هوش مصنوعی.

با توجه به افزایش سریع تحقیق و توسعه هوش مصنوعی در بخش خصوصی و پذیرش سریع هوش مصنوعی از سوی صنعت، بسیاری از پاسخ‌ها به RFI، خواهان مشارکت بیشتر تحقیق و توسعه دولت فدرال با بخش خصوصی بودند. در نتیجه، در نسخه ۲۰۱۹ طرح، یک راهبرد هشتمی گنجانده شده است:

راهبرد ۸: گسترش مشارکت بخش عمومی و خصوصی برای شتاب بخشی پیشرفت‌ها در هوش مصنوعی.

بازخورد گرفته شده از موسسات دولتی و فدرال، چالش‌های خاصی را برای توسعه و پذیرش بیشتر هوش مصنوعی آشکار کرد. بسیاری از این چالش‌ها که موسسات متعددی را درگیر کرده است، در باره روش‌هایی که این طرح راهبردی طی آن‌ها می‌تواند مسیر تحقیق و توسعه هوش مصنوعی را در آمریکا هدایت کند، بینش بهتری را فراهم می‌کنند. بسیاری از این چالش‌ها با مضامین مطرح شده در دستور ریاست جمهوری در باره حفظ جایگاه پیشرو آمریکا در هوش مصنوعی ارتباط نزدیکی دارند. مثال‌های آن شامل موارد زیر هستند:

- «تحقیقات در مرز (دانش).» با وجود اینکه در چند سال گذشته یادگیری ماشین قابلیت‌های جدید خیره‌کننده‌ای را به ارمغان آورده است، برای گسترش مرزهای یادگیری ماشینی و همچنین برای فائق آمدن بر چالش‌های فنی هوش مصنوعی، نیاز به تحقیقات مستمر وجود دارد (راهبرد ۱).
- «تأثیر مثبت». با رشد قابلیت‌های هوش مصنوعی، آمریکا باید بر توسعه روش‌های جدید تأکید بیشتری داشته باشد تا اطمینان حاصل

کند که اثرات هوش مصنوعی در آینده کاملاً مثبت هستند (راهبردهای ۱، ۳ و ۴)

- «اعتماد و توضیح‌پذیری»^۱. هوش مصنوعی قابل‌اعتماد، نیاز به هوش مصنوعی توضیح‌پذیر دارد، به‌خصوص که سیستم‌های هوش مصنوعی به لحاظ مقیاس و پیچیدگی در حال رشد هستند. این امر مستلزم درک جامعی از سیستم هوش مصنوعی از سوی کاربر و طراح انسانی دارد (راهبردهای ۱، ۲، ۳، ۴ و ۶)

- «ایمنی و امنیت». محققان برای حفظ امنیت سیستم‌های هوش مصنوعی و داده‌های مورد استفاده آن‌ها، باید روش‌هایی را ابداع کنند که دولت بتواند توسط این فناوری، فرصت‌هایی را فراهم کند و در عین حال محرمانگی و ایمنی این فناوری را حفظ کند (راهبردهای ۴، ۵ و ۶).

- «استانداردهای فنی». هم‌زمان با اینکه دولت، فنونی را برای گسترش توانایی‌ها و تضمین‌های هوش مصنوعی توسعه می‌دهد، باید آن‌ها را آزمون کرده و محک بزند؛ وقتی این فنون آماده بهره‌برداری شدند، باید به استانداردهای فنی جهانی تبدیل شوند (راهبرد ۶).

- «قابلیت نیروی کار». دستیابی به این اهداف، مستلزم رشد نیروی کار ماهر تحقیق و توسعه هوش مصنوعی است که در حال حاضر محدود بوده و تقاضا برای آن زیاد است. آمریکا باید در آموزش و جذب نیروی کار ماهری که برای پیشگامی در تحقیق و کاربرد هوش مصنوعی لازم دارد، خلاق باشد (راهبرد ۷).

- «شراکت». پیشرفت در تحقیق و توسعه هوش مصنوعی به‌طور فزاینده ای مستلزم مشارکت مؤثر بین دولت فدرال با دانشگاهیان، صنعت

1. explainability

و سایر نهادهای غیر فدرال است تا پیشرفت‌های بنیادی فناورانه در هوش مصنوعی ایجاد کنند و آن‌ها را سریعاً به قابلیت‌ها تبدیل نمایند (راهبرد ۸).

• «همکاری با هم‌پیمانان». بعلاوه، این طرح به اهمیت همکاری بین‌المللی برای اجرای موفقیت‌آمیز اهداف مذکور واقف است؛ در عین حال از شرکتهای تحقیق و توسعه هوش مصنوعی آمریکایی در قبال رقبای راهبردی و کشورهای رقیب متخاصم حمایت می‌کند.

ساختار نسخه ۲۰۱۹ طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی

این نسخه جدید از طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی در برگیرنده متن اصلی نسخه ۲۰۱۶ است و شامل بخش‌های زیر در باره راهبرد تحقیق و توسعه (با تغییرات جزئی) و همان جمله‌بندی اولیه نسخه ۲۰۱۶ از هفت راهبرد اول است. برای هر راهبرد، «نسخه ۲۰۱۹ از طرح راهبردی ملی تحقیق و توسعه ۲۰۱۶ در مربع‌های سایه‌دار در بالای هفت راهبرد اصلی درج می‌شوند. این کار تأکیدی است بر الزامات نسخه جدید و/یا حوزه‌های مورد تأکید جدید برای راهبردها». متن زیر مربع‌های سایه‌دار، درست مثل متن و طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ بوده و نشانگر مشاهدات و محتوایی هستند که امروز حائز اهمیت‌اند (توجه داشته باشید که ممکن است طی این مدت، برخی از جزئیات اصلی قدیمی شده باشند). علاوه بر این، همان‌طور که قبلاً اشاره شد، در نسخه جدید ۲۰۱۹، در جهت توسعه مشارکت‌های دولتی-

خصوصی در تحقیق و توسعه هوش مصنوعی، راهبرد هشتم جدیدی نیز افزوده شده است.

راهبرد تحقیق و توسعه هوش مصنوعی

اولویت‌های پژوهشی مطرح شده در این طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی به حوزه‌هایی می‌پردازد که بعید است صنعت به تنهایی به آن‌ها رسیدگی کند و لذا این اولویت‌ها از مواردی هستند که به احتمال زیاد از سرمایه‌گذاری فدرال بهره‌مند خواهند شد. این اولویت‌ها، همه هوش مصنوعی را در بر می‌گیرند و شامل: نیازهای مشترک حوزه‌های فرعی ادراک، استدلال/برنامه‌ریزی خودکار، سیستم‌های شناختی، یادگیری ماشین، پردازش زبان طبیعی، رباتیک و شاخه‌های مرتبط. به دلیل وسعت هوش مصنوعی، این اولویت‌ها به جای تمرکز بر چالش‌های خاص در حوزه‌های فرعی، کل طیف را در بر می‌گیرند. برای اجرای این طرح، باید نقشه راه‌های تفصیلی توسعه داده شوند تا شکاف‌های بالقوه مشخص شوند.

یکی از مهم‌ترین اولویت‌های تحقیقاتی فدرال که در راهبرد ۱ مطرح شده، به تحقیقات بلندمدت در هوش مصنوعی برای پیشبرد ابداعات و شهودها مربوط است. بسیاری از سرمایه‌گذاری‌های دولت فدرال آمریکا در تحقیقات بنیادی که با ریسک و پاداش زیاد^۱ انجام شده، به پیشرفت‌های فناورانه انقلابی منجر شده که امروز به آن‌ها وابسته هستیم، از جمله اینترنت،^۲ GPS، تشخیص گفتار در گوشی‌های هوشمند، نمایشگرهای قلب، صفحات خورشیدی، باتری‌های پیشرفته، درمان سرطان و موارد

۱. تحقیقات با ریسک و پاداش زیاد، اشاره به پژوهش‌های مبتکرانه‌ای است که از لحاظ فکری چالش‌برانگیز است اما برای تأثیر عمیق و تحول‌برانگیز بر این حوزه مطالعاتی، توان بالقوای دارند.

متعدد دیگر. تعهدات هوش مصنوعی تقریباً همه ابعاد جامعه را تحت تأثیر قرار می‌دهد و قابلیت زیادی برای تامین منافع اجتماعی و اقتصادی دارد. بنابراین، آمریکا برای حفظ رهبری خود در جهان در این حوزه، باید سرمایه‌گذاری‌های خود را بر تحقیقات هوش مصنوعی بنیادی و بلندمدت با اولویت بالا، متمرکز کند.

بسیاری از فناوری‌های هوش مصنوعی به کمک انسان و دوشادوش او کار خواهند کرد؛ لذا چالش‌های مهمی در چگونگی ایجاد سیستم‌های هوش مصنوعی که به نحوه مفیدی با انسانها کار می‌کنند، چالش‌های مهمی به وجود می‌آیند^۱. با ظهور سیستم‌های هوش مصنوعی که قابلیت‌های انسانی را افزایش داده و بهبود می‌بخشند، دیوارهای بین انسان و هوش مصنوعی به تدریج در حال فروریختن است. همان‌گونه که در راهبرد ۲ مطرح شده، برای توسعه روش‌های مؤثری برای تعامل و همکاری انسان و هوش مصنوعی، به تحقیقات بنیادی نیاز داریم.

پیشرفت‌های هوش مصنوعی فواید مثبت زیادی برای جامعه فراهم می‌کنند و در حال ارتقا دست برتر آمریکا هستند^۲. با این حال، همانند اکثر فناوری‌های تحول ساز، هوش مصنوعی نیز با برخی چالشهای اجتماعی، از مشاغل و اقتصاد گرفته تا مسائل مربوط به ایمنی، سؤالات اخلاقی و حقوقی، همراه است. بنابراین، همان‌طور که در راهبرد ۳ درج شده، هم‌زمان با توسعه علم و فناوری هوش مصنوعی، دولت فدرال باید در تحقیقاتی سرمایه‌گذاری کند که درک بهتری از پیامدهای هوش مصنوعی در همه این حوزه‌ها حاصل می‌کنند و سعی کند با توسعه سیستم‌های هوش مصنوعی که با اهداف اخلاقی، قانونی و اجتماعی هم‌راستا هستند، این پیامدها را برطرف کند.

۱. مراجعه کنید به گزارش ۲۰۱۶ صدسال مطالعه بر روی هوش مصنوعی که به مصارف و تأثیرات احتمالی هوش مصنوعی در سال ۲۰۳۰ می‌پردازد؛ <https://ai100.report-2016.stanford.edu>.
۲. جی. فورمن، «با این زمان متفاوتی است؟ فرصت‌ها و چالش‌های هوش مصنوعی»، نظرات شورای مشاوران اقتصادی، دانشگاه نیویورک: همایش اکنون هوش مصنوعی، ۷ جولای ۲۰۱۶.

یک نقصان اساسی در فناوری هوش مصنوعی فعلی، نبود یک روش‌شناسی است که ایمنی و عملکرد پیش‌بینی پذیر سیستم‌های هوش مصنوعی را تضمین کند. به دلیل پیچیدگی غیرمعمول و ماهیت در حال تحول سیستم‌های هوش مصنوعی، تضمین ایمنی این سیستم‌ها یک چالش تلقی می‌شود. پژوهش‌های متعددی، بررسی این چالش ایمنی را در اولویت خود قرار داده‌اند. اول، راهبرد ۴ بر نیاز به سیستم‌های قابل توضیح و شفاف تأکید می‌کند که مورد اعتماد کاربران خود هستند و به شیوه‌ای عمل می‌کنند که مورد قبول کاربران است و می‌توانند تضمین کنند که مطابق میل کاربر عمل می‌کنند. قابلیت‌های بالقوه و پیچیدگی سیستم‌های هوش مصنوعی، توأم با غنای تعاملات میان کاربران انسانی و محیط، موجب می‌شوند تا سرمایه‌گذاری در تحقیقاتی که امنیت و کنترل فناوری هوش مصنوعی را ارتقا می‌دهند، حائز اهمیت باشد. راهبرد ۵ از دولت فدرال می‌خواهد تا برای پیشبرد تحقیقات هوش مصنوعی و امکان انجام مقایسه مؤثرتر بین راه‌حل‌های جایگزین، روی مجموعه داده‌های عمومی برای آموزش و آزمایش هوش مصنوعی سرمایه‌گذاری کند.

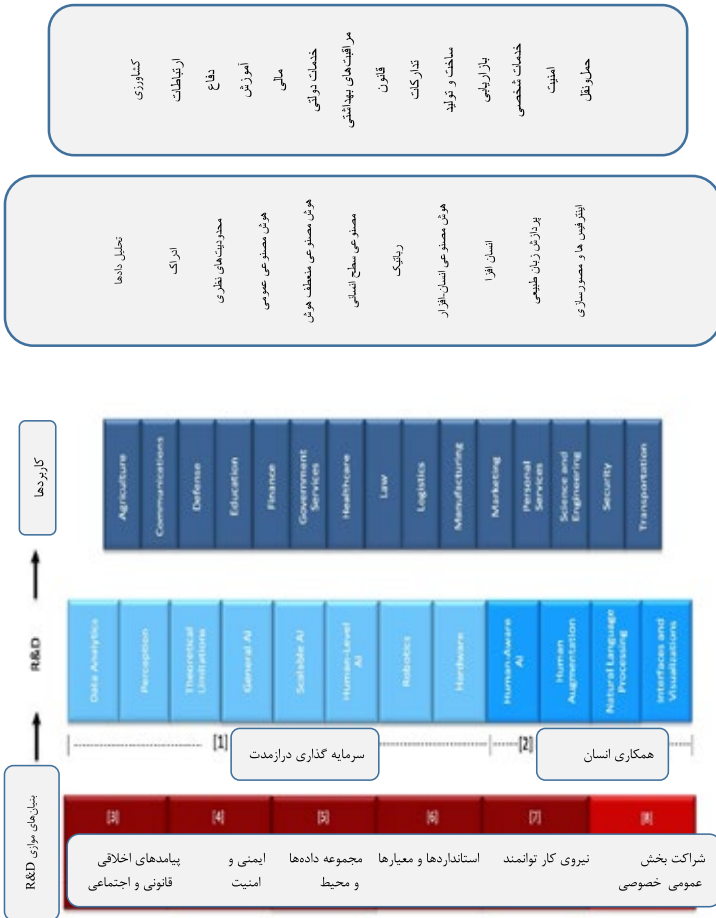
راهبرد ۶ به بحث در این باره می‌پردازد که استانداردها و ملاک‌ها چگونه می‌توانند تحقیق و توسعه را بر تعریف پیشرفت، شکاف‌های کوچک و به پیش راندن راه‌حل‌های نوآورانه برای مشکلات و چالش‌های خاص متمرکز نمایند. استانداردها و ملاک‌ها برای سنجش و ارزیابی سیستم‌های هوش مصنوعی و اطمینان از اینکه فناوری‌های هوش مصنوعی به اهداف مهمی برای کارکرد و تعامل پذیری میرسند، ضروری هستند.

در نهایت، رواج روزافزون فناوری‌های هوش مصنوعی در تمام بخش‌های جامعه، فشارهای جدیدی را بر متخصصان تحقیق و توسعه هوش مصنوعی وارد می‌کند. برای هسته مرکزی دانشمندان و مهندسين هوش مصنوعی که درک عمیقی از فناوری دارند و به دنبال پیشبرد مرزهای دانش در این حوزه هستند و می‌توانند ایده‌های جدیدی تولید کنند، فرصت‌های زیادی وجود دارد. دولت برای کسب اطمینان از وجود مسیری هموار برای حضور افراد مستعد در حوزه هوش مصنوعی، باید اقداماتی انجام دهد. راهبرد ۷ به این چالش می‌پردازد.

شکل ۱ (به روز شده در نسخه ۲۰۱۹ طرح) تصویری گرافیکی از سازمان کلی این طرح راهبردی تحقیق و توسعه هوش مصنوعی را نشان می‌دهد. در امتداد مستطیل‌های ردیف پایینی مربع‌ها، اصول زیربنایی و موازی^۱ وجود دارند که بر توسعه تمامی سیستم‌های هوش مصنوعی تأثیر می‌گذارند. این اصول در راهبردهای ۳-۷ و راهبرد جدید ۸ توضیح داده می‌شوند. لایه بالاتر (مستطیل‌های ردیف میانی) شامل بسیاری از حوزه‌های تحقیقاتی است که برای پیشبرد هوش مصنوعی ضروری هستند. این حوزه‌های تحقیق و توسعه (شامل تحقیقات پایه الهام گرفته شده از کاربرد^۲) در راهبردهای ۱-۲ مطرح شده‌اند^۳. در امتداد مستطیل‌های ردیف بالایی، نمونه‌هایی از برنامه‌هایی هستند که انتظار می‌رود از پیشرفت در هوش مصنوعی حاصل شوند. این مؤلفه‌های طرح راهبردی تحقیق و توسعه هوش مصنوعی، با همدیگر چارچوبی سطح بالایی را برای سرمایه‌گذاری‌های فدرال تعریف می‌کنند که می‌تواند به پیشرفت‌های مؤثر در حوزه هوش مصنوعی و منافع اجتماعی مثبت منجر شوند.

1. crosscutting
2. use-inspired

۳. در سراسر این سند، «تحقیق پایه» شامل هر دو تحقیق اولیه محض و پژوهش اصلی الهام گرفته از کاربرد می‌شود (همان ربع معروف پاستور که در کتاب ۱۹۹۷ دونالد استاکس که به همین نام است). تحقیقات پایه، تحقیقاتی هستند که برای جامعه کاربردی دارند که در ذهن می‌ماند. به عنوان مثال، سرمایه‌گذاری بنیادی NIH در فناوری اطلاعات، اغلب تحقیقات پایه الهام گرفته شده از کاربرد نامیده می‌شود.



شکل ۱. سازماندهی طرح راهبردی تحقیق و توسعه هوش مصنوعی (نسخه ۲۰۱۹، شامل راهبرد ۸). ترکیبی از مبانی موازی تحقیق و توسعه (در ردیف پایین) برای همه تحقیقات هوش مصنوعی مهم هستند. بسیاری از حوزه‌های تحقیق و توسعه هوش مصنوعی (در ردیف میانی) می‌توانند بر این مبانی نظری بنا شوند تا بر طیف گسترده‌ای از کاربردهای اجتماعی تأثیر بگذارند (در ردیف بالا). اعداد داخل پرانتز، شماره راهبردی را در این طرح نشان می‌دهند که هر عنوان را بیشتر توضیح می‌دهند. ترتیب راهبردها، براساس اهمیت بیشتر نیست.

بخش اول

راهبردا: سرمایه‌گذاری بلندمدت
در تحقیقات هوش مصنوعی



حفظ سرمایه‌گذاری‌های بلندمدت در تحقیقات بنیادی هوش مصنوعی، نسخه ۲۰۱۹

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی نسخه ۲۰۱۶، قابلیت‌های جدید قدرتمندی (در درجه اول، کاربردهای ML تا وظایف خوش-تعریف) بروز پیدا کرده است. این قابلیت‌ها در طیف وسیعی از کاربردها مانند طبقه‌بندی توالی‌های ژنتیکی^۱، مدیریت منابع طیف بی‌سیم محدود^۲، تفسیر تصاویر پزشکی^۳ و درجه‌بندی سرطان‌ها^۴ تأثیر خود را نشان داده‌اند. این پیشرفت‌های سریع، مستلزم چندین دهه تحقیق هستند تا فناوری‌ها و کاربردها به بلوغ برسند^۵. دولت فدرال برای حفظ این پیشرفت در ML و برای دستیابی به پیشرفتهایی در سایر حوزه‌های هوش مصنوعی و همچنین تلاش برای رسیدن به اهداف بلندمدت هوش مصنوعی، باید به ترغیب تحقیقات بنیادین و بلندمدت در ML و هوش مصنوعی ادامه دهد. این تحقیق منجر به ظهور فناوری‌های تحولی و به نوبه خود، پیشرفت‌های بنیان شکن در تمامی بخش‌های جامعه خواهد شد.

1. <https://ai.googleblog.com/12/2017/deepvariant-highly-accurate-genomes.html> , <https://irp.nih.gov/catalyst/v26i4/machine-learning>).
2. <https://www.spectrumcollaborationchallenge.com>
3. <https://news-medical.net/news/20190417/Workshop-explores-the-future-of-artificial-intelligence-in-medical-imaging.aspx>
4. <https://www.nature.com/articles/nature21056>
5. <https://www.nitrd.gov/rfi/ai/2018/AI-RFI-Response-2018-Yolanda-Gil-AAAI.pdf>

بسیاری از پیشرفت‌های کنونی در کارهای تخصصی و خوش-تعریف صورت گرفته که اغلب توسط ML طبقه‌بندی، شناخته و خطی سازی (یعنی «سیستم‌های هوش مصنوعی محدود») شده‌اند. تحقیقات انجام شده در این زمینه اشاره کرده‌اند که برای ادامه این پیشرفت‌ها در ML، لازم است تا در تحقیقات بنیادی، سرمایه‌گذاری‌های بلندمدتی صورت بگیرد. علاوه بر این، برای تحقق کامل چشم‌انداز سیستم‌های «هوش مصنوعی همه‌منظوره» که انعطاف‌پذیری و تطبیق‌پذیری هوش انسانی را در دامنه وسیعی از حوزه‌های شناختی نشان می‌دهند، به تلاش‌های مداوم، نیازمندیم^۱.

برای یادگیری تعاملی و مستمر، برقراری ارتباط بین ادراک و توجه و ادغام مدل‌های آموخته شده در ساختار استدلال جامع، تأکید بر توسعه قابلیت‌های بیشتر ML ضرورت دارد^۲. فراتر از ML، تحقیقات مهمی در دیگر حوزه‌های اصلی هوش مصنوعی شامل استدلال عقلانی و حل مسئله، استدلال احتمالی، بهینه‌سازی ترکیبی، دانش‌پدیدی، طرح‌ریزی و زمان‌بندی، پردازش زبان طبیعی، تصمیم‌گیری و تعامل انسان و ماشین موردنیاز هستند. پیشرفت‌ها در این حوزه‌ها، به نوبه خود، رباتیک مشترک و سیستم‌های مشترک و کاملاً خودکار را توانمند می‌سازند (راهبرد ۲ را ببینید). چالش بزرگ در فهم و شناخت هوش انسانی، مستلزم سرمایه‌گذاری قابل توجهی در منابع و زیرساخت مشترک است^۳. توافق گسترده برای سرمایه‌گذاری بنیادی در پیشرانان ML و هوش مصنوعی، یعنی منشأ و کیفیت داده‌ها، پارادایم‌ها و پلتفرم‌های نرم‌افزاری و سخت‌افزاری بدیع

1. https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fml_singles.pdf , <http://cdn.aiindex.org/2018/AI20%Index20%202018%Annual20%Report.pdf> , <https://cra.org/ccc/visioning/visioning-activities/-2018activities/artificial-intelligence-roadmap/> , <https://www.microsoft.com/en-us/research/research-area/artificial-intelligence>

2. <https://cra.org/ccc/events/artificial-intelligence-roadmap-workshop-3-learning-and-robotics>

3. <https://www.nitrd.gov/rfi/ai/2018/AI-RFI-Response-2018-Yolanda-Gil-AAAI.pdf>

و امنیت سیستم‌های هوش مصنوعی نیز وجود دارد. برای مثال، همان‌طور که نرم‌افزار هوش مصنوعی به شکل روزافزونی وظایف پیچیده‌ای را در تمام جنبه‌های زندگی روزمره و تمام بخش‌های اقتصادی انجام می‌دهد، برای ایجاد تحول در برآورده کردن بهره‌وری و کیفیت نرم‌افزار و ضرورت‌های پایداری، به توسعه پارادایم‌های موجود توسعه نرم‌افزار نیاز خواهد بود.

سرمایه‌گذاری‌های اخیر فدرال به حوزه‌های پژوهشی بنیادی ML و AI (مربع کناری را ببینید) و استفاده از ML و AI در تمام بخش‌های کاربردی مختلف از جمله دفاع، امنیت، انرژی، حمل‌ونقل، بهداشت، کشاورزی و مخابرات اولویت داده‌اند. در نهایت، فناوری هوش مصنوعی برای بررسی و تحلیل طیف گسترده‌ای از چالش‌های بلندمدت، مانند ساخت سیستم‌های مراقبت‌های بهداشتی پیشرفته، سیستم حمل‌ونقل هوشمند قوی و شبکه‌های انعطاف‌پذیر انرژی و مخابرات حیاتی هستند.

برای آنکه کاربردهای هوش مصنوعی رواج یابند، باید توضیح‌پذیر و قابل‌درک باشند (راهبرد ۳ را ببینید). این چالش‌ها، به‌ویژه برای ترویج روابط مشترک انسان-هوش مصنوعی، حائز اهمیت هستند (راهبرد ۲ را ببینید). امروزه توانایی درک و تحلیل تصمیمات سیستم‌های هوش مصنوعی و اندازه‌گیری دقت، قابلیت اطمینان و قابلیت بازتولید آن‌ها محدود است. برای افزایش اعتماد به سیستم‌های هوش مصنوعی، به سرمایه‌گذاری‌های پایدار تحقیق و توسعه نیاز است تا اطمینان حاصل شود که آن‌ها نیازهای جامعه را برآورده می‌کنند و به قدر کافی به الزامات مربوط به استحکام، انصاف، توجیه‌پذیری و امنیت می‌پردازند.

برای ادامه و گسترش پیشرفت‌های فنی کنونی، تعهد بلندمدت به تحقیق و توسعه هوش مصنوعی ضروری است و به‌طور وسیع‌تری این اطمینان را فراهم می‌کنند که هوش مصنوعی، تجربه انسانی را تقویت می‌کند. در حقیقت، فرمان ۲۰۱۹ ریاست جمهوری برای حفظ رهبری آمریکا در هوش مصنوعی یادآور می‌شود که:

روسای مؤسسات اجرایی که مجری یا تأمین‌کننده بودجه تحقیق و توسعه نیز هستند (مؤسسات تحقیق و توسعه هوش مصنوعی)، باید هوش مصنوعی را به عنوان یک اولویت تحقیق و توسعه موسسه در نظر بگیرند که مناسب مأموریت‌های مربوطه موسسه است ... روسای این مؤسسات باید این اولویت را در هنگام تهیه طرح‌های بودجه و طرح‌ریزی برای استفاده از وجوه در سال‌های مالی جاری و آینده در نظر بگیرند. روسای این مؤسسات باید اقدامات مدیریتی مناسب را نیز برای افزایش تمرکز بر هوش مصنوعی در نظر داشته باشند. در حوزه‌هایی با بازده بلندمدت بالقوه، به سرمایه‌گذاری‌های تحقیقاتی هوش نیاز داریم. در حالی که یکی از مؤلفه‌های مهم تحقیقات بلندمدت، تحقیقات تدریجی با پیامدهای قابل پیش‌بینی است، سرمایه‌گذاری‌های پایدار بلندمدت در تحقیقات با ریسک بالا می‌توانند به بازده‌هایی با پاداش بالا منجر شوند. این بازده‌ها را می‌توان در پنج سال، ده سال یا بیشتر در نظر گرفت. گزارش ۲۰۱۲ شورای ملی تحقیقات بر نقش حیاتی سرمایه‌گذاری‌های فدرال در تحقیقات بلندمدت تأکید دارد و به «دوره نهفتگی طولانی و غیر قابل پیش‌بینی بین کشف اولیه و استقرار تجاری که مستلزم کار مداوم و سرمایه‌گذاری مداوم است» اشاره دارد.^۱ در ضمن، این گزارش خاطر

۱. کرسی مخابرات و علوم رایانه شورای تحقیقات ملی، ادامه نوآوری در فن‌آوری اطلاعات (انتشارات National Academies، واشینگتن دی. سی، ۲۰۱۲)؛ <https://doi.org/10.17226/13427>.

نشان می‌کند که «زمان لازم از مرحله مفهوم پردازی تا رسیدن به بازار موفق، اغلب در بازه های یک یا چند دهه در نظر گرفته می‌شود.» مثال‌هایی از تلاش‌های تحقیقات بنیادی پایدار که مستند شده بازده خوب داشته اند، شامل وب جهانی و یادگیری عمیق هستند. بنیان‌های اصلی هر دو مورد، در دهه ۱۹۹۰ بنا شدند؛ تنها بیش از ۳۰ سال از تلاش‌های تحقیقاتی می‌گذشت که این ایده‌ها به فناوری‌های تحول سازی مبدل شدند که امروزه در بسیاری از مقوله‌های هوش مصنوعی شاهد آن‌ها هستیم.

تحقیقات هوش مصنوعی بنیادی بلندمدت: برنامه‌های تحقیق و توسعه اخیر موسسه

از زمان انتشار طرح ۲۰۱۶ راهبردی ملی تحقیق و توسعه هوش مصنوعی، مؤسساتی، برنامه‌های تحقیق و توسعه هوش مصنوعی را در جهت راهبرد ۱ آغاز کرده‌اند:

- NSF^۱ به تأمین بودجه تحقیقات بنیادی در هوش مصنوعی، گسترش یادگیری ماشین، استدلال و بازنمایی، بینایی رایانه، عصب‌شناسی محاسباتی، گفتار و زبان، روباتیک و سیستم‌های چندعاملی ادامه داده است. NSF فرصت‌های سرمایه‌گذاری مشترک جدیدی را با مؤسسات دیگر -بخصوص با DARPA^۲ در حوزه عملکرد بالا، سخت‌افزار کارآمد از لحاظ انرژی برای یادگیری ماشین بلادرنگ^۳ و با USDA^۴-NIFA^۵ بر روی هوش مصنوعی برای

2. National Science Foundation
3. Defense Advanced Research Projects Agency
4. https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505640&org=NSF
5. U.S. Department of Agriculture
6. National Institute of Food and Agriculture (USDA)

علم کشاورزی^۱ و با صنعت آغاز کرده است.^۲ بعلاوه، طرح «بهره‌برداری از ایده بزرگ انقلاب داده»^۳ NSF از تحقیقات بر روی بنیان‌های علم داده پشتیبانی می‌کند که به عنوان پیشران یادگیری ماشین و سیستم‌های هوش مصنوعی در آینده عمل می‌کند.

• DARPA در سپتامبر ۲۰۱۸، یک سرمایه‌گذاری چند ساله به نام کمپین «هوش مصنوعی بعدی» را بر روی برنامه‌های جدید و موجود اعلام کرد.^۴ حوزه‌های کلیدی این کمپین عبارت‌اند از: بهبود استحکام و قابلیت اطمینان سیستم‌های هوش مصنوعی؛ بهبود امنیت و انعطاف فناوری‌های ML/AI، کاهش قدرت، داده، و عملکرد ناکارآمدی‌ها؛ و پیشگامی نسل بعدی الگوریتم‌ها و کاربردهای هوش مصنوعی مانند استدلال عقلانی.

• طرح راهبردی NIH^۵ برای علم داده^۶، در سپتامبر ۲۰۱۸ هدف خود را گسترش دسترسی به فناوری علم داده و قابلیت یادگیری ماشین/هوش مصنوعی برای جامعه تحقیقات زیستی-پزشکی به سمت تحقیقات مراقبت‌های پزشکی داده-محور قرار داده است.

بخش‌های فرعی زیر، برخی از این حوزه‌ها را مورد تأکید قرار می‌دهند. مقوله‌های دیگری از تحقیقات هوش مصنوعی در راهبردهای ۲ تا ۶ مورد بحث قرار می‌گیرند.

1. <https://www.nsf.gov/pubs/2019/nsf19051/nsf19051.jsp>
 2. <https://www.nsf.gov/pubs/2019/nsf19018/nsf19018.jsp>
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651
 3. <https://www.nsf.gov/cise/harnessingdata>
 4. <https://www.darpa.mil/work-with-us/ai-next-campaign>
 5. National Institutes of Health
 6. <https://datascience.nih.gov/strategicplan>

پیشبرد روش‌شناسی‌های داده محور برای تولید دانش

همان‌طور که در طرح راهبردی فدرال تحقیق و توسعه کلان داده^۱ بحث شد، برای دستیابی به درکی از داده‌های هوشمند و اکتشاف دانش، به بسیاری از ابزارها و فناوری‌های جدید بنیادی نیاز داریم. برای توسعه الگوریتم‌های ML پیشرفته که می‌توانند همه اطلاعات مفید پنهان‌شده در کلان داده را شناسایی کنند، به پیشرفت‌های بیشتری نیاز داریم. بسیاری از سؤالات تحقیق باز، به ایجاد و استفاده از داده‌ها، از جمله صحت و تناسب آن‌ها برای راه‌اندازی سیستم هوش مصنوعی مربوط می‌شوند. صحت داده‌ها به‌خصوص هنگامی چالش برانگیز است که با مقادیر معتنا به‌ای از داده‌ها سروکار داریم. این موضوع، ارزیابی و استخراج دانش از داده‌ها را برای انسان‌ها دشوار می‌کند. در حالی که تحقیقات بسیاری به موضوع صحت داده‌ها پرداخته‌اند و برای پاک‌سازی داده‌ها و کشف دانش، از روش‌های تضمین کیفیت داده استفاده کرده‌اند، برای بهبود کارایی فنون پاک‌سازی داده‌ها، به روش‌هایی برای کشف ناهماهنگی و بی‌قاعدگی در داده‌ها و همچنین به توسعه رویکردهایی برای اعمال بازخورد انسانی، به مطالعه بیشتری نیاز داریم. محققان باید روش‌های جدیدی را کشف کنند تا داده‌ها و فراداده‌های مربوطه را هم‌زمان مورد توجه قرار دهند.

بسیاری از کاربردهای هوش مصنوعی، ذاتاً چند رشته‌ای هستند و از داده‌های ناهمگن استفاده می‌کنند. برای تولید دانش از انواع مختلفی از داده (مثلاً، داده گسسته، داده پیوسته، داده متنی، داده مکانی، زمانی، مکانی-زمانی، داده‌های نموداری)، به تحقیقات بیشتری در باره ML

1. <https://www.nitrd.gov/PUBS/bigdatardstrategicplan.pdf>

چندوجهی نیاز داریم. محققان هوش مصنوعی باید میزان داده‌های موردنیاز برای آموزش را تعیین کرده و به درستی به نیازهای مربوط به داده‌های مقیاس بزرگ نسبت به داده‌های دنباله‌دار اشاره کنند. آن‌ها همچنین باید موارد زیر را تعیین کنند: نحوه شناسایی و پردازش رویدادهای نادر، ورای رویکردهای صرفاً آماری؛ کار کردن با منابع دانش (مثلاً، هر نوع اطلاعاتی که جهان را تبیین می‌کند، مانند دانش قانون جاذبه یا هنجارهای اجتماعی) و منابع داده؛ ادغام مدل‌ها و هستی‌شناسی‌ها در فرایند یادگیری و به دست آوردن عملکرد یادگیری کارآمد با داده‌های کم، هنگامی که ممکن است منابع کلان داده در دسترس نباشند.

ارتقا قابلیت‌های ادراکی سیستم‌های هوش مصنوعی

ادراک، پنجره سیستم هوشمند به جهان است. ادراک با داده‌های حسی (احتمالاً توزیعی) آغاز می‌شود که شرایط و اشکال مختلفی دارند، مانند وضعیت خود سیستم یا اطلاعات در مورد محیط. برای استخراج اطلاعات مربوط به کار سیستم هوش مصنوعی، مانند ویژگی‌های هندسی، مختصات، مکان و سرعت داده‌های حسی اغلب به همراه دانش پیشینی و مدل‌ها، پردازش می‌شوند. داده‌های حاصل از ادراک، نوعی از آگاهی را شکل می‌دهند که این آگاهی دانشی جامع و یک مدل از وضعیت جهان را فراهم می‌کند که برای طرح‌ریزی و اجرای کارآمد مورد نیاز است. سیستم‌های هوش مصنوعی از پیشرفت‌های سخت‌افزاری و الگوریتم‌ها برای ایجاد درک

قوی و قابل اعتماد استفاده میکند. سنسورهای حسی باید قادر باشند تا اطلاعات را از فواصل طولانی تر با جزئیات بالاتر و در زمان واقعی ثبت کنند. لازم است سیستم‌های ادراکی بتوانند داده‌های سنسورهای مختلف و منابع دیگر، از جمله ابر محاسباتی را ادغام کنند، تا تعیین کنند که سیستم هوش مصنوعی، در حال حاضر چه چیزی را ادراک می‌کند و امکان پیش‌بینی وضعیت‌های آتی را بدهند. آشکارسازی، طبقه‌بندی، شناسایی و تشخیص اشیاء، به خصوص تحت شرایط آشفته و پویا، همچنان چالش برانگیز می‌مانند. علاوه بر این، باید ادراک انسان‌ها با استفاده از ترکیب مناسبی از حسگرها و الگوریتم‌ها، تا حد زیادی بهبود یابد تا سیستم‌های هوش مصنوعی بتوانند به‌طور مؤثر با افراد کار کنند^۱. برای محاسبه و توزیع عدم قطعیت در طول فرآیند ادراک، به چارچوبی نیاز است تا سیستم هوش مصنوعی به واسطه آن بتواند سطح اعتمادی را که از آگاهی خودش دارد، کمی کرده و دقت را بهبود ببخشد.

شناخت قابلیت‌ها و محدودیت‌های هوش مصنوعی

در حالی که هدف نهایی بسیاری از الگوریتم‌های هوش مصنوعی، حل چالش‌های موجود با راه‌حل‌های انسانی است، درک درستی از توانایی‌ها و محدودیت‌های نظری هوش مصنوعی نداریم و حتی نمی‌دانیم ارائه راه‌حل‌های انسانی تا چه اندازه‌ای با الگوریتم‌های هوش مصنوعی میسر است. برای درک بهتر این که چرا فنون هوش مصنوعی، مخصوصاً ML، - اغلب در عمل به خوبی کار می‌کنند، به کار

۱. مراجعه کنید به گزارش سال ۲۰۱۶ از یک‌صد سال مطالعه بر روی هوش مصنوعی که به مصارف و تأثیرات هوش مصنوعی در سال ۲۰۳۰ می‌پردازد؛ <https://ai100.stanford.edu/report-2016>.

نظری، نیاز داریم. در حالی که رشته‌های مختلف (شامل ریاضیات، علوم کنترل و علوم رایانه) همین مسئله را مطالعه می‌کنند، در حال حاضر این حوزه فاقد مدل‌ها یا چارچوب‌های نظری منسجم برای درک عملکرد سیستم هوش مصنوعی هستند. برای قابلیت حل محاسباتی به تحقیقات مضاعف نیاز است. این قابلیت عبارت است از درک کلاس مسائلی که از لحاظ نظری، الگوریتم‌های هوش مصنوعی قادر به حل آن‌ها هستند و همچنین مسائلی که این الگوریتم‌ها قادر به حل آن‌ها نیستند. این ادراک باید در زمینه سخت‌افزار موجود توسعه داده شود تا ببینیم سخت‌افزار چگونه بر عملکرد این الگوریتم‌ها تأثیر می‌گذارد. درک این مسئله که چه مسائلی از لحاظ نظری قابل حل هستند، می‌تواند در ایجاد راه‌حل‌های تقریبی این مسائل یا حتی ایجاد خطوط جدیدی از تحقیقات بر روی سخت‌افزار جدید برای سیستم‌های هوش مصنوعی، هدایتگر محققان باشد. مثلاً، زمانی که در دهه ۱۹۶۰ شبکه‌های عصبی مصنوعی (ANNs) اختراع شدند، تنها می‌توانستند ی در حل مسائل بسیار ساده مورد استفاده قرار گیرند. صرفاً پس از انجام بهبود سخت‌افزارها مانند متوازن سازی و تنظیم الگوریتم‌ها برای استفاده از سخت‌افزارهای جدید بود که استفاده از شبکه‌های عصبی مصنوعی برای حل مسائل پیچیده امکان‌پذیر شد. چنین تحولاتی، عوامل کلیدی پیشرفت‌های مهم امروزی در یادگیری عمیق هستند.

پیگیری تحقیقات در باره هوش مصنوعی همه منظوره

رویکردهای هوش مصنوعی را می‌توان به «هوش مصنوعی محدود» و «هوش مصنوعی عمومی» تقسیم کرد. سیستم‌های هوش مصنوعی محدود، وظایف مجزایی را در حوزه‌های تخصصی و خوش تعریف، مانند تشخیص گفتار، تشخیص تصویر و ترجمه انجام می‌دهند. چند «سیستم هوشمند مصنوعی محدود» مورد توجه اخیر، از جمله IBM واتسون^۱ و آلفاگو^۲ شرکت دیپ مایند^۳ به دستاوردهای مهمی دست یافته‌اند^۴. در واقع، این سیستم‌های خاص را فرا بشری نامیده‌اند، زیرا آن‌ها به ترتیب، بهتر از بهترین بازیگران انسانی در بازی‌های جنوپاردی و GO بوده‌اند. اما این سیستم‌ها، مصادیق هوش مصنوعی محدود هستند چراکه صرفاً در کارهایی قابل استفاده هستند که به‌طور خاص طراحی شده‌اند. استفاده از این سیستم‌ها مسائل بزرگتر، مستلزم بازطراحی مجدد است. در مقابل، هدف بلندمدت هوش مصنوعی عمومی، ایجاد سیستم‌هایی است که انعطاف‌پذیری و تطبیق‌پذیری هوش انسانی را در طیف وسیعی از حوزه‌های شناختی از جمله یادگیری، زبان، ادراک، استدلال، خلاقیت و طرح‌ریزی نشان بدهند. قابلیت‌های یادگیری گسترده، سیستم‌های هوش مصنوعی عمومی را قادر می‌سازند تا دانش را از یک حوزه به حوزه دیگر منتقل کنند و به صورت دوسویه از تجربه و انسان یاد بگیرند. از زمان ظهور هوش مصنوعی، هوش مصنوعی عمومی از آرزوهای محققان بوده است، اما سیستم‌های کنونی هنوز از دستیابی به این هدف فاصله دارند. در حال حاضر، رابطه بین هوش مصنوعی محدود و

1. Watson
2. AlphaGo
3. DeepMind

۴. در سال ۲۰۱۱، واتسون شرکت IBM دو بازیکنی را که در بین بهترین بازیکنان انسانی در بازی جنوپاردی! بودند، شکست داد.
 ۵. در سال ۲۰۱۶، آلفاگو قهرمان بلامنازع جهان در بازی GO، یعنی لی سی دال را شکست داد. قابل ذکر است که AlphaGo ترکیبی از یادگیری عمیق و جستجوی مونت کارلو است. جستجوی مونت کارلو روشی است که در دهه ۱۹۴۰ بر اساس یک روش احتمالاتی کشف شد و در دهه ۱۹۸۰ توسعه یافت.

هوش مصنوعی عمومی مورد کاوش قرار گرفته و این امکان وجود دارد که درس‌های یکی در بهبود دیگری به کار آیند. در حالی که هیچ توافق عمومی وجود ندارد، بیشتر محققان هوش مصنوعی معتقدند که هنوز چند دهه از هوش مصنوعی عمومی فاصله داریم و دستیابی به آن، نیازمند تلاش‌های تحقیقاتی پایدار است.

توسعه سیستم‌های هوش مصنوعی مقیاس پذیر^۱

ممکن است گروه‌ها و شبکه‌های سیستم‌های هوش مصنوعی با یکدیگر هماهنگ شوند و یا به‌طور خودکار همکاری کنند تا وظایفی را انجام دهند که انجام آن با یک سیستم هوش مصنوعی منفرد امکان‌پذیر نیست و ممکن است شامل انسان‌هایی باشند که در کنار همدیگر کار می‌کنند یا تیمی را هدایت می‌کنند. توسعه و استفاده از چنین سیستم‌های چندگانه هوش مصنوعی، چالش‌های تحقیقاتی مهمی را در باره طرح‌ریزی، هماهنگی، کنترل و مقیاس‌پذیری این سیستم‌ها ایجاد می‌کنند. فنون طرح‌ریزی سیستم‌های چندگانه هوش مصنوعی باید به اندازه کافی سریع عمل کنند و در زمان واقعی با تغییرات محیط سازگار شوند. این سیستم‌ها به شکلی سیال، باید با تغییرات در پهنای باند ارتباطی در دسترس یا خرابی و خطاهای سیستم تطبیق پیدا کنند. بسیاری از تلاش‌های قبلی به طرح‌ریزی متمرکز و فنون هماهنگی پرداخته‌اند، اما این رویکردها در معرض خطاهای نقطه‌ای هستند، یعنی مواردی مانند از دست رفتن برنامه‌ریز، از دست رفتن لینک ارتباطی به برنامه‌ریز. دستیابی

1. scalable

الگوریتمی به فنون طرح‌ریزی و کنترل، سخت‌تر است و اغلب کارایی کمتری داشته و ناقص هستند، اما به‌طور بالقوه استحکام بیشتری نسبت به خطاهای نقطه‌ای دارند. تحقیقات آینده باید فنون کارآمدتر، قوی‌تر و توسعه‌پذیرتری برای طرح‌ریزی، کنترل و همکاری تیم‌های متشکل از سیستم‌های چندگانه هوش مصنوعی و انسان‌ها را کشف کنند.

ترغیب تحقیقات به سوی هوش مصنوعی انسان‌گونه

دستیابی به هوش مصنوعی انسان‌گونه مستلزم وجود سیستم‌هایی است که به روش‌های قابل فهم خودشان را تشریح می‌کنند. این امر، منجر به ایجاد نسل جدیدی از سیستم‌های هوشمند، مانند سیستم‌های تدریس هوشمند و دستیاران هوشمند می‌شود که در کمک به افراد هنگام انجام وظایفشان، کارآمد هستند. با این حال، شکاف قابل توجهی بین نحوه کار الگوریتم‌های هوش مصنوعی کنونی و نحوه یادگیری و انجام وظایف افراد وجود دارد. یادگیری افراد از طرق مختلف انجام می‌شود: یا از طریق مثال‌ها و مصادیق محدود یا از طریق دریافت آموزش رسمی و/یا راهنمایی‌هایی برای انجام وظایف، و یا مشاهده افراد دیگری که همان وظایف را انجام می‌دهند. برای مثال، زمانی که در دانشکده‌های پزشکی، دانشجویان پزشکی با مشاهده یک پزشک متخصص که در حال انجام یک روش پزشکی پیچیده است، فراگیری دارند، در واقع این رویکرد اتخاذ شده است. حتی در کارهای سطح بالا، مانند بازی‌های قهرمانی جهان GO، یک

بازیکن حرفه‌ای برای آموزش خود، چند هزار بازی انجام خواهد داد. در مقابل، صدها سال طول می‌کشد تا یک انسان تعداد بازی‌های موردنیاز برای آموزش آلفاگو را انجام دهد. تحقیقات بنیادی بیشتر در زمینه رویکردهای جدید برای رسیدن به هوش مصنوعی انسان گونه، سیستم‌ها را به این هدف نزدیک‌تر خواهد کرد.

توسعه روبات‌هایی با توانایی و اعتبار بیشتر

در طول دهه اخیر پیشرفت‌های چشمگیر در فناوری‌های روباتیک منجر به تنوع کاربردها از جمله تولید، تدارکات، پزشکی، بهداشت، دفاع و امنیت ملی، کشاورزی و محصولات مصرفی شده است. در حالی که روبات‌ها از نظر تاریخی برای محیط‌های صنعتی ایستا پیش‌بینی شده بودند، پیشرفت‌های اخیر شامل همکاری‌های نزدیک میان روبات‌ها و انسان‌ها هستند. در حال حاضر فناوری‌های روباتیک در توانایی خود برای تکمیل، افزایش، بهبود یا تقلید از قابلیت‌های فیزیکی انسانی یا هوش انسانی نویدهایی را می‌دهند. با این حال، لازم است تا دانشمندان این سیستم‌های روباتیکی را توانمندتر، قابل‌اعتمادتر کرده و استفاده از آن‌ها را آسان‌تر کنند.

محققان به فهم بهتری از ادراک روباتیک نیاز دارند تا بتوانند اطلاعات را از سنسورهای مختلف گرفته و از این طریق ربات‌هایی بسازند که آگاهی از موقعیت در آن‌ها بی‌درنگ صورت می‌گیرد. امکان درک بهتر و تعامل با دنیای فیزیکی روبات‌ها، به پیشرفت در شناخت و استدلال نیاز دارد. توانایی پیشرفته برای سازگاری و

یادگیری به روبات‌ها اجازه می‌دهند تا مهارت‌های خود را تعمیم دهند، از عملکرد فعلی خودشان ارزیابی داشته باشند و مجموعه‌ای از حرکات فیزیکی را از معلم انسانی یاد بگیرند. تحرک و تنظیم گری، حوزه‌هایی برای تحقیقات بیشتر هستند، یعنی این مساله که روبات‌ها بتوانند در مناطق ناهموار و نامطمئن حرکت کنند و با مهارت، اشیا مختلف را هدایت کنند. روبات‌ها نیازمند یادگیری هستند تا بطور منسجم تیم تشکیل دهند و به روشی قابل اعتماد و قابل پیش‌بینی با انسان‌ها همکاری کنند.

پیشرفت سخت‌افزار برای تقویت هوش مصنوعی

در حالی که تحقیقات هوش مصنوعی اغلب با پیشرفت نرم‌افزار در ارتباط است، عملکرد سیستم‌های هوش مصنوعی به شدت به سخت‌افزاری بستگی دارد که بر روی آن اجرا می‌شود. رنسانس کنونی در ML ژرف، به‌طور مستقیم با پیشرفت فناوری سخت‌افزاری مبتنی بر GPU و حافظه بهبود یافته آن^۱، ورودی/خروجی، سرعت پالس ساعت، پردازش موازی و بازده انرژی، گره خورده است. توسعه سخت‌افزاری که برای الگوریتم‌های هوش مصنوعی بهینه‌سازی شده، حتی کارایی بالاتری نسبت به GPU ها فراهم می‌کند. ، ، یک مثال، پردازشگرهای نورومورفیک^۲ است که تا حدودی از ساختار مغز انسان الهام گرفته شده و در برخی موارد برای عملکرد شبکه‌های عصبی بهینه‌سازی شده است^۳.

پیشرفت‌های سخت‌افزاری همچنین می‌توانند عملکرد روش‌های

۱. GPU مخفف واحد پردازش گرافیکی است که پردازنده‌ای است با مصرف کم برق و مقرون به‌صرفه که صدها هسته پردازش را شامل می‌شود؛ این طراحی، آن را به خصوص برای کاربردهای ذاتا موازی، از جمله بیشتر سیستم‌های هوش مصنوعی، مناسب می‌سازد.

2. neuromorphic

۳. رایانش neuromorphic به توانایی سخت‌افزار برای یادگیری، سازگاری و بیکربندی فیزیکی اشاره دارد و از زیست‌شناسی و علوم اعصاب الهام گرفته شده است.

هوش مصنوعی را بهبود بخشند که به شدت داده محور هستند. برای روشن و خاموش کردن مسیره‌های داده به شیوه‌ای کنترل شده در طول یک سیستم توزیع شده، نیازمند مطالعه بیشتری در باره روش‌ها هستیم. در ضمن، یادگیری مؤثر الگوریتم‌های ML از داده‌های پر سرعت، مثل الگوریتم‌های ML توزیع شده که هم‌زمان از چندین مسیر داده میگیرند به تحقیقات نیاز دارد. روش‌های بازخورد مبتنی بر ML پیشرفته‌تر، به سیستم‌های هوش مصنوعی امکان خواهند داد تا از داده‌های شبیه‌سازی‌هایی با مقیاس بزرگ، ابزارهای آزمایشگاهی و سیستم‌های حسگر توزیع شده، مانند ساختمان‌های هوشمند و اینترنت اشیا (IoT)، به صورت هوشمندانه‌ای نمونه بگیرند یا این داده‌ها را در اولویت قرار دهند. ممکن است چنین روش‌هایی مستلزم تصمیم‌گیری پویای ورودی/خروجی باشند که در آن، به جای اینکه ذخیره داده با تناوب ثابت انجام شود، این کار در زمان واقعی و بر اساس اهمیت داده‌ها صورت می‌گیرد.

ایجاد هوش مصنوعی برای سخت‌افزارهای پیشرفته

در حالی که بهبود سخت‌افزار می‌تواند به سیستم‌های توانمندتر هوش مصنوعی منجر شود، سیستم‌های هوش مصنوعی نیز می‌توانند عملکرد سخت‌افزار را بهبود بخشند^۱. این عمل متقابل به پیشرفت‌های بیشتری در عملکرد سخت‌افزاری منجر خواهد شد، زیرا محدودیت‌های فیزیکی در محاسبات، به رویکردهای بدیع در طراحی سخت‌افزار نیاز دارند^۲. روش‌های هوش مصنوعی به ویژه می‌توانند

۱. ام. میلانو و آل. بنینی، «مدل‌سازی پیش‌گویانه برای مصرف برق انجام کار در سیستم‌های HPC»، در مجموعه مقالات کنفرانس رایانش با عملکرد بالا: سی و یکمین کنفرانس بین‌المللی، ISC با عملکرد بالا ۲۰۱۶ (اسپرینگر مجلد ۹۶۹۷، ۲۰۱۶).
 ۲. این محدودیت‌های فیزیکی در محاسبات، مقیاس بندی Dennard نامیده می‌شوند و به تراکم انرژی بالا در تراشه منجر می‌شوند. بعلاوه این محدودیت‌ها به پدیده‌ای به نام «سیلیکون تاریک» نیز منجر می‌شوند که در آن، بخش‌های مختلف تراشه باید خاموش شوند تا دما محدود شود و اطمینان از یکپارچگی داده‌ها حاصل شود.

برای بهبود عملکرد سیستم‌های رایانش با کارایی بالا (HPC) مهم باشند. چنین سیستم‌هایی مقادیر زیادی انرژی مصرف می‌کنند. از هوش مصنوعی برای پیش‌بینی عملکرد HPC و مصرف منابع و اتخاذ تصمیم‌های بهینه‌سازی آنلاین استفاده می‌شود که کارایی را افزایش می‌دهند؛ فنون هوش مصنوعی پیشرفته‌تر می‌توانند عملکرد سیستم را افزایش دهند. از هوش مصنوعی می‌توان برای ایجاد سیستم‌های HPC که قادر به پیکربندی خودشان هستند، نیز استفاده کرد. این سیستم‌ها می‌توانند هنگام وقوع خرابی، بدون دخالت انسان، نسبت به رفع آن‌ها اقدام کنند.^۲

الگوریتم‌های هوش مصنوعی تقویت شده، می‌تواند با کاهش جابه‌جایی داده‌ها بین پردازنده‌ها و حافظه (مانع اصلی سیستم‌های محاسباتی با مقیاس‌اگزا که ۱۰ برابر سریع‌تر از ابرکامپیوترهای امروزی عمل می‌کنند)، عملکرد سیستم‌های چند هسته‌ای را افزایش دهد.^۳ در عمل، پیکربندی اجراها در سیستم‌های HPC، هرگز یکسان نیستند و وضعیت هر کد نرم‌افزاری متفاوت، به موقع و به صورت مستقل تغییر می‌کند و لذا کاربردهای مختلف، هم‌زمان اجرا می‌شوند. لازم است به نوعی از الگوریتم‌های هوش مصنوعی طراحی شوند که به صورت آنلاین و در مقیاس مناسب با سیستم‌های HPC، کار می‌کنند.

۱. high-performance computing

۲. ای. کوکانا-فراناندز، جی. رانیلا و آل. سانچز، «تخصیص مؤثر انرژی اسلاتهای گره محاسباتی در خوشه‌های HPC از طریق یادگیری پارامتر و مدل‌سازی سیستم فازی ژنتیکی هیبریدی»، ژورنال سوپر کامپیوتینگ، شماره ۷۱ (۲۰۱۵): ۱۱۶۳-۱۱۷۴.
۳. سیستم‌های رایانش با مقیاس‌اگزا می‌توانند حداقل یک میلیارد محاسبات را در ثانیه انجام دهند.

بخش دوم

راهبرد دوم: توسعه روش های مؤثر
برای همکاری انسان و هوش مصنوعی





راهبرد دوم: توسعه روش‌های مؤثر برای همکاری انسان و هوش مصنوعی

توسعه سیستم‌های هوش مصنوعی که قابلیت‌های انسانی را تکمیل کرده و به آن‌ها می‌افزایند، با افزایش تمرکز بر آینده کار (به روز سازی ۲۰۱۹)

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی (۲۰۱۶)، شاهد افزایش علاقه ملی در همکاری انسان و هوش مصنوعی بوده ایم. زمانی که سیستم‌های هوش مصنوعی، قابلیت‌های انسانی را تکمیل کرده و به آن‌ها می‌افزایند، انسان‌ها و هوش مصنوعی در طیفی از سناریوهای مشترک تا سناریوهای کاملاً مستقل، به شریک هم تبدیل می‌شوند. به‌طور خاص، همکاری انسان و هوش مصنوعی، هم به عنوان یک چالش و هم فرصتی در زمینه آینده کار ارتقا یافته است.

در سه سال گذشته، کنفرانس‌ها، کارگاه‌ها و گروه‌های کاری تازه تأسیس و قدیمی، به‌طور گسترده همکاری انسان و هوش مصنوعی را در اولویت قرار داده‌اند. مثلاً، کنفرانس مربوط به محاسبات انسانی و جمع‌سپاری از یک کارگاه به یک کنفرانس مهم بین‌المللی تبدیل شده است که تحقیقات در فصل مشترک هوش مصنوعی و تعامل

انسان و رایانه را ترویج می‌کند (HCI)^۱. در سال ۲۰۱۸، انجمن پیشرفت هوش مصنوعی، همکاری انسان و هوش مصنوعی را به عنوان موضوع نوظهور کنفرانس سالانه خود انتخاب کرد^۲. در ماه مه ۲۰۱۹، بزرگ‌ترین کنفرانس تعامل انسان و رایانه (CHI)، شامل یک کارگاه «پر کردن شکاف بین هوش مصنوعی و HCI بود^۳. ژورنال تعامل انسان و رایانه در ماه مارس ۲۰۱۰ فراخوانی را برای ارائه شماره ویژه‌ای در باره «یکپارچه کردن تعامل انسان و رایانه و هوش مصنوعی» مطرح کرد^۴. در زمینه کار، کنفرانس‌هایی ظهور یافته‌اند که به بررسی نقش انسان، ماشین و مشارکت آن‌ها می‌پردازند، مانند «علوم رایانه و آزمایشگاه هوش مصنوعی (CSAIL) و طرح اقتصاد دیجیتال که هوش مصنوعی سالانه و کنگره آینده کار را راه‌اندازی کرد^۵».

در سال ۲۰۱۹، کنسرسیوم جامعه محاسبات (CCC) به عنوان بخشی از «نقشه راه ۲۰ ساله جامعه برای تحقیقات هوش مصنوعی در آمریکا»^۶، کارگاهی را برگزار کرد که به تعامل معنادار بین انسان و سیستم‌های هوش مصنوعی می‌پرداخت^۷. علاوه بر این، CCC در سال ۲۰۱۷-۲۰۱۸ به منظور تمرکز بر توان بالقوه فناوری برای افزودن عملکرد انسانی در مواردی (ولی نه صرفاً در این موارد) مثل محیط کار، کلاس و سیستم مراقبت بهداشتی، گروه مطالعاتی مرز فناوری انسانی را عملیاتی کرد^۸.

۱. HCOMP ۲۰۱۹ خوش‌آمدید: <https://www.humancomputation.com>

۲. AAAI18- Emerging Topic Human-AI Collaboration: <http://www.aaai.org/Conferences/AAAI/2018/aaai18emergingcall.php>

۳. بشر کجاست؟ پر کردن شکاف بین AI و HCI: کنفرانس CHI ۲۰۱۹: <https://michaelv.ai/hci-workshop>

۴. فراخوان: شماره ویژه تعامل انسان-رایانه «یکپارچه‌سازی تعامل انسان-رایانه و هوش مصنوعی»
<https://isprinfo/20/02/2019/call-unifying-human-computer-interaction-and-artificial-intelligence-issue-of-human-computer-interaction/>

۵. <https://futureofworkcsail.mit.edu/>

۶. گروه‌هایی ۲۰۱۹ هوش مصنوعی و آینده نوآوری کار: <https://analyticseven.com>

۷. https://cra.org/ccc/wp-content/uploads/sites/03/2019/2/AI_Roadmap_Exec_Summary-FINAL-.pdf

۸. Artificial Intelligence Roadmap Workshop 2 - Interaction: <https://cra.org/cc/events/artificial-intelligence->

۹. <https://cra.org/ccc/human-technology-frontier/>

اصل راهبرد متقابل در طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، یعنی «اعتماد مناسب سیستم‌های هوش مصنوعی مستلزم قابلیت تبیین است، به ویژه هم‌زمان با اینکه هوش مصنوعی از لحاظ مقیاس و پیچیدگی رشد می‌کند»، شاهد دعوت به اقدام تحقیق و توسعه در زمینه همکاری‌های انسان-هوش مصنوعی بوده است. تعدادی از جوامع و موسسه‌های حرفه‌ای، این اصل را به عنوان یک حوزه اولویت شناسایی کرده‌اند (مربع زیر را ملاحظه کنید). این حوزه تحقیقاتی، تلاقی راهبردهای ۲ و ۳ را منعکس می‌کند، زیرا قابلیت تبیین، انصاف و شفافیت اصول کلیدی سیستم‌های هوش مصنوعی برای همکاری مؤثر با انسان‌ها است. به همین ترتیب، چالش درک و طراحی اخلاق انسان و هوش مصنوعی و تنظیم ارزش در سیستم‌ها، حوزه تحقیقاتی باز باقی خواهد ماند. به موازات آن، بخش خصوصی با اصولی برای همکاری مؤثر انسان-هوش مصنوعی پاسخ داده است.^{۱، ۲}

همکاری انسان و هوش مصنوعی: برنامه‌های اخیر تحقیق و توسعه مؤسسات

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی (۲۰۱۶)، چندین موسسه تلاش‌هایی را برای تحقق راهبرد ۲ شروع کرده‌اند:

- طرح «ایده بزرگ» موسسه NSF با نام «آینده کار بر روی مرز انسان-فناوری» از تحقیقات اجتماعی-فنی حمایت می‌کند. این تحقیقات، آینده‌ای را محقق می‌کنند که در آن، فناوری‌های

1. <https://www.microsoft.com/en-us/research/uploads/prod/01/2019/Guidelines-for-Human-AI-Interaction-camera-ready.pdf>

2. <https://www.partnershiponai.org/about/#our-work>

هوشمند به صورت هم افزایانه با انسان‌ها همکاری می‌کنند تا مشارکت گسترده در نیروی کار صورت بگیرد و منافع اجتماعی، اقتصادی و محیطی در گستره‌ای از محیط‌های کاری بهبود یابند.

- NOAA (اداره ملی اقیانوس‌شناسی و هواشناسی) در حال پیشبرد همکاری انسان-هوش مصنوعی در زمینه طوفان، گردباد و دیگر پیش‌بینی‌های آب و هوایی شدید است که در آن‌ها، پیش‌بینی‌کننده انسانی و یک سیستم هوش مصنوعی همکاری می‌کنند تا در روند تولید هشدار آب‌وهوای بد و شناسایی الگوهای متمایزی که گویای رویدادهای شدید هستند، بهبود حاصل شود. در این سیستم که گاهی اوقات از آن به عنوان «انسان در بالای حلقه» یاد می‌شود، پیش‌بینی‌کنندگان انسانی بر پیش‌بینی‌های سیستم هوش مصنوعی نظارت می‌کنند و نتایج را هدایت می‌کنند.
- NIH تحقیقات مستمری در زمینه پردازش زبان طبیعی دارد که بر مبنای یک پایگاه داده مشتمل بر ۹۶,۳ میلیون داده که از استنادهای MEDLINE حفظ شده در کتابخانه ملی پزشکی استخراج شده اند، کار می‌کند.

- گزارش کارگاه^۱ DOE ۲۰۱۹ درباره ML علمی، موارد زیر را شناسایی کرد: اولویت جهت‌گیری‌های تحقیقاتی، موارد اصلی کاربرد علمی و روندهای نوظهوری که از طریق آن‌ها، همکاری‌های انسان و هوش موجب تحول در شیوه علم خواهد شد .

در سه سال گذشته موسسه‌های فدرال، سرمایه‌گذاری‌ها در هوش مصنوعی و اهداف این مأموریت را افزایش داده‌اند و تأکید مشترکی بر شناخت انسان-ماشین، استقلال و عاملیت دارند، مانند پشتیبانی

1. Department of Energy

تصمیم‌گیری، مدل‌سازی ریسک، آگاهی موقعیتی و هوش ماشینی قابل‌اعتماد (مربع بالا را ملاحظه کنید). از طریق چنین سرمایه‌گذاری‌های تحقیق و توسعه‌ای، شراکت‌های پژوهشی در چندین محور رشد می‌کنند و دانشمندی از رشته‌های مختلف را گرد هم می‌آورند، مثل دانشمندان علوم محاسباتی، دانشمندان علوم رفتاری، شناختی و روان‌شناختی و دانشمندان و مهندسان دیگر حوزه‌ها. همکاری‌های جدید بین محققان دانشگاهی و کاربران سیستم‌های هوش مصنوعی داخل و خارج از محل کار شکل گرفته‌اند.

برای حرکت رو به جلو، این نکته حیاتی است که موسسه‌های فدرال به ترویج تحقیق و توسعه هوش مصنوعی در دنیای باز ادامه دهند تا در طراحی سیستم‌های هوش مصنوعی ارتقا صورت بگیرد تا جایی که این سیستم‌ها قادر باشند که وضعیت و اهداف کاربران را به گونه‌ای ترکیب کرده و تطبیق دهند که بتوانند در شرایط پیش‌بینی شده و پیش‌بینی نشده با کاربران کار مشترک انجام دهند.

در حالی که سیستم‌های هوش مصنوعی کاملاً خودکار، در برخی از حوزه‌های کاربردی (مثلاً، اکتشاف زیر آب و یا عمق فضا) مهم خواهند بود، ترکیبی از انسان و سیستم‌های هوش مصنوعی که با همدیگر کار می‌کنند تا به اهداف کاربردی برسند، به بسیاری از حوزه‌های کاربردی دیگر (مثل بازگشت از بیماری‌ها و تشخیص پزشکی) می‌پردازند. این تعامل مشترک، ریشه در مکمل بودن انسان‌ها و سیستم‌های هوش مصنوعی دارد. در حالی که رویکردهای مؤثری برای همکاری انسان-هوش مصنوعی در حال حاضر وجود

دارند، بیشتر آن‌ها «راه‌حل‌های نقطه‌ای^۱» هستند که تنها با استفاده از پلتفرم‌های خاص در محیط‌های خاص به سمت اهداف خاص کار می‌کنند. ایجاد راه‌حل‌های نقطه‌ای برای نمونه برنامه‌های موجود کافی نیست. بنابراین برای فراتر رفتن از این راه‌حل‌های نقطه‌ای به سوی روش‌های کلی‌تر همکاری انسان-هوش مصنوعی به کار بیشتری نیاز داریم. باید به دنبال مصالحه و تعادل در گزینه‌های احتمالی باشیم، مصالحه بین طراحی سیستم‌های عمومی (که در همه نوع مساله‌ای کارایی دارند، برای ایجاد و تسهیلات بیشتر برای عوض کردن کاربردها، نیاز به تلاش انسانی کمتری دارند) و ساخت تعداد زیادی از سیستم‌های مخصوص مشکلات که ممکن است برای هر مشکل مؤثرتر عمل کنند.

برنامه‌های کاربردی آینده از جنبه‌های زیر با هم متفاوت خواهند بود: تقسیم نقش کارکردی بین انسان‌ها و سیستم‌های هوش مصنوعی؛ ماهیت تعاملات بین انسان‌ها و سیستم‌های هوش مصنوعی؛ تعداد انسان‌ها و دیگر سیستم‌های هوش مصنوعی که با یکدیگر کار می‌کنند و نحوه ارتباط و تسهیم آگاهی موقعیتی بین انسان‌ها و سیستم‌های هوش مصنوعی. تقسیمات نقش کارکردی بین انسان‌ها و سیستم‌های هوش مصنوعی به‌طور معمول در یکی از طبقات زیر قرار می‌گیرند:

۱- هوش مصنوعی در کنار انسان اعمال را انجام می‌دهد: سیستم‌های هوش مصنوعی وظایف جانبی را که از تصمیم‌گیری انسان پشتیبانی می‌کنند، انجام می‌دهند. برای مثال، هوش مصنوعی می‌تواند با حافظه کاری، بازیابی حافظه کوتاه‌مدت یا

بلندمدت و وظایف پیشگویی به انسان‌ها کمک کند.

۲- هوش مصنوعی هنگامی عمل می‌کند که انسان با بار شناختی بالا مواجه شود: سیستم‌های هوش مصنوعی وظایف نظارتی پیچیده‌ای را انجام می‌دهند (مانند سیستم‌های هشدار نزدیک شدن به زمین در هواپیما)، تصمیم‌گیری و تشخیص‌های خودکار پزشکی، هنگامی که انسان به کمک نیاز دارد.

۳- هوش مصنوعی به جای یک انسان عمل می‌کند: سیستم‌های هوش مصنوعی وظایفی را انجام می‌دهند که انسان‌ها قابلیت‌های بسیار محدودی برای انجام آن‌ها دارند، مانند عملیات پیچیده ریاضی، هدایت کنترل سیستم‌های پویا در محیط‌های عملیاتی رقابتی، جنبه‌هایی از کنترل سیستم‌های خودکار در محیط‌های مضر یا سمی و در موقعیت‌هایی که یک سیستم باید بسیار سریع واکنش نشان دهد (مثل اتاق‌های کنترل راکتور هسته‌ای).

دستیابی به تعاملات مؤثر بین انسان‌ها و سیستم‌های هوش مصنوعی مستلزم تحقیق و توسعه بیشتری است تا اطمینان حاصل شود که طراحی سیستم، منجر به پیچیدگی بیش از حد، عدم اعتماد یا اعتماد بیش از حد نمی‌شود. آشنایی انسان با سیستم‌های هوش مصنوعی می‌تواند از طریق آموزش و تجربه افزایش یابد تا انسان از قابلیت‌های سیستم هوش مصنوعی و آنچه که سیستم هوش مصنوعی نمی‌تواند انجام دهد، درک خوبی پیدا کند. برای رفع این نگرانی‌ها باید از برخی اصول اتوماسیون انسان‌محور در طراحی و توسعه این سیستم‌ها استفاده کرد^۱:

۱. سی. ویکنز و جی. جی. هولندز «توجه، اشتراک زمان و حجم کار» در نشریه مهندسی، روانشناسی و عملگردهای انسانی (لندن): پیرسون PLC، (۱۹۹۹)، ۴۳۹ - ۴۷۹.

- ۱- در سیستم انسان-هوش مصنوعی از طرح ساده، کنترل‌ها و نمایشگرهایی استفاده کنید که شهودی هستند.
- ۲- اپراتور را مطلع نگه دارید. اطلاعات مهم، حالات سیستم هوش مصنوعی و تغییرات این حالت‌ها را به اطلاع وی برسانید.
- ۳- دائماً اپراتور را آموزش دهید. مرتباً دانش عمومی، مهارت‌ها و توانایی‌ها (KSA) و آموزش الگوریتم‌ها و منطق به کار گرفته شده در سیستم‌های هوش مصنوعی و حالت‌های محتمل خرابی در سیستم را آموزش دهید.
- ۴- اتوماسیون را انعطاف‌پذیر کنید. استقرار سیستم‌های هوش مصنوعی باید به عنوان یک گزینه طراحی برای اپراتورهایی در نظر گرفته شود که می‌خواهند تصمیم بگیرند که آیا می‌خواهند از آن‌ها استفاده کنند یا نه. طراحی و استقرار سیستم‌های هوش مصنوعی انطباق‌پذیر که می‌توان از آن‌ها برای پشتیبانی از اپراتورهای انسانی در خلال دوره‌های کاری سنگین یا فرسوده کننده استفاده کرد نیز حائز اهمیت است^۱.

بسیاری از چالش‌های بنیادی محققان، هنگام ایجاد سیستم‌هایی پیش می‌آیند که با انسان‌ها به‌طور مؤثری کار می‌کنند. چندین چالش مهم در زیربخش‌های بعدی آورده شده‌اند.

1. https://www.nasa.gov/mission_pages/SOFIA/index.html
<https://cloud1.arc.nasa.gov/intex-na/>

جستجوی الگوریتم‌های جدید برای سیستم‌های هوش مصنوعی آگاه- انسانی^۱

در طول سال‌ها، الگوریتم‌های هوش مصنوعی قادر به حل مسائل پیچیده شده‌اند. با این حال، شکافی بین قابلیت‌های این الگوریتم‌ها و قابلیت استفاده از این سیستم‌ها توسط انسان وجود دارد. به سیستم‌های هوشمند آگاه-انسانی نیاز داریم که می‌توانند به‌طور شهودی با کاربران تعامل داشته و امکان همکاری انسان-ماشین را به وجود بیاورند. تعاملات شهودی شامل تعاملات سطحی، مانند زمانی که یک کاربر به گزینه توصیه شده توسط سیستم توجهی نمی‌کند؛ رویکردهای مبتنی بر مدل که اقدامات گذشته کاربران را در نظر می‌گیرند؛ یا حتی مدل‌های عمیق کاربرمحور که مبتنی بر مدل‌های دقیق شناختی انسان هستند. باید مدل‌های مداخله‌ای^۲ توسعه یابند که به یک سیستم هوشمند اجازه می‌دهد تا تنها در صورت لزوم، در کار انسان وقفه ایجاد کند. سیستم‌های هوشمند باید توانایی تقویت شناخت انسان را نیز داشته باشند، بدانند چه اطلاعاتی را بازیابی کنند که کاربر به آن نیاز دارد، حتی زمانی که کاربران صراحتاً به سیستم اعلام نمیکنند که به آن اطلاعات نیاز دارند. سیستم‌های هوشمند آینده باید بتوانند هنجارهای اجتماعی انسان را در نظر بگیرند و بر اساس آن عمل کنند. سیستم‌های هوشمند در صورتی می‌توانند به‌طور مؤثرتری با انسان‌ها کار کنند که دارای درجه‌ای از هوش عاطفی باشند تا به درستی احساسات کاربران خود را بشناسند و به آن‌ها پاسخ دهند. هدف تحقیقاتی دیگر این است که از تعاملات یک انسان و یک ماشین فراتر رفته و به سمت

«سیستم‌های سیستم‌ها» حرکت کنیم، یعنی تیم‌هایی متشکل از چندین ماشین که با چندین انسان تعامل دارند. تعاملات انسان و سیستم هوش مصنوعی طیف وسیعی از اهداف را شامل میشوند. سیستم‌های هوش مصنوعی باید به این توانمندی برسند که انبوهی از اهداف را در نظر گرفته و اقدامات متناسب با آن اهداف را دنبال کنند، همچنین محدودیت‌های مربوط به آن اقدامات و سایر عوامل را در نظر بگیرند و در نهایت به راحتی با تغییرات در اهداف سازگار شوند. بعلاوه، انسان‌ها و سیستم‌های هوش مصنوعی باید اهداف مشترکی را به اشتراک گذارند و از آن‌ها و جنبه‌های مرتبط با آن‌ها درک متقابلی داشته باشند. برای توسعه سیستم‌هایی که نیاز به مهندسی انسانی کمتری دارند، تحقیقات بیشتری لازم است.

توسعه فنون هوش مصنوعی برای افزایش توانایی انسان

در حالی که بیشتر تمرکز تحقیقات پیشین هوش مصنوعی بر روی الگوریتم‌هایی بوده که با افرادی که وظایف محدودی انجام می‌دهند، سازگار بوده یا بهتر از آن‌ها عمل کنند، برای توسعه سیستم‌هایی که قابلیت‌های انسانی را در بسیاری از حوزه‌ها تقویت می‌کنند، کار بیشتری لازم است. تحقیقات مربوط به افزایش توانایی‌های انسان شامل موارد زیر هستند: الگوریتم‌هایی که بر روی یک دستگاه ایستا (مانند رایانه) کار می‌کنند؛ دستگاه‌های پوشیدنی (مانند عینک هوشمند)؛ وسایل تعبیه شده (مانند رابط‌های مغز) و

در محیط‌های کاربری خاص (مانند اتاق‌های جراحی ویژه). برای مثال، ارتقا آگاهی انسان می‌تواند یک دستیار پزشک را قادر سازد تا بر اساس خواندن داده حاصل از چندین دستگاه، یک اشتباه در روال پزشکی را تشخیص دهد. سایر سیستم‌ها می‌توانند با کمک به یادآوری تجربیات گذشته که قابل اعمال به وضعیت فعلی کاربر است، شناخت انسان را افزایش دهند.

نوع دیگری از همکاری بین انسان و سیستم‌های هوش مصنوعی شامل یادگیری فعال برای درک داده‌های هوشمند است. در یادگیری فعال، از یک حوزه تخصصی درخواست ورودی می‌شود و یادگیری فقط در زمانی انجام می‌شود که الگوریتم یادگیری غیرقطعی باشد. این امر، برای کاهش حجم داده‌های آموزشی که ابتدا باید تولید شوند یا میزان یادگیری لازم، تکنیک مهمی تلقی می‌شود. یادگیری فعال، راه کلیدی در به دست آوردن ورودی از حوزه تخصصی و افزایش اعتماد در الگوریتم یادگیری است. تاکنون یادگیری فعال، تنها در یادگیری نظارت شده مورد استفاده قرار گرفته است؛ تحقیقات بیشتر برای گنجاندن یادگیری فعال در یادگیری بدون نظارت (مثلاً خوشه‌بندی، شناسایی آنومالی) و یادگیری تقویتی لازم است.^۱ شبکه‌های احتمالی به حوزه دانشی اجازه می‌دهند به شکل توزیع احتمال قبلی افزایش پیدا کنند. باید روش‌های کلی را جستجو کرد که به الگوریتم‌های ML امکان گنجاندن حوزه دانشی را چه به شکل مدل‌های ریاضی، متن و... می‌دهند.

۱. در حالی که یادگیری نظارت شده مستلزم این است که انسان‌ها پاسخ‌های مبتنی بر حقیقت را ارائه دهند، یادگیری تقویتی و یادگیری بدون نظارت این گونه نیستند.

توسعه فنونی برای مصورسازی و واسطه‌های^۱ انسان-هوش مصنوعی

مصورسازی و رابط‌های کاربری بهتر، حوزه‌های دیگری هستند که نیاز به توسعه بیشتری دارند تا به انسان‌ها کمک کنند حجم انبوه مجموعه داده‌ها و اطلاعات جدید حاصل از منابع مختلف را درک کنند. مصورسازی و رابط‌های کاربری باید به وضوح داده‌های پیچیده روزافزون و اطلاعات به دست آمده از آن‌ها را به صورت قابل‌درکی ارائه کنند. ارائه نتایج بلادرنگ در عملیات بحرانی، حائز اهمیت است و ممکن است با افزایش توان محاسباتی و سیستم‌های متصل به آن قابل دستیابی شود. در چنین وضعیت‌هایی، کاربران به مصورسازی و رابط‌های کاربری نیاز دارند که می‌توانند به سرعت اطلاعات صحیح را برای پاسخ بی‌درنگ منتقل کنند.

همکاری انسان و هوش مصنوعی را می‌توان در طیف وسیعی از محیط‌ها و جاهایی که محدودیت‌هایی در ارتباط وجود دارد، به کار گرفت. در برخی حوزه‌ها، ارتباطات انسان-هوش مصنوعی پایین هستند و ارتباطات سریع و قابل‌اطمینان هستند. در حوزه‌های دیگر (مثل استقرار کاوشگرهای اسپیریت و آپورچیونیتی^۲ در مریخ)، ارتباطات از راه دور بین انسان‌ها و سیستم هوش مصنوعی، تأخیر بسیار بالایی دارند (مثلاً، زمان رفت‌و برگشت ۵-۲۰ دقیقه‌ای بین زمین و مریخ)، بنابراین نیاز به پلتفرم (های) استقرار یافته است که تا حد زیادی مستقل بوده و تنها با اهداف راهبردی سطح بالای منتقل شده به پلتفرم عمل می‌کنند. این الزامات و محدودیت‌های ارتباطی، ملاحظات مهمی برای تحقیق و توسعه واسطه‌های کاربری هستند.

1. interfaces
2. Spirit and Opportunity

توسعه سیستم‌های پردازش زبانی مؤثر تر

مدت‌های مدیدی است که توانمندسازی افراد برای تعامل با سیستم‌های هوش مصنوعی از طریق زبان گفتاری و نوشتاری، هدف محققان هوش مصنوعی بوده است. در حالی که پیشرفت‌های قابل توجهی صورت گرفته، اما قبل از اینکه انسان‌ها بتوانند مثل برقراری ارتباط با سایر انسان‌ها، به‌طور مؤثر با سیستم‌های هوش مصنوعی ارتباط برقرار کنند، باید چالش‌های تحقیقاتی باز قابل توجهی در پردازش زبانی مرتفع شوند. اکثر پیشرفت‌های اخیر در پردازش زبانی به استفاده از رویکردهای ML داده محور اعتبار بخشیده‌اند که منجر به سیستم‌های موفق‌تری شده‌اند که مثلاً، به‌طور موفقیت آمیزی گفتار انگلیسی روان را در محیط‌های آرام، تشخیص می‌دهند. با این وصف، این دستاوردها تنها گام‌های اولیه برای رسیدن به اهداف بلندمدت هستند. سیستم‌های کنونی نمی‌توانند چالش‌های دنیای واقعی مانند گفتار در محیط‌های شلوغ، گفتار با لهجه غلیظ، گفتار کودکان، اختلال تکلم و گفتار برای زبان اشاره را حل و فصل کنند. توسعه سیستم‌های پردازش زبانی که قادر به مشارکت در گفتگوی آنی با انسان‌ها باشند، نیز مورد نیاز هستند. چنین سیستم‌هایی باید اهداف و مقاصد مخاطبان خود را استنتاج کنند، از شدت صدا، سبک و بیان مناسبی برای وضعیت استفاده کنند و در صورت سو تفاهم گفتاری، راهبردهای اصلاحی را به کار گیرند. نیاز به تحقیقات بیشتر در مورد توسعه سیستم‌هایی وجود دارد که به سادگی به زبان‌های مختلف تعمیم می‌یابند. علاوه بر این، در زمینه کسب دانش مفید و ساختار یافته حوزه، به شیوه‌ای

آسان و توسط سیستم‌های پردازش زبانی، به مطالعات بیشتر نیاز است.

در بسیاری از حوزه‌های دیگر نیز برای ایجاد تعاملات بین انسان‌ها و سیستم‌های هوش مصنوعی، به پیشرفت‌های پردازش زبانی نیاز است. برای الگوهایی در زبان گفتاری و نوشتاری، باید مدل‌های محاسباتی قوی ساخته شوند که شواهدی را برای وضعیت احساسی، عاطفی و موضعی ارائه می‌دهند و به شکل گفتار و متن، اطلاعات ضمنی تعیین می‌کنند. برای مبتنی کردن زبان به سیاق محیطی در سیستم‌های هوش مصنوعی که در دنیای فیزیکی مانند روباتیک فعالیت می‌کنند، به فنون جدید پردازش زبان نیاز است. در نهایت، از آنجا که نحوه ارتباط افراد در تعاملات آنلاین می‌تواند کاملاً با تعاملات صوتی تفاوت داشته باشد، مدل‌های زبان‌های مورد استفاده در این زمینه‌ها، باید به گونه‌ای تکمیل شوند که سیستم‌های هوش مصنوعی اجتماعی بتوانند با مردم تعامل مؤثرتری داشته باشند.

بخش سوم

راہبرد سوم: بررسی و فہم پیامدہای اخلاقی
قانونی و اجتماعے ہوش مصنوعی



راهبرد سوم: بررسی و فهم پیامدهای اخلاقی، قانونی و اجتماعی هوش مصنوعی

بررسی ملاحظات اخلاقی، قانونی و اجتماعی هوش مصنوعی (نسخه ۲۰۱۹)

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، فعالیت‌های تحقیق و توسعه که به پیامدها اخلاقی، قانونی و اجتماعی توسعه و استقرار سیستم هوش مصنوعی پرداخته‌اند، افزایش یافته است. درک فزاینده‌ای وجود دارد که سیستم‌های هوش مصنوعی باید «قابل اعتماد» باشند و بتوانند بسیاری از بخش‌های زندگی اجتماعی و اقتصادی شامل اشتغال، مراقبت‌های بهداشتی و تولید را دگرگون کنند. سازمان‌های بین‌المللی مانند سازمان همکاری اقتصادی و توسعه (OECD)^۱ و وزرای نوآوری G7^۲، برای افزایش اعتماد و پذیرش هوش مصنوعی، به تحقیق و توسعه آن ترغیب کرده‌اند.

طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ مطابق برنامه، در شناسایی مضامین تحقیقاتی در حریم خصوصی، بهبود انصاف، شفافیت و پاسخگویی سیستم‌های هوش مصنوعی و طراحی و معماری هوش مصنوعی اخلاقی، اطلاع‌رسانی کرده است. کنفرانس‌های تحقیقاتی اختصاص داده شده به انصاف، پاسخگویی و شفافیت در ML و سیستم‌های هوش مصنوعی

1. "OECD Initiatives on AI": <http://www.oecd.org/going-digital/ai/oecd-initiatives-on-ai.htm>.

2. "G7 Innovation Ministers' Statement on AI": <http://www.g8.utoronto.ca/employment/-2018labour-annex-b-en.html>.

رونق یافته‌اند^۱. موسسه‌های فدرال با طرح‌ها و جلسات تحقیقاتی مختلفی به این حوزه‌های حیاتی واکنش نشان داده‌اند (به بخش زیر توجه کنید).

تبیین پذیری، انصاف و شفافیت: برنامه‌های اخیر مؤسسات تحقیق و توسعه

از زمان انتشار طرح راهبردی تحقیق و توسعه هوش مصنوعی ملی ۲۰۱۶، چندین موسسه، برنامه‌های تحقیق و توسعه هوش مصنوعی برای راهبرد ۳ را آغاز کرده‌اند:

- هدف برنامه هوش مصنوعی قابل تبیین (XAI^۲) موسسه DARPA، ایجاد مجموعه فنون ML است که با حفظ سطح بالایی از عملکرد یادگیری (صحت پیش‌بینی) سیستم‌های هوش مصنوعی تبیین پذیرتری تولید می‌کنند. XAI کاربران انسانی را نیز قادر می‌کند تا نسل نوظهور سیستم‌های هوش مصنوعی را درک کرده، در جای مناسب به آن‌ها اعتماد کنند و به طور مؤثری آن‌ها را مدیریت کنند. به‌طور کلی‌تر،^۳ DOD متعهد به «هدایت اخلاق نظامی و امنیت هوش مصنوعی» به عنوان یکی از پنج اقدام کلیدی مطرح شده در رویکرد راهبردی است که تلاش‌هایش برای تسریع در پذیرش سیستم‌های هوش مصنوعی را هدایت می‌کند.
- NSF و آمازون برای حمایت مشترک از تحقیق متمرکز بر انصاف در هوش مصنوعی با همدیگر همکاری می‌کنند. هدف آنان کمک به سیستم‌های هوش مصنوعی قابل اعتمادی است که به راحتی پذیرفته شده و استقرار می‌یابند تا با چالش‌های بزرگی که جامعه با آن مواجه است، مقابله کنند. موضوعات خاص موردعلاقه شامل

1 <http://www.fatml.org/>; <https://fatconference.org/>; <http://www.aies-conference.com/>

2. explainable AI

3. Department of Defense

موارد زیر هستند (اما محدود به آن‌ها نیستند): شفافیت، قابلیت تبیین، پاسخگویی، جانبداری‌ها و اثرات سوء بالقوه، راهبردهای کاهش، اعتبار سنجی انصاف و ملاحظات فراگیر شدن.

دستور اجرایی ۲۰۱۹ ریاست جمهوری برای حفظ رهبری آمریکا در هوش مصنوعی، تأکید دارد که حفظ رهبری آمریکا در هوش مصنوعی، مستلزم تلاش هماهنگ برای ترویج پیشرفت‌ها در فناوری و نوآوری، همراه با حفاظت از آزادی‌های مدنی، حفظ حریم خصوصی و ارزش‌های آمریکایی است: (مرجع ۱)

ایالات متحده باید اعتماد و اطمینان عمومی به فناوری‌های هوش مصنوعی را ترویج کند و در کاربرد آن‌ها، آزادی‌های مدنی، حریم خصوصی و ارزش‌های آمریکایی را حفظ کند تا برای مردم آمریکا درک کاملی از توان بالقوه فناوری‌های هوش مصنوعی را فراهم کند. توسعه معماری‌های هوش مصنوعی نیازمند تحقیق و توسعه است. این توسعه شامل نگرانی‌های اخلاقی، قانونی و اجتماعی از طریق سازوکارهای فنی مانند شفافیت و تبیین پذیری است. این تحقیق و توسعه، مستلزم همکاری گسترده متخصصان فنی و ذینفعان و متخصصان حوزه‌های دیگر، از جمله علوم اجتماعی و رفتاری، حقوق، اخلاق و فلسفه است. از آنجایی که ممکن است تصمیمات اخلاقی به شدت مبتنی بر زمینه باشند یا وابسته به کاربرد باشند، همکاری با متخصصان این حوزه نیز ضرورت دارد. به منظور درک و لحاظ کردن تصمیمات و اقدامات مربوط به هوش مصنوعی و کاهش پیامدهای ناخواسته، این رویکرد میان‌رشته‌ای می‌تواند در آموزش، طراحی،

آزمون، ارزیابی و پیاده‌سازی هوش مصنوعی گنجانده شود.

بنابراین، نهادهای فدرال باید با حمایت مالی از تحقیق و گردهمایی متخصصان و ذینفعان به تقویت جامعه ذینفع رو به رشد ادامه دهند تا تحقیق و توسعه بیشتری را در باره این مسائل انجام دهند.

وقتی عامل‌های^۱ هوش مصنوعی به طور مستقل عمل می‌کنند، از آن‌ها انتظار داریم تا مطابق هنجارهای رسمی و غیررسمی رفتار کنند که ما انسان‌ها به آن‌ها باور داریم. بنابراین، هم قانون و هم اخلاق به عنوان نیروهای نظم اجتماعی بنیادی، هر دو از رفتار سیستم‌های هوش مصنوعی خبر داده و آن‌ها را مورد قضاوت قرار می‌دهند. نیازهای تحقیقاتی غالب، هم شامل درک مفاهیم اخلاقی، قانونی و پیامدهای اجتماعی هوش مصنوعی و هم شامل روش‌های در حال توسعه برای طراحی هوش مصنوعی هستند که با اصول اخلاقی، قانونی و اجتماعی سازگار باشند. نگرانی‌های حریم خصوصی نیز باید در نظر گرفته شوند؛ اطلاعات بیشتر در مورد این موضوع را می‌توان در راهبرد تحقیقات حفظ حریم خصوصی ملی یافت.^۲

مانند هر فن‌آوری، ما از طریق اصول قانون و اخلاق از مصارف قابل قبول هوش مصنوعی مطلع خواهیم شد؛ چالش این است که چگونه این اصول را به این فناوری جدید اعمال کنیم، به ویژه آن‌هایی که شامل خودمختاری، عاملیت و کنترل می‌شوند.

همان طور که در «اولویت‌های تحقیقاتی هوش مصنوعی قوی و سودمند» روشن می‌شوند^۳، به منظور ساخت سیستم‌هایی که بسیار خوب رفتار می‌کنند، باید تصمیم بگیریم که رفتار خوب در هر حوزه کاربرد چه معنایی دارد.

1 agents

<https://www.nitrd.gov/pubs/NationalPrivacyResearchStrategy.pdf>

۲. «نامه سرگشاده: اولویت‌های تحقیقاتی برای هوش مصنوعی قوی و سودمند» (موسسه آینده زندگی). <http://futureoflife.org/ai-open-letter>.

این بُعد اخلاقی با این سؤالات ارتباط نزدیکی دارند که چه فنون مهندسی در دسترس هستند، این فنون چقدر قابل اعتماد هستند و چه مصالحه‌ای صورت می‌گیرد، یعنی در تمام حوزه‌هایی که علم کامپیوتر، یادگیری ماشین و تخصص وسیع‌تر هوش مصنوعی ارزشمند هستند. تحقیقات در این زمینه می‌توانند از دیدگاه‌های چند رشته‌ای بهره‌مند شوند که شامل متخصصان علوم کامپیوتر، علوم اجتماعی و رفتاری، اخلاق، علوم زیست-پزشکی، روانشناسی، اقتصاد، حقوق و تحقیقات سیاست‌گذاری است. تحقیقات بیشتر در حوزه‌هایی که در داخل و خارج از حوزه NITRD (حوزه فن‌آوری اطلاعات مرتبط، یعنی فن‌آوری اطلاعات و رشته‌هایی که قبلاً ذکر شد) مورد نیاز است تا از تحقیق توسعه و استفاده از سیستم‌های هوش مصنوعی و تأثیرات آن‌ها بر جامعه اطلاع‌رسانی کنند.

در بخش‌های فرعی زیر چالش‌های کلیدی تحقیقات فن‌آوری اطلاعات در این حوزه را بررسی می‌کنیم.

بهبود انصاف، شفافیت و پاسخگویی آگاهانه

نگرانی‌های بسیاری در مورد آسیب‌پذیری الگوریتم‌های داده بنیان هوش مصنوعی نسبت به خطا و سو استفاده و عواقب احتمالی برای جنسیت، سن، نژاد یا طبقات اقتصادی ابراز شده است. جمع‌آوری مناسب و استفاده از داده برای سیستم‌های هوش مصنوعی، در این زمینه، نشان‌دهنده چالش مهمی است. با این وصف، فراتر از مسائلی که صرفاً مربوط به داده‌ها هستند، پرسش‌های بزرگ‌تری در مورد طرح هوش مصنوعی پیش می‌آیند تا به طور ذاتی عادل،

منصف، شفاف و جوابگو باشند. محققان باید یاد بگیرند که چطور این سیستم‌ها را طراحی کنند تا اعمال و تصمیم‌گیری آن‌ها شفاف بوده و به راحتی توسط انسان‌ها تفسیر شوند و در نتیجه برای هر جانب‌داری که ممکن است حاوی آن‌ها باشند، قابل بررسی شوند، نه اینکه این جانب‌داری‌ها را فقط یاد بگیرند و تکرار کنند. در مورد چگونگی بازنمایی و «کدگذاری» سیستم‌های ارزش و باور، مسائل عقلانی جدی وجود دارند. در ضمن، دانشمندان باید مطالعه کنند که چه میزان ملاحظات عدالت و انصاف را می‌توان در سیستم طراحی کرد و چگونه این کار را در محدوده فنون مهندسی کنونی انجام داد.

ساخت هوش مصنوعی اخلاقی

ورای فرضیات بنیادی عدالت و انصاف، نگرانی‌های دیگری در این باره وجود دارد که آیا سیستم‌های هوش مصنوعی می‌توانند رفتاری را نشان دهند که سازگار با اصول کلی اخلاقی باشند. چگونه ممکن است پیشرفت‌های هوش مصنوعی، سؤالات جدید مربوط به ماشین را در اخلاق پیش بیاورند، یا چه مصارفی از هوش مصنوعی بایستی غیراخلاقی تلقی شود؟ اخلاق ذاتاً یک سؤال فلسفی است، درحالی‌که فن‌آوری هوش مصنوعی به مهندسی بستگی دارد و توسط آن محدود می‌شود. بنابراین، محققان باید در محدوده آنچه از لحاظ فن‌آوری امکان‌پذیر است، برای توسعه الگوریتم‌ها و معماری‌های سازگار با قوانین، هنجارهای اجتماعی و اخلاق موجود تلاش کنند؛ بدیهی است که این کار بسیار چالش‌برانگیز است. اصول اخلاقی معمولاً با درجات

مختلفی از ابهام بیان می‌شوند و به سختی می‌توان آن‌ها را به یک سیستم و طراحی الگوریتم دقیق تبدیل کرد. در ضمن، وقتی که سیستم‌های هوش مصنوعی، به ویژه با انواع جدید الگوریتم‌های تصمیم‌گیری مستقل، با دوره‌های اخلاقی مبتنی بر سیستم‌های ارزشی مستقل و احتمالاً متناقض مواجه می‌شوند، مشکلاتی نیز وجود دارند. موضوعات اخلاقی بر اساس فرهنگ، دین و باورها تغییر می‌کنند. با این وصف، چارچوب مرجع اخلاقی قابل قبول را می‌توان توسعه داد تا هدایتگر استدلال و تصمیم‌گیری سیستم هوش مصنوعی باشد و نتایج و اعمال آن را تبیین و توجیه کند. برای تولید مجموعه داده‌هایی برای آموزش به یک رویکرد چند رشته‌ای نیاز است که یک سیستم ارزشی مناسب را منعکس کند، از جمله، مثال‌هایی هستند که هنگام مواجهه با مسائل اخلاقی دشوار یا با ارزش‌های متناقض، نشان‌دهنده رفتار مرجح هستند. این مثال‌ها می‌توانند شامل «موارد گوشه‌ای» قانونی یا اخلاقی باشند که با یک نتیجه یا داوری مشخص می‌شوند که برای کاربر آن شفاف است. هوش مصنوعی برای حل تعارض ناشی از ارزش‌ها به روش‌های مکفی نیاز دارد که در آن‌ها، سیستم شامل اصولی است که می‌تواند در جایی که قوانین سخت عملی نیستند، واقعیت‌های شرایط پیچیده را نشان دهد.

طراحی معماری برای هوش مصنوعی اخلاقی

در تحقیقات بنیادین، برای تعیین نحوه طراحی بهترین معماری‌های طراحی سیستم‌های هوش مصنوعی که از استدلال اخلاقی استفاده می‌کنند، باید پیشرفت بیشتری صورت گیرد. رویکردهای مختلفی پیشنهاد شده‌اند، مانند یک معماری نظارت دو طرفه که هوش

۱. ای. اتزیونی و او. اتزیونی، «طراحی سیستم‌های AI که از قوانین و ارزش‌های ما پیروی می‌کنند»، نشریه ارتباطات ACM، ۲۰۱۶، ۲۹-۳۱.

مصنوعی عملیاتی را از یک کارگزار ناظر جدا می‌کند که مسئول ارزیابی اخلاقی یا قانونی هر اقدام عملیاتی است.^۱ دیدگاه دیگر، ترجیح مهندسی ایمنی است که در آن، یک چارچوب مفهومی دقیق برای معماری عامل هوش مصنوعی مورد استفاده قرار می‌گیرد تا اطمینان حاصل شود که رفتار هوش مصنوعی ایمن بوده و برای انسان‌ها ضرری ندارد.^۲ سومین روش، تدوین یک معماری اخلاقی با استفاده از ترکیب مجموعه‌ای از اصول نظری و محدودیت‌های منطقی رفتار سیستم هوش مصنوعی است که عمل را محدود می‌کند تا با آموزه‌های اخلاقی سازگار باشد.^۳ همان‌طور که سیستم‌های هوش مصنوعی عمومی‌تر می‌شوند، معماری‌های آن‌ها احتمالاً شامل سیستم‌های فرعی خواهند بود که می‌توانند مسائل اخلاقی را در سطوح چندگانه قضاوت در نظر بگیرند، از جمله: الگوی پاسخ سریع که منطبق بر قواعد است، استدلال آگاهانه برای واکنش‌های کندتر برای توصیف و توجیه اعمال، سیگنال دهی اجتماعی برای نشان دادن قابلیت اعتماد به کاربر و فرآیندهای اجتماعی که در مقیاس‌های زمانی طولانی عمل می‌کنند تا این سیستم بتواند از هنجارهای فرهنگی تبعیت کند. محققان نیاز دارند تا بر نحوه رسیدگی مطلوب به طراحی کلی سیستم‌های هوش مصنوعی تمرکز کنند که با اهداف اخلاقی، قانونی و اجتماعی همسو هستند.

1. آر. وی. یامپولسکی، «مهندسی ایمنی هوش مصنوعی»: چرا اخلاق ماشین، رویکرد غلطی است. در کتاب *Philosophy and Theory of Artificial Intelligence*, ed. V.C. Muller (Heidelberg: Springer Verlag, 2013), 396-389.

2. آر. سی. آرکین، «حاکمیت رفتار قانونی: گنجاندن اخلاق در معماری روبات هیبرید عامدانه / راکتیو»، نشریه 2007, 11-07-Georgia Institute of Technology Technical Report, GIT-GVU.

3. بی. کوپرس، «اخلاقیات انسان گونه و اخلاق برای روباتها»، کارگاه AAAI-16 در باره هوش مصنوعی، اخلاق و جامعه، 2016. <https://web.eecs.umich.edu/~kuipers/papers/Kuipers-aaaiws.2016.pdf>

بخش چهارم

راهبرد ۴: تضمین ایمنی و امنیت
سیستم‌های هوش مصنوعی
نسخه ۲۰۱۹



ایجاد سیستم‌های قوی و قابل اعتماد هوش مصنوعی

از زمان انتشار نسخه ۲۰۱۶ طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی، رشد سریعی در درک علمی و اجتماعی از امنیت و ایمنی هوش مصنوعی صورت گرفته است. بیشتر این دانش جدید به شناسایی مسائل جدید کمک کرده است: حالا بسیار روشن‌تر است که چگونه می‌توان سیستم‌های هوش مصنوعی ساخت که کار غلطی را انجام ندهند، کار اشتباهی را یاد بگیرند یا کار اشتباهی را آشکار کنند، مثلاً از طریق مثال‌های معکوس، آلوده کردن داده‌ها و وارونه کردن مدل. متأسفانه، راه‌حل‌های فنی برای حل این مسائل ایمنی و امنیت هوش مصنوعی همچنان مبهم هستند.

برای رفع همه این مسائل، باید ایمنی و امنیت سیستم‌های هوش مصنوعی در تمام مراحل چرخه عمر سیستم هوش مصنوعی، از طراحی اولیه و ساخت داده/مدل گرفته تا تأیید و اعتبار سنجی، استقرار، بهره‌برداری و نظارت در نظر گرفته شوند. در واقع، مفهوم «ایمنی (یا امنیت) آگاهانه» ممکن است تصور نادرستی را القا کند که این‌ها صرفاً به طراحان سیستم مربوط هستند؛ که درست

نیست و باید در سراسر چرخه زندگی سیستم در نظر گرفته شوند، نه فقط در مرحله طراحی و لذا باید بخش مهمی از مراحل تحقیق و توسعه هوش مصنوعی باشند.

وقتی مؤلفه‌های هوش مصنوعی به سیستم‌های یا اطلاعات دیگر متصل می‌شوند که آنها هم باید ایمن یا امن باشند، آسیب‌پذیری‌ها و الزامات عملکرد هوش مصنوعی (مثلاً، میزان بسیار پایین مثبت-کاذب و منفی-کاذب، زمان کار بر روی حجم بالایی از داده‌ها) در سیستم‌های بزرگ‌تر جاری می‌شود. این چالش‌ها ایستا نیستند، چرا که سیستم‌های هوش مصنوعی به رشد خود ادامه می‌دهند، به احتمال زیاد آن‌ها از جهت پیچیدگی رشد می‌کنند و این امر، تأیید و اعتبارسنجی عملکرد صحیح یا حفظ حریم خصوصی اطلاعات را دشوار می‌کند. در ضمن، این پیچیدگی، به طور فزاینده‌ای تبیین تصمیمات را به شیوه‌هایی که سطوح بالایی از اعتماد کاربران انسانی را توجیه می‌کنند، دشوار می‌کند (راهبرد ۳ را ببینید).

ایجاد اعتماد نسبت به استفاده از هوش مصنوعی، هم‌اکنون و در آینده، یک مسئله حیاتی است که نیازمند سرمایه‌گذاری‌های تحقیق و توسعه دولت فدرال (بخش زیر را ببینید) و تلاش‌های مشترک میان دولت، صنعت، دانشگاهیان و جامعه مدنی است. ممکن است قابلیت اعتماد مهندسی سیستم‌های هوش مصنوعی، از شیوه‌های موجود در مهندسی ایمنی در سایر حوزه‌ها عاریت گرفته شود که یاد گرفته‌اند چگونه رفتار سوء بالقوه سیستم‌های غیر-هوش مصنوعی مستقل یا نیمه خودکار را توضیح دهند. با این وصف، مسائل خاص هوش مصنوعی به معنی آن است که برای اثبات این که یک سیستم

مبتنی بر هوش مصنوعی با مشخصات آن مطابقت دارد، فنون بدیع برای تحلیل، آزمایش، تأیید رسمی و ترکیب برنامه، از اهمیت حیاتی برخوردار خواهند بود - یعنی آن سیستم دقیقاً همان کاری را می‌کند که باید انجام دهد و نه بیشتر. این مسائل در سیستم‌های مبتنی بر هوش مصنوعی تشدید می‌شوند که می‌توان به راحتی آن‌ها را به روش‌هایی که ممکن است پیامدهای امنیتی عمیقی داشته باشند، فریب داد، از آن طفره رفت و گمراه کرد. یک حوزه تحقیق نوظهور، ML تخصصی است که به واکاوای تحلیل آسیب‌پذیری در الگوریتم‌های ML و فنون الگوریتمی می‌پردازد که یادگیری قوی‌تری را به وجود می‌آورند. حملات معروف به ML شامل موارد زیر هستند: حملات تخصصی اجتناب کلاسیفایر^۱ است که در آن، مهاجم برای فرار از ردیابی، رفتار خود را تغییر می‌دهد و حملات مسموم کننده که در آن، داده‌های آموزشی تخریب می‌شوند. نیاز فزاینده‌ای به تحقیقاتی وجود دارند که اولاً به طور منظم فضای تخصصی را بررسی می‌کنند که به ML و دیگر سیستم‌های مبتنی بر هوش مصنوعی حمله می‌کنند؛ ثانیاً الگوریتم‌هایی را طراحی می‌کنند که در برابر انواع تخصصات، ثبات و استحکام را تضمین می‌کنند.

باید روش‌هایی ابداع شوند تا ایجاد، ارزیابی، استقرار و جلوگیری از نفوذ در هوش مصنوعی را ایمن و امن کنند. این روش‌ها باید گسترش یابند تا با قابلیت و پیچیدگی هوش مصنوعی تطبیق یابند. ارزیابی این روش‌ها به معیارهای جدید، چارچوب‌های کنترل و محک‌هایی برای آزمایش و ارزیابی ایمنی سیستم‌هایی نیاز خواهند داشت که روز به روز قوی‌تر می‌شوند. هم روش‌ها و هم معیارها

1. adversarial classifier evasion attacks

باید تلفیق کننده عوامل انسانی با موارد زیر باشند: اهداف امن هوش مصنوعی که مطابق با اهداف طراحان انسانی تعریف شده باشند؛ عملیات ایمن هوش مصنوعی که مطابق با عادات کاربران انسانی تعریف شده‌اند و معیارهای هوش مصنوعی ایمن که توسط ادراک ارزیابان انسانی تعریف شده‌اند. تولید روش‌ها و معیارهایی برای ایمنی سیستم‌های هوش مصنوعی که محرک انسانی داشته و از سوی انسان قابل درک هستند، سیاست‌گذاران، بخش خصوصی و عموم را قادر خواهد ساخت تا به دقت چشم‌انداز ایمنی هوش مصنوعی را مورد قضاوت قرار دهند و متناسب با آن گام بردارند.

ایمنی و امنیت هوش مصنوعی: برنامه‌های اخیر تحقیق و توسعه

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، تعدادی از مؤسسات تلاش‌هایی را در حمایت از راهبرد ۴ آغاز کرده‌اند:

- DOT^۱ در اکتبر سال ۲۰۱۸ راهنمای فدرال جدیدی را برای وسایل نقلیه خودکار منتشر کرد که از ادغام ایمن اتوماسیون با سیستم حمل‌ونقل سطح چندگانه پشتیبانی می‌کند. آمادگی برای آینده حمل‌ونقل: وسایل نقلیه خودکار^۲، پیشرفتی است در اصول یکپارچگی DOT در باره ادغام ایمن وسایل نقلیه خودکار. علاوه بر این، این سند راهنمای ایمنی قبلی را تکرار می‌کند، رهنمودهای جدید ایمنی چندگانه را ارائه می‌دهد و هم‌زمان با تکامل این فن‌آوری جدید، فرآیندی را برای کار با DOT طرح می‌کند. از ماه مه ۲۰۱۹، چهارده شرکت در باره نحوه پیش‌بینی ایمنی در

1. Department of Transportation
2. <https://www.transportation.gov/av/3>

طراحی و آزمایش سیستم‌های رانش خودکار، جزئیات خود-ارزیابی ایمنی داوطلبانه را منتشر کردند.^۱

• در دسامبر ۲۰۱۸، IARPA^۲ دو برنامه را در زمینه امنیت هوش مصنوعی اعلام کرد: سیستم‌های یادگیری هوشمند امن، مطمئن (SAILS)^۳ و تروجانها در هوش مصنوعی (TrojAI)^۴. DARPA برنامه دیگری را در فوریه ۲۰۱۹ اعلام کرد، تضمین مقاومت هوش مصنوعی در مقابل کلاهبرداری (GARD)^۵. روی هم رفته، هدف این برنامه‌ها، مبارزه با طیفی از حملات به سیستم‌های هوش مصنوعی است.

• همان طور که در راهبرد ۳ اشاره شد، به عنوان یکی از پنج اقدام کلیدی مطرح شده در رویکرد راهبردی مبنی بر شتاب بخشیدن به تلاش‌های خود برای پذیرش سیستم‌های هوش مصنوعی، DoD متعهد به «پیشسازی در اخلاق نظامی و ایمنی هوش مصنوعی» است.^۶

قبل از اینکه سیستم هوش مصنوعی به طور گسترده‌ای مورد استفاده قرار گیرند، لازم است اطمینان حاصل شود که سیستم به صورت ایمن و امن و به شیوه‌ای کنترل شده عمل می‌کند. برای رفع چالش ایجاد سیستم‌های هوش مصنوعی که قابل اعتماد، قابل اتکا و مطمئن هستند، به تحقیقات نیاز داریم. سیستم‌های هوش مصنوعی، مانند دیگر سیستم‌های پیچیده، با چالش‌های ایمنی و امنیتی مهمی رو به رو هستند که ناشی از موارد زیر است:^۷

1. <https://www.nhtsa.gov/automated-driving-systems/voluntary-safety-self-assessment>
2. Intelligence Advanced Research Projects Activity
3. <https://www.iarpa.gov/index.php/research-programs/sails>
4. <https://www.iarpa.gov/index.php/research-programs/trojai>
5. <https://www.darpa.mil/news-events/06-02-2019>

۶. «خلاصه راهبرد ۲۰۱۸ وزارت دفاع در باره هوش مصنوعی».

<https://media.defense.gov/2019/Feb/1/1-1-2002088963/12/SUMMARY-OF-DOD-AI-STRATEGY.PDF>

۷. جی. برنشتاین، «قشه راه مستقل وزارت دفاع - جامعه ذینفعان مستقل» ارائه شده در NDIA 16th Annual Science & Engineering Technology Conference, March 2015.

- محیط‌های پیچیده و نامطمئن: در بسیاری از موارد، سیستم‌های هوش مصنوعی برای کار در محیط‌های پیچیده طراحی شده‌اند، با تعداد زیادی از حالات بالقوه که نمی‌توانند به طور کامل مورد بررسی یا آزمون قرار بگیرند. ممکن است یک سیستم با شرایطی مواجه شود که هرگز در طول طراحی آن لحاظ نشده است.
- رفتار نوظهور: سیستم‌های هوش مصنوعی که بعد از استقرار، یاد می‌گیرند، ممکن است رفتار سیستم تا حد زیادی توسط دوره‌های زمانی آموزش تحت شرایط بدون نظارت تعیین شود. تحت چنین شرایطی، پیش‌بینی رفتار یک سیستم دشوار است.
- تعیین نادرست هدف^۱: به خاطر دشواری تبدیل اهداف انسانی به دستورالعمل‌های کامپیوتری، ممکن است اهدافی که برای سیستم هوش مصنوعی برنامه‌ریزی شده‌اند، با اهداف برنامه‌نویس مطابقت نداشته باشند.
- تعاملات انسان-ماشین: در بسیاری از موارد، عملکرد سیستم هوش مصنوعی، اساساً تحت تأثیر تعاملات انسانی قرار می‌گیرد. در این موارد، ممکن است تغییر در واکنش‌های انسانی بر کارایی سیستم تأثیر بگذارد.^۲

برای رفع این مسائل و موارد دیگر، به سرمایه‌گذاری‌های بیشتری نیاز داریم تا ایمنی و امنیت هوش مصنوعی را در موارد زیر بالا ببریم^۳: قابلیت تبیین و شفافیت، اعتماد، تأیید و اعتبار سنجی، امنیت در برابر حملات و ایمنی بلندمدت هوش مصنوعی و تنظیم ارزش‌ها.

1. Goal misspecification

۲. جی. ام. بردشاو، آر. آر. هافمن، ام. جانسون و دی. وود، «هفت افسانه‌کننده در باره سیستم‌های مستقل» نشریه IEEE Intelligent Systems ۳۸(۳):۱۳-۲۸، ۲۰۱۳، ۵۴-۶۱.

۳. برای مثال نگاه کنید به: دی. آمودی، سی. اولاج، جی. اشتینهارت، جی. شولمان و دی. مبین، «مسائل عینی در ایمنی هوش مصنوعی»، ۲۰۱۶، arXiv:۱۶۰۶.۰۶۵۶۵۷۲؛ اس. راسل، دی. دیوی و ام. تگمارک، «اولویت‌های پژوهشی هوش مصنوعی مقاوم و سودمند»، ۲۰۱۶، arXiv:۱۶۰۲.۰۳۵۰۶؛ تی. جی. دایتریج و ای. جی. هورویتزف «اوج‌گیری دغدغه‌های مربوط به هوش مصنوعی: تأملات و جهت‌گیری‌ها»، نشریه ارتباطات ACM، ۵۸(۱۰):۱۰۵۸-۱۰۶۳، ۲۰۱۵؛ و کی. سوتالا و آر. یامپولسکی، «پاسخ به ریسک فاجعه‌آمیز AGI: مرور»، نشریه Physica Scripta، ۹۰(۱)، ۱۹ December ۲۰۱۴.

بهبود قابلیت تبیین و شفافیت

چالش اصلی تحقیق، افزایش «قابلیت تبیین» یا «شفافیت» هوش مصنوعی است. بسیاری از الگوریتم‌ها از جمله آن‌هایی که بر اساس یادگیری عمیق کار می‌کنند، برای کاربران شفاف نیستند و برای توضیح نتایج آن‌ها سازوکارهای معدودی وجود دارند. این موضوع به ویژه برای حوزه‌هایی مانند بهداشت و درمان که در آن، پزشکان برای توجیه یک تشخیص خاص یا دوره درمان به توضیحات نیاز دارند، مسئله‌ساز است. فنون هوش مصنوعی مثل القای درخت تصمیم‌گیری^۱، تبیین‌های آماده‌ای را فراهم می‌کنند، اما به طور کلی دقت کم‌تری دارند. بنابراین، محققان باید سیستم‌هایی را ابداع کنند که شفاف بوده و فی‌نفسه قادر به تبیین دلایل نتایج خود برای کاربران هستند.

ایجاد اعتماد

طراحان سیستم هوش مصنوعی برای دستیابی به اعتماد، باید سیستم‌های دقیق و قابل اطمینانی را به همراه واسط‌های کاربرپسند، طراحی کنند؛ در حالی که اپراتورها باید زمان کافی برای آموزش کافی جهت درک کارکرد سیستم و محدودیت‌های عملکرد داشته باشند. سیستم‌های پیچیده‌ای که به طور گسترده‌ای توسط مورد اعتماد کاربران قرار می‌گیرند، مانند کنترل‌های دستی برای وسایل نقلیه، معمولاً باید شفاف عمل کنند (سیستم به گونه‌ای عمل می‌کند که برای کاربر قابل رؤیت است)، قابل اعتماد باشند

(خروجی‌های سیستم از سوی کاربر مورد پذیرش قرار گیرند)، قابل حسابرسی باشند (سیستم قابل ارزیابی باشد)، قابل اتکا باشد (سیستم مطابق نظر کاربر عمل می‌کند) و قابل بازیابی باشد (کاربر بتواند در موقع ضروری کنترل را بازیابی کند). چالش مهم سیستم‌های هوش مصنوعی فعلی و آینده، کیفیت ناسازگار فن‌آوری تولید نرم‌افزار است. هم‌زمان با پیشرفت‌هایی که پیوندهای بزرگ‌تری بین انسان‌ها و سیستم‌های هوش مصنوعی برقرار می‌کنند، چالش در حوزه اعتماد، عبارت است از: همگام بودن با تغییر و افزایش قابلیت‌ها، پیش‌بینی پیشرفت‌های فناورانه در پذیرش و استفاده بلندمدت و ایجاد اصول و سیاست‌های حاکم بر مطالعه بهترین نمونه‌های طراحی، ساخت و استفاده، از جمله آموزش مناسب اپراتور برای عملیات ایمن است.

ارتقا تأیید و اعتبارسنجی

برای تأیید و اعتبارسنجی سیستم‌های هوش مصنوعی به روش‌های جدید نیاز است. «تأیید» ثابت می‌کند که یک سیستم با مشخصات رسمی مطابقت دارد؛ در حالی که «اعتبارسنجی» ثابت می‌کند که یک سیستم نیازهای عملیاتی کاربر را برآورده می‌کند. ممکن است سیستم‌های هوش مصنوعی نیاز به ابزارهای جدید داشته باشند، ابزارهای جدید ارزیابی (تعیین اینکه آیا سیستم سوء عملکرد دارد یا خیر، شاید هنگامی که خارج از محدوده پارامترهای مورد انتظار عمل می‌کند)، تشخیص (تعیین علل عملکرد نادرست) و تعمیر (تنظیم سیستم برای رفع خرابی). برای سیستم‌هایی که در مدت

زمان طولانی به صورت خودکار عمل می‌کنند، ممکن است طراحان سیستم، هر شرطی را که سیستم با آن مواجه می‌شود، در نظر نگرفته باشند. ممکن است چنین سیستم‌هایی به منظور قوی بودن و قابلیت اعتماد، به قابلیت‌هایی برای خود-ارزیابی، خود-تشخیصی و خود-تعمیری نیاز داشته باشند.

ایمن‌سازی در برابر حملات

هوش مصنوعی تعبیه شده در سیستم‌های حیاتی، باید قوی باشد تا بتواند به حوادث رسیدگی کند، اما باید برای دامنه وسیعی از حملات سایبری عمده امن باشد. مهندسی امنیت شامل درک آسیب‌پذیری یک سیستم و اقدامات کنشگرانی است که ممکن است به حمله به آن سیستم علاقه‌مند باشند. در حالی که نیازهای تحقیق و توسعه امنیت سایبری با جزییات بیشتر در برنامه راهبردی تحقیق و توسعه امنیت سایبری فدرال ۲۰۱۶ NITRD مورد اشاره قرار می‌گیرد¹، اما برخی از ریسک‌های امنیت سایبری، مختص سیستم‌های هوش مصنوعی هستند. برای مثال، یک حوزه تحقیقاتی کلیدی «یادگیری ماشین تخصصی» است که به واکاوی حد و اندازه‌ای می‌پردازد که سیستم‌های هوش مصنوعی می‌توانند با موارد زیر سازش داشته باشند: «آلوده کردن» داده‌های آموزشی، اصلاح الگوریتم‌ها یا ایجاد تغییرات ظریف در شیئی که مانع تشخیص صحیح خودش می‌شود (مثلاً، اندام مصنوعی که از کار سیستم‌های تشخیص چهره جلوگیری می‌کنند). پیاده‌سازی هوش مصنوعی در سیستم‌های امنیت سایبری

1. <https://www.nitrd.gov/pubs/-2016Federal-Cybersecurity-Research-and-Development-Strategic-Plan.pdf>; this is being updated in 2019.

که نیازمند درجه بالایی از استقلال هستند نیز، حوزه‌های برای مطالعه بیشتر است. یک مثال جدید از پژوهش در این زمینه، چالش بزرگ سایبری^۱ ARPA است که شامل کارگزاران هوش مصنوعی است که به صورت خودکار به تحلیل و مقابله با حملات سایبری می‌پردازند.^۲

دستیابی به ایمنی بلند مدت هوش مصنوعی و همترازی ارزش‌ها

ممکن است سیستم‌های هوش مصنوعی در نهایت، قادر به «خود بهبودی بازگشتی» باشند که در آن، به جای اینکه تغییرات نرم‌افزاری اساسی، توسط برنامه نویسان انسانی انجام شود، توسط خود نرم‌افزار صورت می‌گیرد. برای اطمینان از ایمنی سیستم‌های خوداصلاح، نیاز به تحقیقات بیشتری برای توسعه موارد زیر وجود دارد: معماری‌های خود-پایشی که سازگاری رفتاری سیستم‌ها را با اهداف اصلی طراحان انسانی مورد بررسی قرار می‌دهند؛ راهبردهای محصورسازی برای جلوگیری از انتشار سیستم‌هایی که در حال ارزیابی هستند؛ یادگیری ارزش که در آن، ارزش‌ها، اهداف یا نیات کاربران توسط سیستم قابل استنتاج هستند و چارچوب‌های ارزشی که احتمالاً در مقابل خود-اصلاحی مقاوم هستند.

1. DARPA's Cyber Grand Challenge
2. https://archive.darpa.mil/CyberGrandChallenge_CompetitorSite/

بخش پنجم

راهبرد ۵: توسعه مجموعه داده‌ها
و محیط‌های عمومی مشترک



راهبرد ۵: توسعه مجموعه داده‌ها و محیط‌های عمومی مشترک برای آموزش و آزمایش هوش مصنوعی: نسخه ۲۰۱۹

افزایش دسترسی به مجموعه داده‌ها و چالش‌های مربوطه

در زمان انتشار برنامه راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، مجموعه داده‌ها و محیط‌های در دسترس عموم نقش مهمی در پیشبرد هوش مصنوعی، به خصوص در حوزه‌هایی مانند بینایی رایانه‌ای، پردازش زبان طبیعی و تشخیص گفتار ایفا می‌کردند. ImageNet^۱، با بیش از ۱۴ میلیون شی برچسب دار، همراه با چالش‌های جامعه بینایی رایانه‌ای (مثلاً، چالش تشخیص بینایی با مقیاس بزرگ ImageNet^۲ که الگوریتم‌های آشکارسازی اشیا و طبقه‌بندی تصویر را ارزیابی می‌کند) نقش حیاتی ویژه‌ای در جامعه ایفا کرده‌اند. از آنجا که اپلیکیشن‌های تبدیلی ML در بسیاری از حوزه‌های کاربردی مانند مراقبت‌های بهداشتی، پزشکی و جوامع هوشمند و متصل یافت می‌شوند، نیاز به مجموعه داده‌های در دسترس عموم در حوزه‌های خاص دامنه^۳ رو به فزونی بوده است. اهمیت مجموعه داده‌ها و مدل‌ها - به خصوص، داده‌ها و مدل‌های دولت فدرال - به صراحت در فرمان ۲۰۱۹ ریاست جمهوری در باره حفظ رهبری آمریکا در هوش مصنوعی بیان می‌شود (مرجع ۱):

1. <http://www.image-net.org/>
 2. <http://www.image-net.org/challenges/LSVRC/>
 3. domain-specific areas

روسای تمام مؤسسات باید اطلاعات و مدل‌های فدرال خود را مورد بررسی قرار دهند تا دسترسی و استفاده جامعه تحقیقاتی غیر فدرال بزرگ‌تر را به شیوه‌ای افزایش دهند که آن جامعه منتفع شود، در عین حال از ایمنی، امنیت، حریم خصوصی و محرمانگی حفاظت شود. به ویژه، مؤسسات باید داده‌ها و مستندسازی موجودی مدل خود را بهبود بخشند تا کشف و قابلیت استفاده را فعال کرده و بر اساس بازخوردهای جامعه تحقیق هوش مصنوعی نسبت به اولویت‌بندی بهبود دسترسی و کیفیت داده‌ها و مدل‌های هوش مصنوعی اقدام کنند.

در سال ۲۰۱۸ در NTSC، برای هماهنگ کردن تلاش‌های فدرال در داده‌های آزاد و FAIR (قابل یافتن، در دسترس، قابل تعامل^۱ و قابل استفاده مجدد)، یک کمیته فرعی جدید در زمینه علوم آزاد^۲ ایجاد شد. برای توسعه ابزارها و منابعی که شناسایی، استفاده و دست‌کاری مجموعه داده‌های مرتبط (شامل مجموعه داده‌های فدرال)، تأیید منشأ داده‌ها و احترام به سیاست استفاده مناسب را آسان‌تر می‌کنند، به سرمایه‌گذاری‌های تحقیق و توسعه نیاز است. ممکن است بسیاری از این مجموعه داده‌ها، بدون سرمایه‌گذاری در برچسب‌زنی و گلچین کردن، خودشان استفاده محدودی در زمینه هوش مصنوعی داشته باشند. مؤسسات فدرال باید با ذینفعان هوش مصنوعی همکاری و تعامل داشته باشند تا اطمینان حاصل شود که مجموعه داده‌ها و مدل‌هایی که به خوبی بررسی شده‌اند و برای به اشتراک‌گذاری منتشر می‌شوند، آماده و مناسب استفاده هستند و اینکه آن‌ها به عنوان استانداردها و هنجارها تکامل می‌یابند. در نهایت، توسعه و

1. interoperable
2. Open Science

اتخاذ بهترین شیوه‌ها و استانداردها در مستندسازی داده‌ها و خاستگاه مدل، موجب ارتقا قابلیت اعتماد و استفاده مسئولانه از فن‌آوری‌های هوش مصنوعی خواهد شد.

از سال ۲۰۱۶ نیز نگرانی‌ها در مورد محتوای داده‌ها، مانند جانب‌داری بالقوه (راهبرد ۳ را ببینید)^۱ یا نشت اطلاعات خصوصی افزایش یافته است. طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ اشاره کرده که «توسعه و به اشتراک‌گذاری مجموعه داده‌ها باید ... از قوانین و مقررات حاکم پیروی کنند و به روشی اخلاقی اجرا شوند». پروژه InSight که تحت حمایت DOT است، چنین امکان دسترسی سازمان‌یافته دقیقی به اطلاعات جمع‌آوری‌شده در خلال مطالعه Naturalistic Driving را فراهم می‌کند (بخش پایین را ببینید). طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی نیز متذکر شده که «برای اطمینان از اشتراک‌گذاری ایمن داده‌ها، نیاز به فن‌آوری‌های جدید وجود دارد، چرا که صاحبان داده‌ها، هنگام به اشتراک‌گذاری داده‌های خود با جامعه پژوهشی، مرتکب ریسک می‌شوند.»

مثلاً، CryptoNets^۲ به شبکه‌های عصبی امکان می‌دهد تا روی داده‌های رمزنگاری شده کار کنند. این امر موجب حصول اطمینان در باره حفظ محرمانگی داده‌ها می‌شود، زیرا در شبکه‌های عصبی نیازی به کلیدهای رمزگشایی نیست. در ضمن، پژوهشگران توسعه فنون جدید ML را آغاز کرده‌اند. این فنون از چارچوب حریم

۱. امیلی ام. بندر و باتالیا فریدمن، «گزاره‌های داده‌ای NLP: به سوی کاهش سوگیری سیستم و توانمندسازی علوم بهتر»، نشریه

Transactions of the Association for Computational Linguistics ۶:۰۴-۵۸۷(۲۰۱۸) ۶.

۲. لیلین کالینسکان، جوانا جی. برابسون و آرونند نارایانان، «معناشناسی استخراج شده خودکار از زبان‌شناسی پیکره‌ای، حلای سوگیری‌های انسان گونه»، نشریه

Science ۲۰۱۷. Apr ۳۵۶(۶۳۳۴):۱۸۳-۱۸۶.

۳. ران گیلا-باچاراج، ناتان داوین، کیم لین، کریستین لاور، مایکل ناهریج، جان ورنسینگ، «کریپتوتن: اعمال شبکه‌های عصبی به داده‌های رمزنگاری شده با بازده و دقت بالا»، ۲۰۱۶، کنفرانس بین‌المللی یادگیری ماشین <http://proceedings.mlr.press/v48/> ۲۰۱۰-۴۸.۲۰۱

خصوصی تفاضلی استفاده می‌کنند تا بر روی داده‌های مورد استفاده، ضمانت‌های حریم خصوصی قابل کمی سازی را ارائه کنند^۱. در عین حال، روش‌های حفظ حریم خصوصی باید به اندازه کافی قابل تبیین و شفاف بمانند تا به محققان کمک کنند تا آن‌ها را تصحیح کرده و آن‌ها را ایمن، کارآمد و دقیق کنند. علاوه بر این، هوش مصنوعی می‌تواند کشفیاتی فراتر از حوزه اصلی یا قلمرو خود داشته باشد، بنابراین، محققان باید در دسترسی به داده‌ها یا کشفیات نهادهای متخصص، از خطرات بالقوه آگاه باشند.

مجموعه داده‌ها و محیط‌های عمومی مشترک برای آموزش و آزمایش هوش مصنوعی: برنامه‌های اخیر مؤسسات تحقیق و توسعه.

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، تعدادی از مؤسسات تلاش‌هایی را در حمایت از راهبرد ۵ آغاز کرده‌اند:

- DOT حامی مالی برنامه تحقیقاتی دومین شاهراه راهبردی (SHRP2)، مطالعه پیشران طبیعت‌گرایانه (NDS)^۲ شد که بیش از ۵٫۴ میلیون سفر بیش از ۳٫۴۰۰ راننده و وسیله را نقلیه ثبت کرد. یک سیستم گردآوری داده خودرو (DAS)، موارد زیر را با هم تلفیق می‌کند: داده‌های گردآوری شده و ذخیره شده رادار جلو، چهار دوربین ویدیو، سرعت سنج‌ها، اطلاعات شبکه وسیله نقلیه، سیستم موقعیت‌یابی جغرافیایی و یک ردیاب خط خودرو. هنگامی

۱. مارتین آبادی، اندی چو، یان گودفلو، اچ. برندان مک موهان، ایلینا میرونوف کونال تالوار و لی ژانگ، «یادگیری عمیق با حریم خصوصی تفاضلی»، کنفرانس ACM ۲۳ در باره امنیت کامپیوتر و ارتباطات، ۲۰۱۶: ۳۰۸-۳۱۸.

2. Naturalistic Driving Study (DOT)

3. <https://insight.shrp2nds.us/>

که وسایل نقلیه شرکت کنندگان در حال کار بودند، داده‌های DAS به طور مداوم ثبت شدند. در حالی که خلاصه‌ای از داده‌های DAS، عمومی هستند، دسترسی به جزئیات مجموعه داده‌ها، مستلزم آموزش ماهرانه اخلاق پژوهشی است.

- **VA¹ Data Commons**، با ابزارهایی برای توانمندسازی ML و هوش مصنوعی در حال ساخت بزرگ‌ترین مجموعه داده پزشکی-ژنوم شناسی در جهان است که با ترجیحات کهنه سربازان هدایت می‌شود. این تلاش، در حال ارتقا استانداردها، قوانین و دستورهای اجرایی NIST^۲ است.

- **GSA^۳** (اداره خدمات عمومی) در حال کار است تا استفاده از منابع رایانش ابری را برای تحقیق و توسعه هوش مصنوعی با بودجه فدرال میسر کند. **Data.gov** و **code.gov** مستقر در **GSA**، حاوی بیش از ۲۴۶،۰۰۰ مجموعه داده و کد از کلیه مؤسسات هستند و به طور خودکار از مجموعه داده‌های منتشر شده از سوی مؤسسات استفاده می‌کند.

- طرح زیرساخت تحقیقات علم و فن‌آوری برای اکتشاف، آزمایش و پایداری (STRIDES) موسسه NIH، شراکتی است با ارائه‌دهندگان خدمات ابری پیشرو در صنعت که امکان دسترسی محقق به دارایی‌های داده‌های اصلی را فراهم می‌کنند و از سوی کل NIH تأمین مالی می‌شود و در محیط‌های ابری ذخیره می‌شوند.

داده‌ها، بدون توانایی تأثیرگذاری منابع محاسباتی بر مجموعه داده‌های عمومی با مقیاس بزرگ، به تنهایی کاربرد چندانی ندارند.

اهمیت منابع محاسباتی برای تحقیق و توسعه هوش مصنوعی در فرمان ۲۰۱۹ ریاست جمهوری در باره حفظ رهبری آمریکا در هوش مصنوعی، ذکر شده است: (مرجع ۱)

وزرای دفاع، بازرگانی، بهداشت و خدمات انسانی و انرژی، رئیس اداره ملی علوم هوانوردی و فضایی و مدیر بنیاد ملی علوم باید تا حد لزوم و مطابق با قانون حاکم، از طریق موارد زیر به تخصیص منابع محاسباتی با عملکرد بالا برای کاربردهای مربوط به هوش مصنوعی اولویت بدهند: (۱) افزایش تخصیص احتیاطی منابع و ذخیره منابع یا (۲) هر سازوکار مناسب دیگر.

و:

... کمیته منتخب، با هماهنگی با اداره خدمات عمومی (GSA)، گزارشی را به رئیس جمهور ارائه کند و طی آن، در مورد استفاده بهتر از منابع رایانش ابری در تحقیق و توسعه هوش مصنوعی فدرال توصیه‌هایی را ارائه کند.

نیاز به ظرفیت محاسباتی در بسیاری از چالش‌های هوش مصنوعی به سرعت در حال افزایش است.^۱ ممکن است سرمایه‌گذاری فدرال، امکانات محاسباتی برای تحقیقات با سرمایه‌گذاری فدرال را تأمین کند. با این حال، ممکن است برخی شرکت‌ها و دانشگاه‌ها نیازهای محاسباتی بیشتری داشته باشند. به طور کلی، برای مطالعه و سرمایه‌گذاری در منابع محاسباتی مشترک برای ترویج تحقیق و توسعه هوش مصنوعی، یک نیاز ملی وجود دارد.

افزایش منافع هوش مصنوعی ادامه خواهد داشت، اما تنها تا جایی که منابع آموزشی و آزمایش هوش مصنوعی توسعه یافته و در

1. <https://openai.com/blog/ai-and-compute/>

دسترس قرار می‌گیرند. تنوع، عمق، کیفیت و دقت مجموعه داده‌های آموزشی و سایر منابع به طور قابل توجهی بر عملکرد هوش مصنوعی اثر می‌گذارند. بسیاری از فناوری‌های مختلف هوش مصنوعی به داده‌های کیفیت بالا برای آموزش و آزمایش و همچنین بسترهای آزمایش پویا و محیط‌های شبیه‌سازی احتیاج دارند. این موضوع بیش از یک سؤال فنی صرف بوده و یک چالش «خیر عمومی» است، زیرا در صورتی که آموزش و آزمایش هوش مصنوعی تنها به چند نهاد محدود شود که از قبل دارای مجموعه داده‌ها و منابع ارزشمندی دارند، پیشرفتی صورت نخواهد گرفت. در عین حال ما باید به طور هم‌زمان به حقوق و منافع تجاری فردی موجود در داده‌ها احترام بگذاریم. برای توسعه مجموعه داده‌ها و محیط‌های با کیفیت بالا برای طیف وسیعی از کاربردهای هوش مصنوعی و فراهم کردن امکان دسترسی مسئولانه به مجموعه داده‌های خوب و منابع آزمایش و آموزش، به تحقیقات نیاز است. بانک‌ها و جعبه‌ابزار نرم‌افزارهای منبع-باز بیشتری نیز لازم هستند تا پیشرفت تحقیق و توسعه هوش مصنوعی سرعت بگیرد. در زیربخش‌های زیر، این حوزه‌های کلیدی از اهمیت مطرح می‌شوند.

توسعه و در دسترس قرار دادن طیف گسترده‌ای از مجموعه داده‌ها برای برآورده کردن نیازهای طیف متنوعی از ذینفعان و کاربردهای هوش مصنوعی

انسجام و در دسترس بودن مجموعه داده‌های آموزش و آزمایش هوش

مصنوعی، برای حصول اطمینان از نتایج علمی معتبر حیاتی هستند. زیرساخت فنی و اجتماعی-فنی لازم برای پشتیبانی از تحقیقات قابل تولید مجدد در حوزه دیجیتال، به عنوان یک چالش مهم شناخته شده است و برای فن‌آوری‌های هوش مصنوعی نیز ضروری است. فقدان مجموعه داده‌های بررسی شده و موجود، با منشأ مشخص، برای فراهم کردن قابلیت تولید مجدد، یک عامل حیاتی برای پیشرفت مطمئن در هوش مصنوعی است^۱. همانند سایر رشته‌های علوم داده بنیان، ثبت منشأ داده‌ها حیاتی است. محققان باید بتوانند نتایج را با مجموعه داده‌های مشابه و متفاوت بازتولید کنند. مجموعه داده‌ها باید نماینده کاربردهای چالشی دنیای واقعی باشند، نه اینکه فقط نسخه‌های ساده شده باشند. برای سریع‌تر کردن پیشرفت، باید بر قابل دسترس کردن مجموعه داده‌های موجود دولت تأکید شود، آن‌هایی که می‌توانند با بودجه فدرال توسعه یابند و تا حد امکان، آن‌هایی که به وسیله صنعت نگهداری می‌شوند.

جنبه یادگیری ماشین چالش هوش مصنوعی، اغلب به تحلیل «کلان داده» مربوط است. با در نظر گرفتن تنوع گسترده مجموعه داده‌های مرتبط، نمایندگی مناسب، دسترسی و تحلیل داده‌های ساختار یافته یا نیمه ساختار یافته همچنان یک چالش در حال رشد باقی می‌ماند. چگونه می‌توان داده‌ها را با اصطلاحات مطلق و رابطه‌ای (وابسته به سیاق) نشان داد؟ پایگاه‌های داده کنونی دنیای واقعی می‌توانند به شدت مستعد داده‌های متناقض، ناقص و نویزی باشند. بنابراین، برای ایجاد مجموعه داده‌های مفید برای کاربردهای هوش مصنوعی، تعدادی از فنون پیش پردازش اطلاعات (مثل تمیز کردن

1. در این راستا، در سال 2016، فعالیت پروژه‌های تحقیقاتی هوش پیشرفته، درخواستی را برای اطلاعات در باره مجموعه داده‌ها و محیط‌های آموزشی بدیع برای پیشبرد هوش مصنوعی را منتشر کرد. نگاه کنید به <https://iarpa.gov/index.php/requests-for-information/novel-training-datasets-and-environments-working-with-iarpa> /to-advance-artificial-intelligence

داده‌ها، یکپارچه سازی داده‌ها، تبدیل داده‌ها، کاهش و بازنمایی) حائز اهمیت هستند. چگونه داده‌های پیش‌پردازش بر کیفیت داده‌ها تأثیر می‌گذارند، به خصوص زمانی که تحلیل اضافی انجام می‌شود؟

به احتمال زیاد تشویق به اشتراک گذاری مجموعه داده‌های هوش مصنوعی - به خصوص برای پژوهش‌هایی با بودجه دولتی - محرک رویکردها و راه‌حل‌های نوآورانه هوش مصنوعی هستند. با این وصف، برای اطمینان از به اشتراک گذاری ایمن داده‌ها، به فن‌آوری‌هایی نیاز داریم، زیرا صاحبان داده‌ها هنگام به اشتراک گذاری داده‌های خود با جامعه تحقیقاتی، متحمل ریسک می‌شوند. توسعه و به اشتراک گذاری مجموعه داده نیز باید مطابق قوانین و مقررات حاکم و به شیوه‌ای اخلاقی انجام شوند. ریسک‌ها می‌توانند به شیوه‌های مختلفی به وقوع بپیوندند: استفاده نامناسب از مجموعه داده‌ها، افشا نادرست یا نامناسب و محدودیت در فنون شناسایی زدایی داده‌ها برای تضمین حفاظت از حریم خصوصی و محرمانگی.

پاسخگو کردن منابع آموزش و آزمایش به ذینفعان تجاری و عمومی

با ادامه انفجار داده‌ها، منابع داده و فن‌آوری اطلاعات در سراسر جهان، هم تعداد و هم اندازه مجموعه‌های داده در حال افزایش هستند. فنون و فن‌آوری‌های تحلیل داده‌ها با حجم بالای منابع اطلاعات خام همگام نیستند. ثبت، گلچین کردن، تحلیل و مصورسازی داده‌ها، همگی چالش‌های اصلی تحقیقات هستند و دانش موردنیاز برای استخراج دانش ارزشمند از انبوهی از داده‌ها دچار عقب ماندگی است.

در حالی که مخازن داده وجود دارند، اما اغلب قادر نیستند با رشد مجموعه داده‌ها کنار بیایند، اطلاعات محدودی در باره منشأ داده‌ها دارند و از جستجوهای که از نظر معنایی، غنی از داده هستند، پشتیبانی نمی‌کنند. لذا به مخازن پویا و چابک نیاز داریم.

نمونه‌ای از نوع برنامه زیرساخت باز/مشترکی که برای پشتیبانی از نیازهای تحقیق هوش مصنوعی موردنیاز است، برنامه^۱ IMPACT (بازار خط‌مشی گذاری و تحلیل ریسک سایبری و اعتماد) است که توسط واحد امنیت داخلی (DHS^۲) توسعه یافته است.^۳ این برنامه با هماهنگی و توسعه قابلیت‌های به اشتراک گذاری داده و اطلاعات دنیای واقعی، از جمله ابزارها، مدل‌ها و روش‌شناسی‌ها، از تلاش‌های تحقیقاتی جهانی ریسک امنیت سایبری پشتیبانی می‌کند. IMPACT نیز از به اشتراک گذاری داده‌های تجربی بین جامعه بین‌المللی تحقیق و توسعه امنیت سایبری، ارائه دهندگان کنندگان زیرساخت حیاتی و حامیان دولت آن‌ها پشتیبانی می‌کند. تحقیق و توسعه هوش مصنوعی از برنامه‌های قابل مقایسه در تمامی کاربردهای هوش مصنوعی بهره‌مند خواهد شد.

توسعه بانک‌ها و جعبه‌ابزار نرم‌افزارهای منبع باز

دسترسی روزافزون به بانک‌ها و جعبه‌ابزارهای نرم‌افزارهای منبع باز، امکان دسترسی به فناوری‌های پیشرفته هوش مصنوعی را برای هر توسعه دهنده نرم‌افزاری متصل به اینترنت را فراهم می‌کند. منابعی مانند جعبه‌ابزار MALLE^۴، Weka^۵، و OpenNLP^۶، در میان بسیاری دیگر،

1. Information Marketplace for Policy and Analysis of Cyber-risk & Trust (DHS)
 2. Department of Homeland Security
 3. <https://www.dhs.gov/csd-impact>
 4. <https://sourceforge.net/projects/weka/>
 5. <http://mallet.cs.umass.edu>
 6. <https://opennlp.apache.org>

توسعه و کاربرد هوش مصنوعی را تسریع کرده‌اند. ابزارهای توسعه، از جمله مخزن کد ارزان قیمت و سیستم‌های کنترل ورژن^۱ و همچنین زبان‌های توسعه رایگان یا ارزان قیمت (مثلاً، R، Octave، و پایتون) موانع کمی را برای استفاده و گسترش این بانک‌ها فراهم می‌کنند. علاوه بر این، برای کسانی که مایل نیستند به طور مستقیم این کتابخانه‌ها را یکپارچه کنند، هر تعداد از خدمات یادگیری ماشینی مبتنی بر فناوری ابری وجود دارند که می‌توانند از طریق پروتکل‌های وب کم تأخیر که برای استفاده، نیاز به برنامه‌ریزی کمی دارند یا اصلاً نیازی به برنامه‌ریزی ندارند، وظایفی مانند طبقه‌بندی تصویر را مطابق تقاضا انجام دهند. در نهایت، بسیاری از این خدمات وب استفاده از سخت‌افزار تخصصی، از جمله سیستم‌های مبتنی بر GPU را نیز ارائه می‌دهند. منطقی است فرض کنیم که سخت‌افزار تخصصی برای الگوریتم‌های هوش مصنوعی، از جمله پردازنده‌های نورومورفیک، به طور گسترده از طریق این سرویس‌ها در دسترس قرار خواهند گرفت.

روی هم رفته، این منابع، زیرساختی از فن‌آوری هوش مصنوعی را فراهم می‌کنند که نوآوری در بازار را تشویق کرده و به کارآفرینان امکان می‌دهد تا راه‌حلهایی را توسعه دهند که مسائلی با دامنه محدود را بدون نیاز به سخت‌افزار یا نرم‌افزار گران‌قیمت، بدون نیاز به سطح بالایی از تخصص هوش مصنوعی حل کنند و اجازه ارتقا و رشد سیستم‌ها را مطابق تقاضا فراهم می‌کند. برای حوزه‌های محدود هوش مصنوعی، موانع بر سر راه نوآوری بازار نسبت به بسیاری از سایر حوزه‌های فن‌آوری، بسیار پایین هستند.

برای کمک به پشتیبانی از سطح بالای مداومی از نوآوری در این حوزه، دولت آمریکا می‌تواند تلاش‌ها برای توسعه، حمایت و استفاده از فن‌آوری‌های باز هوش مصنوعی^۱ را افزایش دهد. به‌ویژه، منابع باز که برای نمایش اطلاعات معنایی، از جمله هستی‌شناسی دامنه (در صورت موجود بودنش) از فرمت‌های استاندارد یا باز و استانداردهای باز استفاده می‌کنند، مفید فایده خواهند بود.

در ضمن، ممکن است دولت با تسریع استفاده از فن‌آوری‌های باز هوش مصنوعی درون خود دولت، موجب تشویق استفاده بیشتر از منابع هوش مصنوعی شود و در نتیجه، به پایین نگه داشتن سطح موانع برای ورود نوآوران کمک کند. هر زمان که میسر باشد، دولت باید به پروژه‌های متن‌باز، الگوریتم‌ها و نرم‌افزارهایی اهدا کند. از آنجا که دولت نگرانی‌های خاص خود را دارد، از جمله تأکید بیشتر بر حریم خصوصی و امنیت داده‌ها، ممکن است لازم باشد تا سازوکارهایی را توسعه دهد تا استفاده دولت از سیستم‌های هوش مصنوعی را تسهیل کند. برای مثال، ممکن است تأسیس یک گروه مطالعاتی مفید فایده باشد که بتواند برای یافتن حوزه‌های کاربردی خاص هوش مصنوعی، یک «جستجوی حوزه^۲» را در کل ادارات دولتی انجام دهد و سپس نگرانی‌های خاصی را مشخص کند که لازم است برطرف شوند تا امکان استفاده از چنین فنونی در این ادارات فراهم شود.

1. open AI technologies
2. horizon scan

بخش هشتم

راهبرد اندازه‌گیری و ارزیابی فن‌آوری‌های
هوش مصنوعی از طریق استانداردها و ملاحظات
نسخه ۲۰۱۹



راهبرد اندازه‌گیری و ارزیابی فن‌آوری‌های هوش مصنوعی از طریق استانداردها و ملاک‌ها: نسخه ۲۰۱۹

حمایت از توسعه استانداردهای فنی هوش مصنوعی و ابزارهای مربوطه

طرح راهبردی ملی تحقق و توسعه هوش مصنوعی اعلام می‌کند که «استانداردها، ملاک‌ها، بسترهای آزمایش و اتخاذ آن‌ها از سوی جامعه هوش مصنوعی برای هدایت و ترویج تحقیق و توسعه فناوری‌های هوش مصنوعی اساسی ضروری هستند.» در خلال این سه سال، تأکید بر استانداردها و ملاک‌ها همچنان در ایالات متحده و در سطح جهانی افزایش یافته است. فرمان ریاست جمهوری در باره حفظ رهبری آمریکا در هوش مصنوعی به صراحت بر اهمیت چنین استانداردهایی تأکید می‌کند: (مرجع ۱)

... وزیر بازرگانی، از طریق رئیس [NIST]، طرحی را برای مشارکت فدرال در توسعه استانداردهای فنی و ابزارهای مربوطه در جهت حمایت از سیستم‌های قابل اطمینان، مقاوم و قابل اعتمادی راه‌اندازی خواهد کرد که از فن‌آوری‌های هوش مصنوعی استفاده می‌کند.

با توجه به تأثیر بالقوه نوآوری‌های هوش مصنوعی بر همه بخش‌ها و حوزه‌های اجتماعی، بسیاری از سازمان‌های توسعه استاندارد

ملاحظات و اقدام کاری جدید مرتبط با هوش مصنوعی دارند که تحت بررسی هستند، از جمله این موارد می‌توان به فعالیت‌های مربوط به اخلاق هوش مصنوعی و قابلیت اعتماد سیستم‌های هوش مصنوعی اشاره کرد (مراجعه کنید به راهبرد ۳). سازمان بین‌المللی استاندارد گذاری (ISO^۱) و کمیسیون بین‌المللی الکتروتکنیک (IEC^۲), یک کمیته فرعی فنی مشترک هوش مصنوعی (ISO / IEC Joint) کمیته فنی مشترک ISO/IEC, کمیته فرعی ۴۲ هوش مصنوعی^۳) تشکیل داده‌اند تا استانداردهایی را برای سیستم‌های هوش مصنوعی و ملاحظات مربوطه توسعه دهند. این امر حیاتی است که پژوهشگران فدرال، صنعت و پژوهشگران دانشگاهی به اطلاع‌رسانی در باره این فعالیت‌ها ادامه دهند، به ویژه هم‌زمان با پیشرفت‌های هوش مصنوعی و سیستم‌هایی که به حوزه‌هایی مانند حمل‌ونقل، مراقبت‌های بهداشتی و غذا گسترش می‌یابند که با مأموریت‌های ادارات دولتی هم‌راستا است.

از سال ۲۰۱۶، جهش در فعالیت‌های مرتبط با استانداردهای هوش مصنوعی از راه‌اندازی ملاک‌ها و ارزیابی‌های متمرکز بر هوش مصنوعی، به ویژه موارد مربوط به قابلیت اعتماد سیستم‌های هوش مصنوعی پیشی گرفته‌اند. با این وصف، در خلال این مدت، با توجه به اینکه محققان در حال بررسی مجموعه داده‌های جدید تشخیص چهره هستند که به دنبال به حداقل رساندن جانب‌داری است، ملاحظات انصاف و جانب‌داری در مجموعه داده‌های معیار به طور فزاینده‌ای حائز اهمیت شده‌اند. موارد بسیار فراوان‌تر عبارت‌اند از معیارهایی که عملکرد سطح-کاربرد الگوریتم‌های هوش مصنوعی

1. International Organization for Standardization
2. International Electrotechnical Commission
3. <https://www.iso.org/committee/6794475.html>

(مثلاً، نرخ‌های مثبت-کاذب یا منفی-کاذب الگوریتم‌های طبقه‌بندی) را آزمون می‌کنند و همچنین ملاک‌هایی که عملکرد سطح-محاسباتی سیستم‌های نرم‌افزاری و سخت‌افزاری هوش مصنوعی را تعیین می‌کنند. دو مورد اخیر از چنین فعالیت‌هایی MLPerf^۱ و DAWNbench^۲ هستند.

ارزیابی، ترویج و اطمینان از همه جوانب قابل اعتماد بودن هوش مصنوعی، نیازمند اندازه‌گیری و ارزیابی عملکرد فن‌آوری هوش مصنوعی از طریق ملاک‌ها و استانداردها است. هوش مصنوعی و رای قابلیت اعتماد، قابلیت اطمینان، انعطاف‌پذیری، قابلیت تبیین و شفافیت، باید در هنگام تشخیص، حریم خصوصی را حفظ کرده و از جانب‌داری نامناسب اجتناب کند. با تکامل فن‌آوری‌های هوش مصنوعی، نیاز به توسعه معیارهای جدید و آزمودن الزامات برای اعتبار سنجی این مشخصات، ضروری خواهد بود.

استانداردها، ملاک‌ها و ابزارهای مرتبط: برنامه‌های اخیر مؤسسات تحقیق و توسعه

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، NIST به طور خاص، تلاش‌های خود برای تحقق راهبرد ۶ را آغاز کرده است:

- NIST در برنامه استاندارد گذاری ISO/IEC JTC 1 SC 42 بر روی هوش مصنوعی مشارکت کرده است^۳. یک متخصص NIST، مسئول گردآوری وظیفه کاری «کلان داده» در SC 42 است. هیئت

نمایندگی آمریکا در SC 42 شامل NIST و دیگر متخصصان موسسه فدرال و نمایندگان از صنعت و دانشگاهیان است. ورود آمریکا به SC 42، توسط کمیته بین‌المللی استانداردهای فن‌آوری اطلاعات (INCITS) تسهیل شده است.

- کارکنان NIST از طریق سازمان‌های استانداردها مانند انجمن آمریکایی مهندسان مکانیک^۱، IEEE، و ISO/IEC در فعالیتهای دیگر مربوط به استانداردهای هوش مصنوعی شرکت دارند. فعالیتهای آنها، موضوعاتی همچون مدل‌سازی محاسباتی برای تولید پیشرفته، هستی‌شناسی رباتیک و اتوماسیون، حریم خصوصی داده‌های شخصی و سوگیری الگوریتمی را پوشش می‌دهند.

- متخصصان NIST در حال افزایش آگاهی‌ها در مورد اهمیت استانداردهای مورد اتفاق برای هوش مصنوعی در مجامع چندجانبه، از جمله نهادهایی مانند G20 و G7 هستند^۲. NIST تخصص منحصر به فرد دولت فدرال را به همراه دارد که پایه‌گذار بحث‌های سیاست‌گذاری عملی است، به ویژه از طریق همکاری نزدیک با بخش خصوصی. بعلاوه، NIST استانداردهای خود و تجربیات مرتبط با آنها را در اختیار بحث‌های دوجانبه میان دولت‌ها قرار می‌دهد.

استانداردها، ملاک‌ها، بسترهای آزمایش و اتخاذ آنها از سوی جامعه هوش مصنوعی برای هدایت و ترویج تحقیق و توسعه فن‌آوری هوش مصنوعی ضروری هستند. در زیر بخش‌های بعدی، زمینه‌هایی را ترسیم می‌کنیم که در آن پیشرفت‌های بیشتری باید صورت گیرد.

1. Institute of Electrical and Electronics Engineers

2. <https://home.treasury.gov/policy-issues/international/g-7-and-g20->

توسعه طیف وسیعی از استانداردهای هوش مصنوعی

توسعه استانداردها باید به گونه‌ای باشد که با تکامل توانایی‌ها و گسترش حوزه‌های کاربردی هوش مصنوعی، همگام باشد. استانداردها، تأمین‌کننده نیازمندی‌ها، مشخصات، رهنمودها یا خصوصیات هستند که می‌توانند به طور پیوسته مورد استفاده قرار گیرند تا اطمینان حاصل شود که فن‌آوری‌های هوش مصنوعی با اهداف مهم عملکرد و قابلیت همکاری مطابقت دارند و به شکلی قابل اطمینان و ایمن عمل می‌کنند. پذیرش استانداردها، به پیشرفت‌های فناوری اعتبار می‌بخشد و بازاری را که قابلیت همکاری تسهیل کرده و گسترش می‌دهد. نمونه‌ای از استاندارد مربوط به هوش مصنوعی که از سوی انستیتوی مهندسی برق و الکترونیک توسعه یافته، P1872-2015 (استاندارد هستی‌شناسی رباتیک و اتوماسیون) است. این استاندارد، یک روش منظم برای ارائه دانش و یک مجموعه مشترک از اصطلاحات و تعاریف است. این موارد، امکان انتقال دانش بدون ابهام بین انسان‌ها، روبات‌ها و دیگر سیستم‌های مصنوعی را فراهم می‌کنند و برای کاربرد فن‌آوری‌های هوش مصنوعی در رباتیک، مبنایی اساسی را ارائه می‌کند. در تمامی حوزه‌های فرعی هوش مصنوعی به کار بیشتری در توسعه استانداردهای هوش مصنوعی نیاز داریم.

استانداردها برای رفع موارد زیر لازم هستند:

مهندسی نرم‌افزار: مدیریت کردن پیچیدگی، پایداری، امنیت سیستم و نظارت و کنترل رفتارهای نوظهور؛
عملکرد: تضمین صحت، قابلیت اطمینان، مقاوم بودن، قابلیت

دسترسی و توسعه پذیری؛
 معیارها: کمی کردن عوامل مؤثر بر عملکرد و انطباق با استانداردها؛
 ایمنی: ارزیابی مدیریت ریسک و تحلیل خطر سیستم‌ها، تعاملات
 انسان-رایانه، سیستم‌های کنترل و پیروی از مقررات؛
 قابلیت استفاده: تضمین این که واسطه‌ها و کنترل‌ها مؤثر،
 کارآمد و شهودی هستند؛
 قابلیت همکاری: تعریف مؤلفه‌ها، داده‌ها و مدل‌های تراکنش قابل
 جابجایی از طریق واسطه‌های استاندارد و سازگار؛
 امنیت: برای رسیدگی به محرمانگی، یکپارچگی و در دسترس بودن
 اطلاعات و همچنین امنیت سایبری؛
 حریم خصوصی: برای کنترل حفاظت از اطلاعات در هنگام پردازش،
 هنگام انتقال، یا ذخیره‌سازی؛
 قابلیت ردیابی: برای فراهم کردن پرونده‌ای از رخدادها (پیاده‌سازی،
 آزمایش و تکمیل آن‌ها) و برای گردآوری داده‌ها؛ و
 دامنه‌ها: برای تعریف فرهنگ لغت استاندارد خاص-دامنه و
 چارچوب‌های مربوطه.

ایجاد ملاک‌های فناوری هوش مصنوعی

ملاک‌ها، متشکل از آزمایش‌ها و ارزیابی‌ها هستند، برای توسعه
 استانداردها و ارزیابی انطباق با استانداردها، سنجه‌های کمی فراهم
 می‌کنند. ملاک‌ها با ارتقای پیشرفت‌هایی که هدفشان پرداختن به
 سناریوهای منتخب راهبردی است، پیشران نوآوری هستند. علاوه بر این،

ملاک‌ها برای ردیابی تکامل علم و فن‌آوری هوش مصنوعی داده‌های عینی را ارائه می‌کنند. برای ارزیابی مؤثر فن‌آوری‌های هوش مصنوعی، باید روش‌شناسی‌ها و معیارهای مناسب و مؤثر آزمایش، توسعه یابند و استاندارد شوند. روش‌های آزمایش استاندارد، برای ارزیابی، مقایسه و مدیریت عملکرد فن‌آوری‌های هوش مصنوعی، پروتکل‌ها و روال‌هایی را تجویز خواهند کرد. برای تعریف سنجه‌های قابل‌اندازه‌گیری برای مشخص کردن فناوری‌های هوش مصنوعی، به معیارهای استاندارد نیاز داریم که شامل موارد زیر هستند (ولی محدود به آن‌ها نیستند): دقت، پیچیدگی، اعتماد و شایستگی، ریسک و عدم اطمینان، قابلیت تبیین، جانب‌داری ناخواسته، مقایسه با عملکرد انسانی و تأثیر اقتصادی. شایان ذکر است که ملاک‌ها داده محور هستند. راهبرد ۵ در باره اهمیت مجموعه داده‌هایی برای آموزش و آزمایش بحث می‌کند.

به عنوان نمونه‌ای موفق از معیارهای مرتبط با هوش مصنوعی، موسسه ملی استانداردها و فناوری برای ارزیابی قابلیت‌های کلیدی ربات‌های واکنش اضطراری، مجموعه جامعی از روش‌های آزمون استاندارد و معیارهای عملکردی مرتبط را توسعه داده است. هدف از انجام این کار، تسهیل در مقایسه کمی مدل‌های مختلف ربات با استفاده از داده‌های آماری معنی‌دار بر روی قابلیت‌های ربات بود که با استفاده از روش‌های آزمون استاندارد به دست آمد. این مقایسه‌ها می‌توانند تصمیمات خرید را هدایت کنند و به توسعه‌دهندگان کمک کنند تا قابلیت‌های استقرار را درک کنند. روش‌های آزمون حاصل، از طریق کمیته استانداردهای بین‌المللی کاربردهای وطنی امنیت برای

تجهیزات عملیاتی رباتیک (استاندارد ۱، ۰۸، ۰۸، E54) استاندارد سازی می‌شوند.^۱ برای به چالش کشیدن جامعه تحقیق از طریق رقابت‌های موسوم به مسابقات لیگ ربات‌های نجات (RoboCup Rescue Robot)^۲، نسخه‌های روش‌های آزمون مورد استفاده قرار می‌گیرند که تأکیدی است بر قابلیت‌های خودکار. مثال دیگر، رباتیک چابک IEEE برای رقابت اتوماسیون صنعتی (ARIAC)^۳، تلاش مشترک بین IEEE و NIST^۴ است که با استفاده از آخرین پیشرفت‌های هوش مصنوعی و طرح‌ریزی ربات، چابکی ربات را افزایش می‌دهد. هدف اصلی این رقابت، آزمون چابکی سیستم‌های ربات صنعتی است تا کسانی را که در کف فروشگاه‌ها کار می‌کنند، مولدتر و مستقل‌تر کنند و در وقت آن‌ها صرفه‌جویی به وجود آورند.

در حالی که این تلاش‌ها، بنیان قوی برای به پیش بردن ملاک سازی برای هوش مصنوعی فراهم می‌کنند، محدودیت آن‌ها مختص بودن به حوزه است. استانداردها، بسترهای آزمایش و ملاک‌های بیشتری در دامنه وسیعی از حوزه‌ها موردنیاز هستند تا اطمینان حاصل شود که راه‌حل‌های هوش مصنوعی کاربرد گسترده‌ای داشته باشند و به طور عام مورد پذیرش قرار گیرند.

افزایش دسترسی به بسترهای آزمایش هوش مصنوعی

اهمیت بسترهای آزمایش در گزارش «آینده آزمایش سایبری» بیان شد: «بسترهای آزمایش ضروری هستند تا محققان بتوانند از داده‌های عملیاتی واقعی برای مدل‌سازی و اجرای آزمایش‌ها بر روی

1. نسخه 2019: هم‌اکنون روش‌های آزمون‌های حاصل، استانداردهایی هستند که توسط کمیته ASTM در زمینه کاربردهای امن خانگی ربات‌های Response منتشر می‌شوند (E54.09).

2. <http://www.robocup2016.org/en/>

3. <http://robotagility.wixsite.com/competition>

4. نسخه 2019: IEEE دیگر شریک ARIAC نیست که هم‌اکنون در سال سوم تأسیس خود قرار دارد.

سیستم‌های دنیای واقعی استفاده کنند ... و سناریوهایی در محیط‌های آزمایش خوب.^۱ همه حوزه‌های هوش مصنوعی، داشتن بسترهای آزمایش کافی یک ضرورت است. دولت مقادیر زیادی داده‌های حساس-ماموریت دارد که منحصر به خود دولت است، اما بخش اعظمی از این داده‌ها را نمی‌توان بین جامعه تحقیقاتی خارجی توزیع کرد. می‌توان برنامه‌های مناسبی را برای محققان دانشگاهی و صنعتی ایجاد کرد تا تحقیقات را درون محیط‌های امن و گلچین شده بستر آزمایش هدایت کنند که توسط مؤسسات خاصی پایه‌ریزی شده‌اند. جوامع تحقیقاتی می‌توانند مدل‌های هوش مصنوعی و روش‌های آزمایشگاهی را به اشتراک گذاشته و اعتبارسنجی کنند. این کار را می‌توان با دسترسی آنان به محیط‌های آزمون، تأمین فرصت‌های تحقیقاتی منحصر به فرد (که در غیر این صورت در دسترس نیستند) برای دانشمندان، مهندسیین و دانشجویان انجام داد.

مشارکت جامعه هوش مصنوعی در استانداردها و معیارها (الگوها)

برای هدایت استاندارد و تشویق کاربرد گسترده آن در دولت، دانشگاه و صنعت به رهبری و هماهنگی دولت نیاز است. جامعه (اجتماع) هوش مصنوعی متشکل از کاربران، صنعت، دانشگاهیان و دولت، باید انرژی بگیرند تا در توسعه استانداردها و برنامه‌های معیارسنجی مشارکت کنند. همان‌طور که هر نهاد دولتی، بر اساس نقش و مأموریت خود به شیوه‌های مختلفی در این اجتماع مشارکت

می‌کنند، تعاملات این اجتماع را می‌توان از طریق هماهنگی ارتقا داد تا تأثیرشان تقویت شود. برای جمع‌آوری دسته‌جمعی الزامات کاربر-محور، پیش‌بینی استانداردهای توسعه دهنده-محور و ارتقا فرصت‌های آموزشی به این هماهنگی نیاز داریم. الزامات کاربر-محور اهداف و طراحی مسائل چالشی را شکل داده و ارزیابی فن‌آوری را ممکن می‌سازند. وجود معیارها در این اجتماع، موجب تمرکز تحقیق و توسعه برای تعریف پیشرفت، شکاف‌های باریک شده و پیشران راه‌حل‌های نوآورانه برای مسائل خاص است. این معیارها باید شامل روش‌هایی برای تعریف و تعیین حقیقت زیربنایی باشند. ایجاد ابزارهای شبیه‌سازی و تحلیل معیار نیز موجب تسریع در پیشرفت‌های هوش مصنوعی خواهد شد. نتایج این معیارها به تطبیق با فن‌آوری مناسب برای نیاز کاربر، تشکیل معیارهای هدف برای انطباق استانداردها، فهرست‌های محصول واجد شرایط و انتخاب منبع بالقوه کمک می‌کنند.

صنعت و دانشگاه منابع اصلی برای فناوری‌های نوظهور هوش مصنوعی هستند. ترویج و هماهنگی مشارکت آن‌ها در استانداردها و فعالیت‌های الگوبرداری بسیار مهم است. همان‌طور که راه‌حل‌ها ظاهر می‌شوند، به وفور فرصت‌هایی به وجود می‌آیند. فرصت‌های مذکور در موارد زیر ایجاد می‌شوند: پیش‌بینی استانداردهای توسعه دهنده-محور و کاربر-محور از طریق به اشتراک‌گذاری چشم‌اندازهای مربوط به معماری‌های فنی، توسعه مرجع پیاده‌سازی استانداردهای نوظهور برای نشان دادن امکان‌سنجی و انجام آزمایش پیش-رقابتی برای تضمین راه‌حل‌های مناسب و قابل تعامل و همچنین توسعه

بهترین تجارب برای کاربردهای فن آوری.

یک نمونه موفق از یک برنامه الگوبرداری با تأثیر بالا، متکی به جامعه و مرتبط با هوش مصنوعی، کنفرانس بازیابی متن (TREC)^۱ است که در سال ۱۹۹۲ توسط NIST آغاز شد تا زیرساخت لازم برای ارزیابی مقیاس بزرگ روش‌شناسی‌های بازیابی اطلاعات را فراهم کند. بیش از ۲۵۰ گروه، از جمله سازمان‌های دانشگاهی و تجاری بزرگ و کوچک در TREC شرکت کرده‌اند. مجموعه داده استاندارد، به طور گسترده در دسترس و به دقت ساخته شده که از سوی TREC ارائه شده است، به وسیله تقویت تحقیقات بر روی بازیابی اطلاعات مورد تأیید قرار گرفته است.^۲ دومین مثال، برنامه الگوبرداری متناوب NIST در حوزه بینایی ماشین است که در بیومتری، به ویژه تشخیص چهره استفاده می‌شود.^۳ این کار با ارزیابی فن آوری تشخیص چهره (FERET) در سال ۱۹۹۳ آغاز شد که مجموعه داده استاندارد از عکس‌های چهره طراحی شده برای پشتیبانی از توسعه الگوریتم تشخیص چهره و پروتکل ارزیابی را ارائه کرد. این تلاش طی سال‌ها در FRVT^۴ (آزمایش فروشنده تشخیص چهره)،^۵ شامل توزیع مجموعه داده‌ها، میزبانی مسائلی چالشی و انجام ارزیابی‌های منزوی فن آوری تکامل یافته‌اند. این برنامه الگوبرداری تا حدود زیادی به بهبود فن آوری تشخیص چهره کمک کرده است. هم TREC و هم FRVT می‌توانند به عنوان نمونه‌هایی از فعالیت‌های الگوبرداری مؤثر جامعه مرتبط با هوش مصنوعی عمل کنند، اما در سایر حوزه‌های هوش مصنوعی نیز به تلاش‌های مشابهی نیاز داریم.

1. Text Retrieval Conference

2. <http://trec.nist.gov>

۳. ای. ام. وورهیز و دی. کی. هارمن، آزمایش و ارزیابی TREC در بازیابی اطلاعات (Cambridge: MIT Press, ۲۰۰۵).
<http://googleblog.blogspot.com/03/2008/why-data-matters.html>

4. <http://biometrics.nist.gov>

5. P. J. Phillips, "Improving Face Recognition Technology," Computer 96-84: (2011)(3)44.

6. Face Recognition Vendor Test

لازم به ذکر است که توسعه و اتخاذ استانداردها و مشارکت در فعالیتهای الگوبرداری، با هزینه همراه است. هنگامی که سازمانهای تحقیق و توسعه منفعت قابل توجهی را مشاهده می‌کنند، انگیزه می‌گیرند. به روز رسانی فرایندهای تملک در تمامی مؤسسات به منظور گنجانیدن الزامات خاص استانداردهای هوش مصنوعی در «درخواست‌ها برای پیشنهادها»، جامعه را به مشارکت بیشتر در توسعه و پذیرش استانداردها تشویق خواهد کرد. علاوه بر این الگوهای مبتنی بر جامعه، مانند TREC و FRVT، به کمک موارد زیر، موانع را کم کرده و مشوق‌ها را تقویت می‌کنند: ارائه انواع آموزش و آزمایش داده‌هایی که صرفاً در این مثال‌ها در دسترس قرار هستند، ترغیب به رقابت سالم بین توسعه‌دهندگان فن‌آوری برای پیشبرد بهترین الگوریتم‌های موجود و ارائه هدف و معیارهای مقایسه‌ای عملکرد برای انتخاب منابع مرتبط.

بخش هفتم

راهبرد ۷: درک بهتر نیازهای نیروی
کارمله تحقیق و توسعه هوش مصنوعی
نسخه ۲۰۱۹



راهبرد ۷: درک بهترین نیازهای نیروی کار مله تحقیق و توسعه هوش مصنوعی: نسخه ۲۰۱۹

به جلو راندن نیروی کار تحقیق و توسعه هوش مصنوعی، شامل کسانی که بر روی سیستم‌های هوش مصنوعی کار می‌کنند و کسانی که در کنار آن‌ها کار می‌کنند تا رهبری آمریکا را حفظ کنند

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، تقاضا برای پژوهشگران و عمل‌گرایان هوش مصنوعی به سرعت افزایش یافته است. مطالعات نشان داده‌اند که انتظار می‌رود در دهه آینده تعداد فرصت‌های شغلی به میلیون‌ها افزایش یابد. به عنوان یک نقطه از داده‌ها، دفتر آمار نیروی کار آمریکا نشان می‌دهد که از سال ۲۰۱۶ تا ۲۰۲۶، تعداد موقعیت‌های شغلی برای دانشمندان و مهندسان کامپیوتر و اطلاعات تا ۱۹٪ افزایش خواهد یافت. این رقم تقریباً سه برابر سریع‌تر از میانگین تمامی مشاغل است.^۱ علاوه بر این، تا سال ۲۰۲۸، انتظار می‌رود که پژوهشگران هوش مصنوعی بالغ بر ۱۱،۵ هزار میلیارد دلار به رشد کلی کمک کنند، رقمی که صرفاً توسط فناوری‌های هوشمند در کشورهای G20 وعده داده شده است.^۲

1. <https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm>

2. https://www.accenture.com/t20180920T094705Z_w_/us-en/_acnmedia/Thought-Leadership-Assets/PDF/Accenture-Education-and-Technology-Skills-Research.pdf

مؤسسات دانشگاهی آمریکایی در تلاش هستند تا با رشد انفجاری علاقه و ثبت‌نام دانشجویان در کلاس‌های هوش مصنوعی همگام شوند.^۱ در عین حال، صنعت با حمایت مالی پایدار و دسترسی به امکانات و مجموعه داده‌های محاسباتی پیشرفته، جذابیت زیادی را برای تحقیقات دانشگاهی و استعداد تدریس به وجود می‌آورد.^۲ حفظ یک زیست‌بوم تحقیقاتی دانشگاهی قوی هوش مصنوعی حائز اهمیت است که با همکاری تحقیق و توسعه صنعت، می‌تواند با پیشبرد بهداشت، رونق و رفاه ملی و تأمین امنیت ملی، به ارائه منافع عظیم^۳ ادامه دهد.

سه سال پس از انتشار طرح راهبردی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، گزارش‌های مختلف، خواستار پشتیبانی مداوم از توسعه مطالب آموزشی و توسعه حرفه‌ای معلمان علوم کامپیوتر در همه سطوح شده‌اند. لازم است بر سطوح K-12 تأکید شود تا در طول چندین دهه، کانال و مسیر کشوری پژوهشگران هوش مصنوعی هموار شود.^۴ با توجه به نقش فزاینده محاسبات در رشته‌های مختلف، لازم است تا در سطح کارشناسی، بر یکپارچه‌سازی مهارت‌ها و روش‌های محاسباتی پیشرفته با دانش خاص این حوزه تأکید شود که از رشته‌های دیگر به دست آمده‌اند.^۵ در ضمن در سطح فارغ‌التحصیلان نیز که در آن، دانشجویان در حال انجام تحقیقات بنیادین در ML و هوش مصنوعی هستند، نیاز به حمایت مداوم وجود دارد. در واقع، فرمان ۲۰۱۹ ریاست جمهوری در باره حفظ رهبری آمریکا در هوش مصنوعی، ملزم می‌دارد که: (مرجع ۱)

1. <https://cra.org/data/generation-cs/>
<https://cra.org/wp-content/uploads/-2017/05/2018Taulbee-Survey-Report.pdf>
<http://web.cs.wpi.edu/~cew/papers/CSareas19.pdf>
2. <https://www.nitrd.gov/rfi/ai/2018/AI-RFI-Response-2018-Yolanda-Gil-AAAI.pdf>
3. <https://www.nap.edu/catalog/13427/continuing-innovation-in-information-technology>
4. <https://github.com/touretzkyds/ai4k12/wiki>
5. <https://www.nap.edu/catalog/24926/assessing-and-responding-to-the-growth-of-computer-science-undergraduate-enrollments>

روسای مؤسسات اجرایی که بورس‌های تحصیلی فراهم می‌کنند، تا اندازه‌ای که با قانون حاکم مطابقت دارد، هوش مصنوعی را به عنوان یک اولویت در سازمان‌های فدرال موجود و برنامه‌های خدمت‌رسانی ... (شامل)...

مورد توجه قرار خواهند داد

الف- دبیرستان، دوره کارشناسی، تحصیلات تکمیلی، آموزش جایگزین و برنامه‌های آموزشی؛

ب- برنامه‌هایی برای شناسایی و تأمین بودجه دانشکده دانشگاهی شغل-اولیه که تحقیق و توسعه هوش مصنوعی را از جمله از طریق جوایز و قدردانی هدایت می‌کنند؛

ج- بورس تحصیلی برای برنامه‌های خدماتی؛

د- اجرای مستقیم برنامه‌های نیروهای مسلح آمریکا؛ و

ه- برنامه‌هایی که برای تسهیل تجارب یادگیری شخصی و تطبیقی از آموزش و تربیت رسمی و غیررسمی، از توسعه برنامه‌های آموزشی و برنامه‌های درسی حمایت می‌کنند که یکپارچگی فناوری‌های هوش مصنوعی را با دروس مورد تشویق قرار می‌دهند.

به طور گسترده‌تر، نیاز به یک بنیان قدرتمند در تفکر محاسباتی، از جمله از طریق آموزش علوم کامپیوتر، مؤکدا در برنامه راهبردی پنج ساله دولت فدرال (دسامبر ۲۰۱۸) برای آموزش علم، فن‌آوری، مهندسی و ریاضیات (STEM^۱) نیز مورد توجه قرار گرفته است.^۲ به‌علاوه، توسعه مشارکت میان گروه‌هایی که به طور سنتی در حوزه محاسبات و زمینه‌های مرتبط نشان داده نشده‌اند، ضروری است.

1. science, technology, engineering, and mathematics

2. <https://www.whitehouse.gov/wp-content/uploads/12/2018/STEM-Education-Strategic-Plan2018-.pdf>

در نهایت، نیروی کار تحقیق و توسعه هوش مصنوعی نه تنها متشکل از تیم‌های چند رشته‌ای دانشمندان و دانشمندان و مهندسين خواهد بود، بلکه متخصصان رشته‌های دیگر در زمینه علوم شناختی و روانشناسی، علم اقتصاد و نظریه بازی، اصول اخلاقی، زبان‌شناسی، فلسفه و بسیاری از حوزه‌های دیگر را شامل خواهد شد که هوش مصنوعی به آن‌ها اعمال می‌شود.

مؤسسات فدرال در حال اولویت دادن به برنامه‌های آموزشی و کمک هزینه تحصیلی در تمام سطوح هستند تا از طریق دوره‌های کارآموزی، برنامه‌های مهارتی، بورس تحصیلی و البته کار در رشته‌های مربوطه، نیروی کار مجهز به مهارت‌های هوش مصنوعی را آماده نمایند (مربع کناری را ببینید). چنین فرصت‌های آموزشی، هم دانشمندان و مهندسين را هدف قرار می‌دهند که به نوآوری‌های تحقیق و توسعه هوش مصنوعی کمک می‌کنند و هم به کاربران تحقیق و توسعه هوش مصنوعی توجه دارند که ممکن است دارای دانش مرتبط با این حوزه باشند. همان طور که در راهبرد ۱ توضیح داده شد، در مورد گروه اول، سرمایه‌گذاری فدرال طولانی مدت در تحقیق و توسعه هوش مصنوعی، هم از طریق آموزش نسل بعدی محققان و هم با ایجاد موقعیت‌های هیئت‌علمی برای دانشجویان دوره تحصیلات تکمیلی و دکترا از رشد این نیروی کار پشتیبانی می‌کند. در مورد گروه دوم، برنامه‌های جدید، مهارت‌های مربوط به هوش مصنوعی را به کاربران فعلی و آینده سیستم‌های هوش مصنوعی عرضه می‌کنند (مربع کناری را ببینید). بنابراین، مؤسسات فدرال باید به پرورش راهبردی تخصص در نیروی کار تحقیق و توسعه

هوش مصنوعی که رشته‌های متعدد و دسته‌های مهارتی متعددی را در بر دارند، ادامه دهند تا از رهبری پایدار ملی اطمینان حاصل کنند.

نیروی کار ملی تحقیق و توسعه هوش مصنوعی: فعالیت‌های اخیر مؤسسات

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، تعدادی از مؤسسات تلاش‌هایی را برای حمایت از راهبرد ۷ آغاز کرده‌اند:

- صرف‌نظر از حمایت از دانشجویان کارشناسی و فارغ‌التحصیلان از طریق بورس‌های استاندارد هوش مصنوعی، مؤسسات در حال اولویت‌بخشی به علم و مهندسی محاسباتی و داده محور در برنامه‌های بورس تحصیلی خود هستند. برای مثال، در سال ۲۰۱۸، DOE یک مسیر جدید را به برنامه بورس تحصیلی لیسانس علم محاسبات خود اضافه کرد. این مسیر از دانشجویانی پشتیبانی می‌کند که درجات پیشرفته در ریاضیات کاربردی، آمار یا علوم رایانه را دنبال می‌کنند و استفاده مؤثر از سیستم‌های عملکرد بالا را ارتقا می‌دهند، از جمله در حوزه‌های هوش مصنوعی، ML و یادگیری عمیق.^۱ در ضمن، در سال ۲۰۱۸ NSF شروع به اولویت دادن به علم و مهندسی محاسباتی و داده محور در یک زیرمجموعه از جوایز برنامه بورس تحصیلی پژوهش تحصیلات تکمیلی خود کرد.
- اداره سرشماری طرح نوسازی داده‌های آماری (SDM) را ایجاد کرده تا نیروی کار، عملیات و فن‌آوری‌هایش را تا وضعیت کنونی ارتقا دهد و در جامعه مبتنی بر داده امروز، استاندارد نهادهای آماری را تدوین کند.

مؤلفه تحول نیروی کار SDM، به استخدام دانشمندان جدید داده با تخصص در روش‌ها و تحلیل جدید، شامل استفاده از روش‌ها و ابزار هوش مصنوعی برای پردازش و تحلیل کلان داده کمک خواهد کرد. تحول نیروی کار، به ارتقا نیروی کار علم اطلاعات فعلی نیز خواهد پرداخت.

دستیابی به پیشرفت‌های موردنیاز در تحقیق و توسعه هوش مصنوعی که در این راهبرد درج شده، مستلزم نیروی کار کافی در تحقیق و توسعه هوش مصنوعی است. کشورهایی با قوی‌ترین حضور در تحقیق و توسعه هوش مصنوعی، موقعیت‌های رهبری را در اتوماسیون آینده به وجود خواهند آورد. آن‌ها به پیشگامان شایستگی‌ها مبدل خواهند شد، شایستگی‌هایی همچون ایجاد و توسعه الگوریتم، نمایش قابلیت و تجاری‌سازی. توسعه تخصص فنی، پایه‌ای برای این پیشرفت‌ها است.

در حالی که در حال حاضر هیچ آمار رسمی از نیروی کار وجود ندارد، گزارش‌های اخیر از بخش‌های تجاری و دانشگاهی حاکی از کمبود فزاینده متخصصان موجود در هوش مصنوعی است. گزارش‌ها حاکی از آن است که کمبود متخصصان هوش مصنوعی وجود دارد^۱ و انتظار می‌رود که این کمبود شدیدتر شود^۲. گزارش‌ها حاکی از آن است که شرکت دارای فن‌آوری پیشرفته، منابع قابل توجهی را در استخدام اعضای هیئت‌علمی و دانشجویان با مهارت هوش مصنوعی سرمایه‌گذاری می‌کنند^۳. گزارش شده که دانشگاه و صنایع در رقابتی برای استخدام و حفظ استعداد هوش مصنوعی قرار دارند^۴.

1. Startups Aim to Exploit a Deep-Learning Skills Gap," MIT Technology Review, January 2016, 6.

2. "AI talent grab sparks excitement and concern," Nature, April 2016, 26.

۳. «تقاضای بالایی برای متخصصان هوش مصنوعی وجود دارد».

The Wall Street Journal, May 2015, 1.

۴. «محبوب‌های میلیون دلاری: نبردهای سیلیکون ولی برای جذب استعداد؛ دانشگاه‌ها در تلاش‌اند تا ستاره‌های خود را نگه دارند»
," The Economist, April 2016, 2.

برای درک بهتر نیازهای نیروی کار ملی کنونی و آتی برای هوش مصنوعی، مطالعات بیشتری لازم است. برای مشخص کردن وضعیت فعلی نیروی کار تحقیق و توسعه هوش مصنوعی، شامل نیازهای دانشگاه، دولت و صنعت، به داده نیاز داریم. مطالعات باید در محل کار هوش مصنوعی، نیروهای عرضه و تقاضا را مورد بررسی قرار دهند تا به پیش‌بینی نیازهای نیروی کار آینده کمک کنند. درک مسیر پیش‌بینی شده نیروی کار تحقیق و توسعه هوش مصنوعی، ضروری است. ملاحظات مسیره‌های آموزشی و فرصت‌های آموزشی بالقوه را تیز باید با این موارد افزود. موضوعات گوناگونی نیز باید مورد بررسی قرار گیرند، زیرا مطالعات نشان داده‌اند که یک نیروی کار متنوع فن‌آوری اطلاعات می‌تواند منجر به نتایج بهتری شود^۱. زمانی که نیازهای نیروی کار فعلی و آینده تحقیق و توسعه هوش مصنوعی بهتر درک شوند، آنگاه می‌توان برنامه‌ها و اقدامات مناسبی را در نظر گرفت تا چالش‌های نیروی کار موجود یا پیش‌بینی‌شده را بر طرف کنند.

۱. جی. دالیلیو، موودیف سی. ام، بی. بی، وویینسکی و ام. ای. میبرز، «تنوع و نیروی کار فناوری اطلاعات: موانع و فرصت‌ها»، Journal of Computer Information Systems 71-63: (2003) 43.



بخش هشتم

راهبرد ۸ گسترش مشارکت
بخش های عمومی - خصوصی



راهبرد ۸: گسترش مشارکت بخش‌های عمومی - خصوصاً برای سرعت بخشیدن به پیشرفت هادروش مصنوعی

راهبرد ۸ در نسخه سال ۲۰۱۹ جدید است و منعکس‌کننده اهمیت فزاینده مشارکت بخش‌های عمومی - خصوصی است که تحقیق و توسعه هوش مصنوعی را توانمند می‌سازند.

رهبری آمریکا در تحقیقات و نوآوری علوم و مهندسی، ریشه در زیست‌بوم منحصربه‌فرد تحقیق و پژوهش کشور - دانشگاه - صنعت دارد. همان‌طور که انجمن آمریکایی هنر و علوم آمریکا نوشته است: «موقعیت آمریکا به عنوان یک رهبر نوآوری» متکی بر «ایجاد یک مشارکت تحقیقاتی ملی قوی‌تر بین دانشگاه و صنعت است»^۱.

از زمان انتشار طرح راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، دولت به تقویت این چشم‌انداز پرداخته است: ترویج « سرمایه‌گذاری مداوم در AI R&D و همکاری با دانشگاهیان، صنعت، شرکای بین‌المللی و هم‌پیمانان و دیگر نهادهای غیردولتی برای تولید پیشرفت‌های فناورانه ساختار شکن در فن‌آوری AI و فن‌آوری مربوطه و گذار سریع از آن پیشرفت‌ها به قابلیت‌هایی که به امنیت اقتصادی و ملی آمریکا کمک می‌کنند.» (مرجع ۱)

در چند دهه گذشته، تحقیقات بنیادینی که در زمینه فن‌آوری

۱. ذخیره مجدد مینا: نقش حیاتی پژوهش در حفظ روای آمریکایی (American Academy of Arts and Sciences, Cambridge, MA, 2014); https://www.amacad.org/multimedia/pdfs/publications/researchpapersmonographs/AmericanAcad_Restoring_the_Foundation_Brief.pdf.

اطلاعات با بودجه فدرال در دانشگاه‌ها و صنعت انجام شده، به بخش‌های جدیدی در اقتصاد چند میلیارد دلاری اقتصاد کشور منجر شده است.^۱ پیشرفت‌های هم‌زمان در سراسر دولت، دانشگاه‌ها و صنایع، همدیگر را تقویت کرده و منجر به یک بخش نوآورانه و فعال در زمینه هوش مصنوعی شده‌اند. بسیاری از سیستم‌های هوش مصنوعی امروزی توسط زیست‌بوم تحقیق و توسعه دولت آمریکا- دانشگاه- صنعت فعال شده‌اند (مربع کناری را ببینید).

پیشبرد زیست‌بوم نوآوری هوش مصنوعی کشور، از طریق پل زدن میان دولت، دانشگاه‌ها و صنعت

• ثابت شده که شبکه‌های عصبی کونولوشنال، نوآوری کلیدی هستند که ریشه در تحقیقات هوش مصنوعی دارند. اگرچه این رویکرد مدل‌سازی، از سرمایه‌گذاری‌های اولیه فدرال در اواخر دهه ۱۹۸۰ سرچشمه می‌گیرد، اما در آن زمان، برای پیش‌بینی‌های دقیق توسط شبکه‌های عصبی نه داده‌های کافی وجود داشت و نه قابلیت‌های محاسباتی. امروزه با ترکیبی از موارد زیر، شبکه‌های عصبی به عنوان یک روش مفید برای بهبود دقت مدل‌های هوش مصنوعی پدیدار شده‌اند: جهش در کلان داده، روش‌های علمی داده‌بنیان و پیشرفت‌های مفهومی در نحوه ساختاردهی و بهینه‌سازی شبکه‌ها. در سال‌های اخیر تعاملات بین دانشگاه‌ها و بخش خصوصی، از جمله تأمین مالی دولت، به کاهش نرخ خطا در سیستم‌های تشخیص گفتار کمک کرده و نوآوری‌هایی مانند

۱. شورای ملی تحقیقات هیئت ارتباطات علوم رایانه، ادامه نوآوری در فناوری اطلاعات (The National Academies Press, Washington, D.C., 2012); <https://doi.org/13427/10.17226>.

ترجمه بلادرنگ را امکان‌پذیر ساخته است.^۱

• به همین ترتیب، در دهه‌های ۱۹۸۰ و ۱۹۹۰، سرمایه‌گذاری فدرال در منجر به سیستم‌های یادگیری عمیق امروزی شده است. رویکرد یادگیری تقویتی، رویکردی است که ریشه در روان‌شناسی رفتاری دارد و شامل یادگیری برای تداعی رفتارهایی با پیامدهای مطلوب است. کامپیوترها از طریق تعاملات با بخش‌های مختلف، به طور فزاینده‌ای مانند انسان‌ها (بدون دستور صریح) یاد می‌گیرند. یادگیری تقویتی، محرک پیشرفت در ماشین‌های خودکار و دیگر اشکال اتوماسیون است که در آن‌ها، ماشین‌ها می‌توانند مهارت‌ها را از طریق تجربه ارتقا دهند. یادگیری تقویتی فناوری اصلی زیربنایی آلفاگو بود، برنامه‌ای که بهترین بازیکن Go جهان را شکست داد و از سال ۲۰۱۶، شاهد افزایش تعداد پیروزی آن بر بازیکنان حرفه‌ای هستیم.^۲

از زمان انتشار طرح راهبردی تحقیق و توسعه هوش مصنوعی ۲۰۱۶، تأکید بیشتری بر منافع حاصل از مشارکت بخش‌های دولتی-خصوصی صورت گرفته است. این منافع شامل ارتقا راهبردی منابع، از جمله امکانات، مجموعه داده‌ها و تخصص هستند تا نوآوری‌های علم و مهندسی را به ببرند، تبدیل این نوآوری‌ها را به عمل تسریع کنند و ارتقا تحصیل و آموزش را برای پژوهشگران، تکنسین‌ها و رهبران نسل بعدی محقق کنند. مشارکت تحقیق و توسعه دولت-دانشگاه-

1. Restoring the Foundation: The Vital Role of Research in Preserving the American Dream (American Academy of Arts and Sciences, Cambridge, MA, 2014); https://www.amacad.org/multimedia/pdfs/publications/researchpapersmonographs/AmericanAcad_RestoringtheFoundation_Brief.pdf.

2. Restoring the Foundation: The Vital Role of Research in Preserving the American Dream (American Academy of Arts and Sciences, Cambridge, MA, 2014); https://www.amacad.org/multimedia/pdfs/publications/researchpapersmonographs/AmericanAcad_RestoringtheFoundation_Brief.pdf.

صنعت، برای محققان دانشگاهی چالش‌های جدی دنیای واقعی را در مواجهه با صنعت به ارمغان می‌آورد؛ «پژوهش الهام گرفته از کاربرد» را میسر می‌کند؛ تخصص صنعت را ارتقا می‌دهد تا تبدیل نتایج تحقیقات آزاد و منتشر شده را به محصولات و خدمات قابل اطمینان سرعت ببخشد تا رشد اقتصادی اتفاق بیفتد و از طریق پیوند هیئت‌علمی دانشگاه و دانشجویان با نمایندگان صنعت، محیط‌های صنعتی و مشاغل صنعتی، تحقیقات و ظرفیت نیروی کار را افزایش می‌دهد (مربع کناری را ببینید)^۱. این مشارکت‌ها، در میان مؤسسات فدرال تعاملات مشترک را شکل می‌دهد که در حوزه‌هایی که مأموریت مؤسسات با همدیگر تلاقی دارند، موجب هم‌افزایی می‌شود. در ضمن، کشور از روابط میان مؤسسات فدرال و سرمایه‌گذاران بین‌المللی برخوردار می‌شود که می‌توانند با هم کار کنند و چالش‌های کلیدی مورد علاقه دو طرف را در طیفی از رشته‌ها بر طرف کنند.

شراکت‌های دولتی - خصوصی: برنامه‌های اخیر مؤسسات تحقیق و توسعه.

شماری از مؤسسات پیش از این، در حمایت از تحقیق و توسعه هوش مصنوعی، مشارکت دولتی - خصوصی را آغاز کرده‌اند:

- واحد نوآوری دفاع^۲ (DIU) یک سازمان وابسته به وزارت دفاع است

۱. گزارش موسسه تحقیقات علوم ریاضی، «مشارکت: کارگاهی در باره همکاری بین NSF/MPS و بنیادهای خصوصی»، ۲۰۱۵، <http://library.msri.org/msri/Partnerships.pdf>.

کنسر سیوم اجتماع محاسبات، «آینده تحقیقات محاسباتی: همکاری‌های صنعت-دانشگاه»، ۲۰۱۶، <http://cra.org/ccc/wp-content/uploads/sites/-15125/06/2016/2CCC-Industry-Whitepaper-v1-4.pdf>.

کنسر سیوم اجتماع محاسبات، «تکامل روابط دانشگاه/صنعت در تحقیقات محاسباتی: گزارش موقت منتشر شده از سوی CCC»، ۲۰۱۹، <https://www.cccblog.org/wp-content/uploads/03/2019/Industry-Interim-Report-w-footnotes.pdf>.
<https://www.diu.mil/>

که به ارائه راه‌حل‌های تجاری می‌پردازد که قادر به رفع نیازهای وزارت دفاع هستند. DIU به نوبه خود قراردادهای آزمایشی را ارائه می‌دهد که می‌توانند شامل سخت‌افزار، نرم‌افزار و یا دیگر خدمات منحصر به فرد باشند. در صورت موفقیت، قراردادهای آزمایشی منجر به قراردادهای بعدی بین شرکت‌ها و هر نهاد DOD می‌شود. یکی از ویژگی‌های کلیدی DID، سرعت قراردادهای آزمایشی و بعدی است.

- NSF و مشارکت در هوش مصنوعی، سازمانی با ذینفعان مختلف که برای درک بهتر تأثیرات هوش مصنوعی کار می‌کند و هدف از مشارکتش، حمایت از تحقیقات مشترک با ریسک و پاداش بالا در فصل مشترک ابعاد اجتماعی و فنی هوش مصنوعی است.
- برنامه نوآوری دره سیلیکون (SV²IP) که متعلق به هیئت‌مدیره شرکت DHS Science and Technology است، نگاهی دارد به مهارت زیست‌بوم‌های نوآوری تحقیق و توسعه تجاری کشور و جهان است تا فن‌آوری‌های را بیابد که کاربردهای دولتی دارند. SVIP از یک اپلیکیشن و فرآیند تبلیغ ساده استفاده می‌کند؛ دولت، کارآفرینان و صنعت را گرد هم می‌آورد تا راه‌حل‌های لبه دانش را بیابند و بر روی گذار به بازار و تسریع آن سرمایه‌گذاری مشترک می‌کنند.

- وزارت بهداشت و خدمات انسانی (HHS²) طرح آزمایشی Health Tech Sprint را راه‌اندازی کرده است که در اولین نسخه، «سلامت برتر» نیز نامیده می‌شود. این طرح تا حدودی پس از پروژه فرصت اداره سرشماری، مدل‌سازی شده است. این تلاش

برای همکاری‌های بخش عمومی-خصوصی در باره پیوندهای دوسویه داده‌ها اطلاعات، یک چارچوب چالاک ایجاد کرد. این طرح، مدل‌های جدید تکرار انتشار داده برای آموزش و آزمایش هوش مصنوعی را مورد آزمون قرار داد و یک چارچوب داوطلبانه تشویق را برای زیست‌بوم هوش مصنوعی بخش‌های عمومی-خصوصی ابداع نمود.

• بخش تحقیق، نوآوری و ونچرهای HHS، بخشی از اداره تحقیق و توسعه پیشرفته زیست-پزشکی در دفتر آمادگی و پاسخ معاون وزیر است. این بخش بر یک شبکه شتاب‌دهنده نظارت می‌کند و یک شریک غیرانتفاعی را استخدام می‌کند که می‌تواند با سرمایه‌گذاران خصوصی همکاری کند تا برای حل چالش‌های امنیتی سلامت بر روی فناوری‌ها و محصولات نوآورانه سرمایه‌گذاری کنند. کاربردهای هوش مصنوعی یکی از زمینه‌های مورد علاقه این بخش است. شتاب‌دهنده‌ها، اقدام به اتصال شرکت‌های استارت‌آپی و سایر کسب‌وکارها به توسعه محصول و خدمات پشتیبانی از کسب‌وکارها خواهند کرد.

در حالی که ساختارها و سازوکارهای زیادی برای مشارکت بخش‌های دولتی-خصوصی وجود دارند، برخی از مقولات مشترک برای مشارکت عبارت‌اند از:

۱- همکاری‌های فردی مبتنی بر پروژه. این تلاش‌ها به منظور شناسایی و بهره‌برداری از هم‌افزایی در حوزه‌های منافع مشترک، تعامل محققان دانشگاهی با محققان بخش‌های دیگر از جمله ادارات

فدرال، صنعت و سازمان‌های بین‌المللی را ممکن می‌سازند.

۲- برنامه‌های مشترک برای پیشبرد، تحقیقات بنیادین پیش‌رقابتی. شراکت مستقیم میان سازمان‌ها در بخش‌های مختلف، امکان تأمین بودجه و حمایت از تحقیقات بنیادین باز و پیش‌رقابتی را در حوزه‌هایی با منافع متقابل شرکا فراهم می‌کنند. به طور کلی، شرکای غیر فدرالی که به منابع تحقیقاتی کمک می‌کنند، همان حق مالکیت معنوی را دریافت می‌کنند که توسط قانون Bayh-Dole به دولت آمریکا اعطا شده است.^۱

۳- همکاری‌هایی برای استقرار و افزایش زیرساخت تحقیق. همکاری‌ها در میان مؤسسات فدرال، صنعت و سازمان‌های بین‌المللی به طور قابل توجهی توان بالقوه توسعه زیرساخت پژوهشی جدید و موجود را افزایش می‌دهند که به نوبه خود، باعث انجام آزمایش‌های پیشگامانه توسط محققان می‌شود. شرکا می‌توانند پشتیبانی مالی و/یا غیر نقدی برای توسعه و/یا ارتقا زیرساخت تحقیق را ارائه کنند.

۴- همکاری‌ها برای بهبود توسعه نیروی کار از جمله گسترش مشارکت. شراکت‌های چندبخشی، مبنایی را برای مواد آموزشی دقیق، جذاب و الهام‌بخش تعیین می‌کنند که توسعه نیروی کار و تنوع در حرفه‌های STEM را افزایش می‌دهند.

در هر یک از این موارد، ارتقا نقاط قوت هر شریک در جهت برخورداری همه شرکا از منافع حاصل، برای دستیابی به موفقیت اهمیت حیاتی دارد.

پیشرفت‌ها در حوزه تحقیق و توسعه در گرو بهره‌مندی از همه انواع این شراکت‌های دولتی - خصوصی است. شراکت می‌تواند

1. <https://history.nih.gov/research/downloads/PL517-96.pdf>

موجب ترغیب تحقیق و توسعه باز و پیش رقابتی هوش مصنوعی شود؛ دسترسی به منابع تحقیق مانند مجموعه داده‌ها، مدل‌ها و قابلیت‌های محاسباتی پیشرفته را افزایش دهد و تبادلات و/یا ملاقات‌های مشترک محققان را بین دولت، دانشگاه‌ها و صنعت برای به اشتراک گذاشتن تخصص تحقیق و توسعه هوش مصنوعی ترویج کند. در ضمن، شراکت می‌تواند ماهیت ذاتاً میان‌رشته‌ای تحقیق و توسعه هوش مصنوعی را به اطلاع عموم برساند. این ماهیت مستلزم همگرایی بین کامپیوتر و علم اطلاعات، علوم شناختی و روانشناسی، اقتصاد و نظریه بازی‌ها، مهندسی و نظریه کنترل، اخلاق، زبان‌شناسی، ریاضی و آمار و فلسفه است تا محرکی باشد برای توسعه و ارزیابی سیستم‌های هوش مصنوعی آتی که منصفانه، شفاف و ایمن باشند. مؤسسات فدرال فعالانه برای دستیابی به این اهداف مشارکت می‌کنند (مربع کناری را ببینید).

بنابراین مؤسسات فدرال باید به دنبال پیگیری و تقویت مشارکت‌های دولتی و خصوصی در زمینه تحقیق و توسعه هوش مصنوعی باشند تا با ارتقا سرمایه‌گذاری و تخصص در زمینه‌های مورد علاقه متقابل دولت، صنعت و دانشگاه، محرکی برای توسعه فناوری و رشد اقتصادی باشند. در انجام این کار، دولت آمریکا بر روی یک زیست‌بوم نوآوری منحصربه‌فرد آمریکایی سرمایه‌گذاری خواهد کرد که در طول دو دهه گذشته تقریباً هر جنبه از اقتصاد و جامعه را از طریق فن‌آوری‌های اطلاعاتی بدیع متحول کرده است.¹

1. National Research Council Computer Science Telecommunications Board, Continuing Innovation in Information Technology (The National Academies Press, Washington, D.C., 2012); <https://doi.org/10.17226>.

Abbreviations

AFOSR	Air Force Office of Scientific Research	NASA	National Aeronautics and Space Administration
AI	artificial intelligence	NCO	National Coordination Office for NITRD
DARPA	Defense Advanced Research Projects Agency	NDS	Naturalistic Driving Study (DOT)
DHS	Department of Homeland Security	NIFA	National Institute of Food and Agriculture (USDA)
DoD	Department of Defense	NIH	National Institutes of Health
DOE	Department of Energy	NIST	National Institute of Standards and Technology
DOT	Department of Transportation	NITRD	Networking and Information Technology Research and Development program
FDA	Food and Drug Administration	NLM	National Library of Medicine (NIH)
FRVT	Face Recognition Vendor Test	NSF	National Science Foundation
GPS	Global Positioning System	NSTC	National Science and Technology Council
GPU	graphics processing unit	NTIA	National Telecommunications and Information Administration
GSA	General Services Administration	ODNI	Office of the Director of National Intelligence
HHS	Department of Health and Human Services	OSTP	Office of Science and Technology Policy
HPC	high-performance computing	R&D	research and development
IARPA	Intelligence Advanced Research Projects Activity	RFI	Request for Information
IEC	International Electrotechnical Commission	STEM	science, technology, engineering, and mathematics
IEEE	Institute of Electrical and Electronics Engineers	SWIP	Silicon Valley Innovation Program (DHS)
IMPACT	Information Marketplace for Policy and Analysis of Cyber-risk & Trust (DHS)	TREC	Text Retrieval Conference
ISO	International Organization for Standardization	USDA	U.S. Department of Agriculture
IT	information technology	VA	U.S. Department of Veterans Affairs
IWG	interagency working group	XAI	explainable AI
ML	machine learning		
MLAI	Machine Learning and Artificial Intelligence (Subcommittee of the NSTC)		



بخش نهم

درباره شورای ملی علوم و فناوری



شورای ملی علوم و فن‌آوری (NSTC) ابزاری اصلی است که به وسیله آن، شاخه اجرایی به هماهنگ کردن علم و سیاست فناوری در میان نهادهای مختلف می‌پردازد که موسسه تحقیقات و توسعه فدرال را تشکیل می‌دهند. هدف اصلی NTSC تضمین این است که برنامه‌ها و تصمیمات سیاست علم و فن‌آوری با اهداف بیان شده رئیس‌جمهور سازگار باشند. NTSC راهبردهای تحقیق و توسعه را آماده می‌کند که با هدف دستیابی به اهداف چندگانه ملی، در میان همه مؤسسات فدرال هماهنگ می‌شوند. کار NTSC توسط کمیته‌هایی سازمان‌دهی می‌شود که سرپرستی گروه‌های فرعی و گروه‌های کاری بر عهده دارند. این گروه‌ها بر جنبه‌های مختلف علوم و فن‌آوری متمرکز هستند. اطلاعات بیشتر در سایت زیر در دسترس است:

<https://www.whitehouse.gov/ostp/nstc>

در باره دفتر سیاست‌گذاری علم و فناوری

اداره سیاست‌گذاری علم و فن‌آوری (OSTP) بر اساس قانون سیاست‌گذاری ملی علوم و فناوری، سازمان و اولویت‌ها و در سال

۱۹۷۶ برای ارائه مشاوره در باره جنبه‌های علمی، مهندسی، امنیت داخلی، سلامت، روابط خارجی، محیط‌زیست و بهبود فن‌آوری و استفاده از منابع تأسیس شد. OSTP تلاش‌های هماهنگ کننده بین مؤسسات علم و فناوری را هدایت می‌کند؛ با بررسی سالانه و تحلیل تحقیقات و توسعه فدرال (تحقیق و توسعه) در بودجه‌ها به دفتر مدیریت و بودجه کمک می‌کند و با توجه به سیاست‌های اصلی، طرح‌ها و برنامه‌های دولت فدرال، به عنوان یک منبع تحلیل علمی و فنی و قضاوت برای رئیس‌جمهور عمل می‌کند. اطلاعات بیشتر در سایت زیر در دسترس است:

<https://www.whitehouse.gov/ostp>

در باره کمیته منتخب هوش مصنوعی

کمیته منتخب هوش مصنوعی برای بهبود کارایی کلی و بهره‌وری تلاش‌های تحقیق و توسعه فدرال مرتبط با هوش مصنوعی به NSTC توصیه کرده و کمک می‌کند تا از ادامه رهبری آمریکا در این زمینه اطمینان حاصل شود. این کمیته به موضوعات سیاست ملی و بین‌المللی می‌پردازد که مرزهای عاملت را شامل می‌شود و سازوکارهای رسمی هماهنگی و توسعه فعالیت‌های همکاری بین مؤسسات را نشان می‌دهد، فعالیت‌هایی مانند آن‌هایی که به سیستم‌های مستقل، شناسایی بیومتریک، بینایی رایانه‌ای، پردازش زبان طبیعی، پردازش زبان طبیعی، و رباتیک مربوط هستند. در ضمن، این کمیته در باره اولویت‌های تحقیق و توسعه هوش مصنوعی به دفتر اجرایی می‌دهد؛ برای ایجاد برنامه‌های توسعه متوازن و جامع

هوش مصنوعی تحقیق می‌کند؛ داده‌های فدرال و منابع محاسباتی را در سراسر مأموریت‌های اداره و موسسه پشتیبانی می‌کند و از یک نیروی کار فنی و ملی هوش مصنوعی پشتیبانی می‌کند.

در مورد کمیته فرعی یادگیری ماشینی و هوش مصنوعی

کمیته فرعی یادگیری ماشین و هوش مصنوعی (MLAI)، وضعیت کنونی یادگیری ماشین (ML) و هوش مصنوعی را در دولت فدرال، در بخش خصوصی پایش می‌کند. بعلاوه به صورت بین‌المللی مراقب وقوع نقاط عطف فناوری در توسعه هوش مصنوعی است تا استفاده و به اشتراک‌گذاری دانش و بهترین تجارب در مورد ML و هوش مصنوعی توسط دولت فدرال را هماهنگ کند. این کمیته در باره توسعه اولویت‌های تحقیق و توسعه MLAI فدرال نیز مشورت می‌دهد. کمیته فرعی MLAI به کمیته فن‌آوری و کمیته منتخب هوش مصنوعی گزارش می‌دهد. کمیته فرعی MLAI همچنین هوش مصنوعی را با گروه کاری توسعه بین‌سازمانی بین‌سازمانی هماهنگ می‌کند.

در باره کمیته فرعی تحقیق و توسعه شبکه‌سازی و فن‌آوری اطلاعات

برنامه تحقیق و توسعه شبکه‌سازی و فن‌آوری اطلاعات (NITRD)، منبع اصلی کار بر روی فن‌آوری‌های پیشرو اطلاعات (IT) در محاسبات، شبکه‌سازی و نرم‌افزار است که بودجه فدرال دارد. کمیته فرعی NITRD برنامه بین‌موسسه‌ای NITRD را در کارش هدایت می‌کند. هدف این موسسه فراهم کردن مبانی تحقیق و توسعه برای تضمین ادامه

رهبری فناوریانه آمریکا و برآورده کردن نیازهای کشور برای فن‌آوری پیشرفته است. این کمیته به کمیته NSTC در زمینه علم و فن‌آوری گزارش می‌دهد. کمیته فرعی توسط گروه‌های کاری بین موسسه‌ای که به آن گزارش می‌دهند و دفتر هماهنگی ملی آن پشتیبانی می‌شود. اطلاعات بیشتر در این باره در سایت زیر در دسترس است:

<https://www.nitrd.gov/about/>

در باره گروه کاری بین موسسه‌ای تحقیق و توسعه هوش مصنوعی

گروه کاری بین موسسه‌ای تحقیق و توسعه هوش مصنوعی (IWD) کمیته NITRD، تحقیق و توسعه فدرال در زمینه هوش مصنوعی را هماهنگ می‌کند؛ همچنین پشتیبانی و هماهنگ فعالیت‌هایی را که توسط کمیته منتخب بر روی هوش مصنوعی و کمیته فرعی NSTC بر آموزش ماشینی و هوش مصنوعی انجام می‌شود، به عهده دارد. این کار حیاتی باعث ترویج رهبری آمریکا و رقابت جهانی در هوش مصنوعی می‌شود. این گروه کاری به روز رسانی این برنامه راهبردی تحقیق و توسعه هوش مصنوعی را هدایت کرد. اطلاعات بیشتر در سایت زیر در دسترس است

<https://www.nitrd.gov/groups/AI>

در باره این نوشتار

این سند شامل متن اصلی برنامه راهبردی ملی تحقیق و توسعه هوش مصنوعی ۲۰۱۶ و نسخه آماده شده در سال ۲۰۱۹ است که پس از مدیریت و ارزیابی بین موسسه‌ای طرح ۲۰۱۶ و پاسخ‌های جامعه

به درخواست برای اطلاعات در باره به روز رسانی برنامه تدوین شد. راهبردهای ۲۰۱۶ به طور گسترده مصمم بودند تا با برخی تأکيدات مکرر و فراخوان برای یک راهبرد جدید در مشارکت خصوصی - عمومی در هوش مصنوعی، پیش رفته و همچنان معتبر باشند. یک جعبه فراخوان سایه‌دار در بالای هر راهبرد قرار داده شده است تا الزامات به روز رسانی شده و/یا حوزه‌های تمرکز جدید را مشخص کند. به روز رسانی سال ۲۰۱۹ یک راهبرد کاملاً جدید ۸ را به مشارکت خصوصی - عمومی در هوش مصنوعی اضافه می‌کند.

اطلاعات کپی‌رایت

این سند کار دولت آمریکا است و در حوزه عمومی قرار دارد (۱۷) U.S.C. را ببینید). این سند می‌تواند آزادانه توزیع شده، کپی شده و تبدیل شود و از OSTP تشکر شود؛ درخواست برای استفاده از هر گونه تصویر باید به OSTP داده شود. منتشر شده در آمریکا ۲۰۱۹.



مرکز ملی فضای مجازی
پژوهشگاه فضای مجازی

حوزه فضای مجازی به اندازه انقلاب اسلامی اهمیت دارد. این فضا مثل یک رودخانه پر از آب و خروشان است که می آید و دائماً هم بر آب آن افزوده و خروشان تر می شود. اگر ما بر این رودخانه تدبیر کنیم و برنامه داشته باشیم، زهکشی کنیم و هدایت کنیم این رودخانه را تا به سد بریزد، می شود فرصت. اگر رهاش کنیم و برنامه ای برای آن نداشته باشیم می شود یک تهدید.



csri.majazi.ir