

توصیه‌های یونسکو درباره اخلاق هوش مصنوعی

<https://saebi.ir/blog/unaiethics#:~:text=%D8%AE%D9%84%D8%A7%D8%B5%D9%87%20%D8%A7%D8%B5%D9%88%D9%84%20%D8%A9%D9%84%D8%B8%D8%AF%D8%B8%20%D8%AA%D9%88%D8%B5%D8%BC%D9%87%E2%80%8C%D9%87%D8%A7%D8%B8%D8%BC%20%D8%B8%D9%88%D9%86%D8%B3%DA%A9%D9%88.%D8%AA%D8%A8%D8%B9%D8%BC%D8%B6%20%D9%88%20%D9%86%D8%A7%D8%A8%D8%B1%DA%A7%D8%A8%D8%B8%1%D8%B8%20%D8%AC%D8%B4%D9%88%DA%AF%D8%B8%D8%B1%DB%8C%20%D8%A9%D9%86%D8%AF>

توصیه‌های یونسکو درباره اخلاق هوش مصنوعی

در سال ۲۰۲۱، یونسکو «توصیه‌نامه‌ای در مورد اخلاق هوش مصنوعی» را به‌عنوان اولین چارچوب جهانی برای استفاده اخلاقی از هوش مصنوعی به تصویب رساند. این سند ۱۹ اصل کلیدی را ارائه می‌دهد که برای هدایت توسعه و استفاده انسانی از هوش مصنوعی تدوین شده است و هدف آن‌ها این است که سامانه‌های هوش مصنوعی با احترام به حقوق بشر، کرامت انسانی و تنوع فرهنگی توسعه یابند. به دلیل حجم بودن مرجع اصلی این توصیه‌ها، خلاصه‌ای از این موارد به شرح زیر در این مطلب درج شده است.

خلاصه اصول کلیدی توصیه‌های یونسکو درباره اخلاق هوش مصنوعی:

- ۱- **انسان‌محوری:** هوش مصنوعی باید در خدمت بشریت باشد و رفاه و کرامت انسانی را ارتقا دهد.
- ۲- **عدالت و انصاف:** هوش مصنوعی باید به‌گونه‌ای عادلانه و منصفانه توسعه یافته و مورد استفاده قرار گیرد و از تبعیض و نابرابری جلوگیری کند.
- ۳- **فراگیری:** هوش مصنوعی باید برای همه افراد، صرف‌نظر از پیشینه یا توانایی‌های ایشان، قابل‌دسترس و مفید باشد.
- ۴- **مشارکت و مسئولیت:** ذی‌نفعان مختلف باید بتوانند در توسعه و استفاده از هوش مصنوعی مشارکت داشته باشند و همه افراد (اعم از بهره‌برداران و توسعه‌دهندگان) در قبال اقدامات خود در سامانه‌های هوش مصنوعی مسئول باشند.
- ۵- **شفافیت و توضیح‌پذیری:** سیستم‌های هوش مصنوعی باید شفاف باشند و عملکرد آنها قابل‌توضیح باشد.
- ۶- **امنیت و حریم خصوصی:** هوش مصنوعی باید به‌گونه‌ای توسعه و استفاده شود که از امنیت و حریم خصوصی افراد محافظت کند.
- ۷- **پاسخگویی:** باید مکانیسم‌هایی برای پاسخگویی در قبال آسیب‌های ناشی از هوش مصنوعی وجود داشته باشد.
- ۸- **احتیاط (پیشگیری):** (در توسعه و استفاده از هوش مصنوعی باید احتیاط کرد و از خطرات احتمالی آن آگاه بود.
- ۹- **پایداری:** هوش مصنوعی باید به‌گونه‌ای توسعه و استفاده شود که از محیط‌زیست و نسل‌های آینده محافظت کند.
- ۱۰- **تنوع فرهنگی:** هوش مصنوعی باید تنوع فرهنگی را در نظر بگیرد و به آن احترام بگذارد.
- ۱۱- **آگاهی و آموزش:** باید آگاهی عمومی در مورد هوش مصنوعی و مخاطرات و مزایای آن افزایش یابد.
- ۱۲- **همکاری بین‌المللی:** همکاری‌های بین‌المللی برای توسعه و استفاده مسئولانه از هوش مصنوعی ضروری است.
- ۱۳- **حفاظت از داده‌ها:** داده‌ها باید به صورتی مسئولانه جمع‌آوری، استفاده و ذخیره شوند.
- ۱۴- **الگوریتم‌های عادلانه:** الگوریتم‌ها باید به‌گونه‌ای طراحی شوند که از تبعیض و نابرابری جلوگیری کنند.
- ۱۵- **نظارت و ارزیابی مستمر:** توسعه و استفاده از هوش مصنوعی باید به طور مداوم نظارت و ارزیابی شود.
- ۱۶- **قابلیت استمرار:** سیستم‌های هوش مصنوعی باید به‌گونه‌ای طراحی شوند که قابل بهره‌برداری مستمر و انعطاف‌پذیر باشند.
- ۱۷- **مسئولیت‌پذیری سازمانی:** سازمان‌ها باید در قبال توسعه و استفاده از هوش مصنوعی در سازمان خود مسئول باشند.

۱۸- **حاکمیت هوش مصنوعی**: باید چارچوب‌های حاکمیتی برای توسعه و استفاده از هوش مصنوعی ایجاد شود.

۱۹- **آینده هوش مصنوعی**: گفتگو در مورد آینده هوش مصنوعی و خطرات و مزایای آن باید به‌صورت مداوم و یک بحث مستمر انجام شود.

به نظر می‌رسد علاوه بر اقداماتی و تعهداتی که باتوجه به این اصول برای نظام حاکم در هر کشوری پدید می‌آید، لازم است هریک از توسعه‌دهندگان سامانه‌های اطلاعاتی هوشمند (در هر سطحی) این فهرست را مقابل روی خود قرار داده و سامانه‌ها و راه‌حل‌های توسعه داده شده را از منظر رعایت این اصول بسنجند و چنانچه نیاز به اقدامات اصلاحی در این مورد وجود دارد، حتماً آن اقدامات را در مراحل مختلف طراحی و توسعه سامانه‌های اطلاعاتی در نظر گیرند.

به‌عنوان نمونه به همکاران عزیزم در طراحی و توسعه نرم‌افزارها و سامانه‌های هوشمند، پیشنهاد می‌کنم یک ماتریس از کارکردهای سامانه در هر سطر و عناوین این اصول در هر ستون در نظر بگیرند و در مورد هر اصل چند سؤال ساده را در نظر گرفته و بین ۱ تا ۱۰ رتبه‌ای برای رعایت آن اصل در سامانه مورد طراحی در نظر بگیرند. آنگاه به ترتیب از کمترین امتیاز (بیشترین انحراف با این اصول) شروع به بازطراحی و توسعه مجدد سامانه بر اساس اصل یاد شده نمایند. به‌صورت طبیعی پس از مدت کوتاهی، تیم طراحی و توسعه سامانه‌ها و راه‌حل‌های هوشمند، به‌صورت ناخودآگاه این اصول را در کار خود در نظر خواهند گرفت.

برخی از پیامدهای این توصیه‌ها را می‌توان به شرح زیر ذکر نمود:

- ۱- نیاز به تدوین قوانین و مقررات جدید برای تنظیم توسعه و استفاده از هوش مصنوعی
- ۲- افزایش سرمایه‌گذاری در تحقیقات و آموزش در زمینه اخلاق هوش مصنوعی
- ۳- ایجاد آگاهی عمومی در مورد هوش مصنوعی و خطرات و مزایای آن
- ۴- تشویق به گفتگوی بین ذی‌نفعان مختلف در مورد آینده هوش مصنوعی
- ۵- افزایش آگاهی و تعهد توسعه‌دهندگان سامانه‌های اطلاعاتی هوشمند به رعایت این اصول

توصیه‌نامه یونسکو در مورد اخلاق هوش مصنوعی نقطه عطفی در تلاش‌های جهانی برای اطمینان از استفاده مسئولانه از این فناوری قدرتمند است. این سند به‌صورتی طراحی شده است که بتواند به‌عنوان یک راهنمای مهم برای دولت‌ها، شرکت‌ها و افراد در مسیر توسعه و استفاده از هوش مصنوعی به‌گونه‌ای که به نفع همه باشد، عمل کند. این خلاصه فقط به برخی از اصول کلیدی و پیامدهای توصیه‌نامه یونسکو می‌پردازد. برای اطلاعات بیشتر، می‌توانید به متن کامل سند مراجعه کنید.

بحث در مورد اخلاق هوش مصنوعی یک موضوع پیچیده و دائماً در حال تکامل تدریجی است. مهم است که در مورد این موضوع گفتگوی فعال داشته باشیم و به دنبال راه‌هایی برای تضمین استفاده از هوش مصنوعی به‌گونه‌ای باشیم که به نفع همه باشد. شایان ذکر است که برای ارزیابی سامانه‌ها در رعایت این اصول چارچوب‌ها و معیارهای مشخصی تدوین شده است که امیدوارم در آینده نزدیک امکان انتشار مطالب بیشتری در آن حوزه فراهم آید.

لینک کوتاه b2n.ir/unaiethics برای ارجاع ساده این مطلب در نظر گرفته شده است.