

Zawody przyszłości w przemyśle 5.0

etyk AI

główny specjalista ds. etycznego wykorzystania sztucznej inteligencji

Dokument ekspercki, będący źródłem informacji dla materiałów e-learningowych
w projekcie "Zawody przyszłości w przemyśle 5.0".

Opracowanie: Futurama, proofreading: Marek Ostafil.



1 WPROWADZENIE

Wielki dylemat etyczny

„Ufamy sztucznej inteligencji, czy nie ufamy – oto jest pytanie”.

W miarę jak sztuczna inteligencja wnika w rdzeń społeczeństwa, wyzwanie nie polega już tylko na uczynieniu AI mądrzejszą, lecz na zapewnieniu, aby działała sprawiedliwie, etycznie i odpowiedzialnie. Specjalista ds. etyki AI powinien kształtować polityki, audytować algorytmy i zapewniać, że decyzje podejmowane przez zaawansowane algorytmy są zgodne z prawami człowieka i wartościami społecznymi. W świecie, w którym wpływ AI poszerza się szybciej niż powstają regulacje, pojawia się pytanie: kto może zapewnić, że AI służy ludzkości, a nie odwrotnie?

2 ZAWÓD PRZYSZŁOŚCI

Rosnące znaczenie etyki AI

Szybki rozwój sztucznej inteligencji wywołał transformacyjne zmiany w różnych branżach, w tym w opiece zdrowotnej, finansach, transporcie i zarządzaniu. Systemy oparte na AI mają obecnie wpływ na kluczowe procesy decyzyjne, oddziałując na jednostki i społeczeństwa w sposób bezprecedensowy. Jednak obok tych przełomów technologicznych pojawiły się poważne wyzwania etyczne, w tym uprzedzenia w modelach AI, kwestie prywatności, brak przejrzystości oraz możliwość niewłaściwego wykorzystania systemów autonomicznych. Wyobraźmy sobie chaos, który mógłby wyniknąć, gdyby uprzedzone algorytmy AI powodowały błędne diagnozy medyczne lub gdyby systemy sądowe sterowane AI skazywały niewinnych ludzi.

W miarę jak sztuczna inteligencja przekształca branże, rosną obawy dotyczące implikacji etycznych. Raport Światowego Forum Ekonomicznego z 2023 roku stwierdza, że ponad 60% globalnych liderów biznesu uznaje etykę AI za priorytet. Jednak tylko 20% ma jasne wytyczne dotyczące odpowiedzialnego korzystania z tego rodzaju algorytmów. Podobnie raport MIT Sloan z 2022 roku wykazał, że uprzedzenia w podejmowaniu decyzji przez AI wpływają na obszary od rekrutacji po udzielanie pożyczek, wzmacniając nierówności systemowe.

Zapobieganie naruszeniom praw człowieka, budowanie zaufania do systemów, promowanie odpowiedzialnych badań i innowacji, zapewnienie przejrzystości

działania algorytmów oraz podejmowania decyzji, uwzględnianie skutków ekonomicznych, ograniczanie realnych szkód, czy zapewnienie, że systemy AI w obszarach krytycznych działają zgodnie z najwyższymi standardami bezpieczeństwa - to kluczowe wyzwania w dynamicznie rozwijającym się nowoczesnym świecie technologicznym.

W odpowiedzi na te wyzwania zaproponowano ramy regulacyjne, takie jak unijny AI Act, amerykański AI Bill of Rights oraz powstające standardy ISO (ISO/IEC 42001). Ich zadaniem jest ustanowienie globalnych wytycznych dotyczących odpowiedzialnego zarządzania AI. Już w 2020 roku amerykański Departament Obrony przyjął pięć zasad etyki sztucznej inteligencji:

1. Odpowiedzialność.
2. Sprawiedliwość.
3. Kontrola.
4. Niezawodność.
5. Zarządzanie.

Komisja Europejska wdrożyła wspomniany AI Act, pierwsze na świecie kompleksowe ramy prawne dotyczące tego obszaru. AI Act wszedł w życie 1 sierpnia 2024 roku. Regulacja klasyfikuje zastosowania AI według poziomu ryzyka:

- Ryzyko niedopuszczalne: Systemy AI, które stwarzają istotne zagrożenia, takie jak manipulacja zachowaniem ludzi lub umożliwianie oceniania społecznego - są zabronione.
- Ryzyko wysokie: Zastosowania w sektorach krytycznych, takich jak ochrona zdrowia, edukacja czy wymiar sprawiedliwości, muszą spełniać rygorystyczne wymagania dotyczące jakości danych, przejrzystości, nadzoru człowieka i bezpieczeństwa.
- Ryzyko ograniczone: Systemy o ograniczonym ryzyku podlegają obowiązkowi przejrzystości, przekazując użytkownikom informacje w trakcie interakcji z AI.
- Ryzyko minimalne: Aplikacje o minimalnym ryzyku, takie jak AI w grach wideo czy filtry antyspamowe - są w dużej mierze nieuregulowane.

Ustawa ustanawia również struktury zarządcze, w tym Europejską Radę ds. Sztucznej Inteligencji i krajowe organy nadzorcze, które monitorują zgodność i egzekwowanie przepisów.

W lutym 2025 roku Komisja Europejska wydała wytyczne zabraniające niektórych praktyk AI, takich jak wykorzystywanie przez pracodawców algorytmów do monitorowania emocji pracowników czy stosowanie w internecie narzędzi do manipulowania wydatkami użytkowników. Przepisy mają wejść w życie w sierpniu 2026 roku.

Jednak same ramy prawne nie wystarczą. Organizacje potrzebują specjalistów, którzy zapewnią, że systemy AI są zgodne z zasadami etycznymi, wymogami prawnymi i oczekiwaniami społecznymi.

W ten sposób powstał nowy zawód: główny specjalista ds. etycznego wykorzystania sztucznej inteligencji, czyli etyk AI. Rola łączy innowacje technologiczne, etykę i politykę, zapewniając, że AI jest stosowana sprawiedliwie, przejrzystie, odpowiedzialnie i społecznie odpowiedzialnie.

Wyzwania związane z uprzedzeniami i sprawiedliwością AI

Jedną z najlepiej udokumentowanych porażek AI są uprzedzenia rasowe i płciowe w algorytmach rekrutacyjnych. Badanie MIT Media Lab z 2018 roku wykazało, że systemy rozpoznawania twarzy miały wskaźnik błędu 34,7% dla kobiet o ciemnej karnacji w porównaniu do zaledwie 0,8% dla mężczyzn o jasnej karnacji. Etycy AI pracują nad zapobieganiem takim uprzedzeniom w narzędziach rekrutacyjnych, algorytmach bankowych i systemach egzekwowania prawa.

Przykład dobrej praktyki: IBM wprowadziło AI Fairness 360, otwartoźródłowe narzędzie, które pomaga firmom wykrywać i redukować uprzedzenia w modelach uczenia maszynowego.

Prywatność i ochrona danych

Systemy AI zbierają i przetwarzają ogromne ilości danych osobowych. Badanie Pew Research z 2021 roku ujawniło, że 79% ludzi obawia się, jak firmy wykorzystują ich dane.

Dobra praktyka: Apple's Differential Privacy Model zapewnia, że personalizacja oparta na AI nie narusza prywatności jednostki - poprzez użycie zanonimizowanych zestawów danych.

Profesja etyka AI staje się kluczowym zawodem przyszłości, ponieważ szybkie rozpowszechnianie generatywnej AI, narzędzi do śledzenia emocji i systemów podejmowania autonomicznych decyzji zdecydowanie wyprzedza istniejące ramy etyczne. Poza samym przeglądaniem algorytmów, etycy kształtują teraz wewnętrzne kultury firm i projektują produkty od samego początku z uwzględnieniem długoterminowych skutków społecznych. Ich wiedza będzie niezbędna nie tylko do zapobiegania incydentom, ale także do budowania systemów wspierających włączającą, demokratyczną i zorientowaną na człowieka cyfrową przyszłość.

3 GŁÓWNE OBOWIĄZKI

Kim jest etyk AI?

Etyk AI to ekspert interdyscyplinarny, który doradza organizacjom w odpowiedzialnym rozwoju, wdrażaniu i zarządzaniu systemami sztucznej inteligencji. Główny specjalista ds. etycznego wykorzystania AI odgrywa kluczową rolę w równoważeniu postępu technologicznego z odpowiedzialnością etyczną. Jego głównym celem jest zapobieganie uprzedzeniom, dyskryminacji i niezamierzonym negatywnym konsekwencjom, przy jednoczesnym promowaniu sprawiedliwości, odpowiedzialności i zgodności ze standardami etycznymi. Etycy AI ściśle współpracują z inżynierami, decydentami, zespołami prawnymi i liderami biznesu, aby promować kwestie etyczne w rozwój zaawansowanych algorytmów.

Kluczowe obowiązki etyka AI

- Tworzenie polityk etycznej AI – ustanawianie wytycznych i ram dla odpowiedzialnego korzystania z algorytmów.
- Wprowadzanie standardów odpowiedzialnego rozwoju AI – zapewnianie, że systemy są rozwijane zgodnie z wartościami etycznymi i oczekiwaniami społecznymi.
- Audyt uprzedzeń i sprawiedliwości algorytmicznej – identyfikacja i łagodzenie uprzedzeń w modelach uczenia maszynowego.
- Przegląd etyczny projektów AI – przed wdrożeniem nowego narzędzia etyk dokonuje przeglądu, aby zidentyfikować potencjalne problemy etyczne i zaproponować rozwiązania.
- Zapewnienie zgodności z regulacjami światowymi – dopasowanie polityk AI do przepisów takich jak EU AI Act, RODO czy do standardów etycznych IEEE.

- Ocena ryzyka i wpływu – ocena społecznych i etycznych konsekwencji decyzji wspieranych przez AI.
- Współpraca z zainteresowanymi stronami – praca z inżynierami, deweloperami, menadżerami produktów, prawnikami, decydentami i społecznościami w celu budowy odpowiedzialnego ekosystemu AI oraz integracji kwestii etycznych na każdym etapie rozwoju algorytmów.
- Przejrzystość i wyjaśnialność AI – zapewnienie, że modele generują jasne, zrozumiałe i uzasadnione wyniki.
- Edukacja i podnoszenie świadomości na temat etycznego korzystania z AI – poprzez edukację członków zespołów organizacja rozwija kulturę odpowiedzialności i świadomości etycznej.

W miarę kształtowania się regulacji AI na całym świecie, zapotrzebowanie na etyków gwałtownie rośnie. Raport PwC z 2023 roku przewiduje, że globalny rynek zarządzania AI wzrośnie do 3,7 mld USD do 2026 roku, a liczba ofert pracy w tym obszarze powiększy się o 200% w ciągu zaledwie dwóch lat. Giganci branży technologicznej, tacy jak Google, Microsoft, IBM i OpenAI, tworzą specjalne zespoły ds. etyki AI w celu kształtowania odpowiedzialnych polityk.

Poza branżą technologiczną etycy stają się niezbędni w finansach, opiece zdrowotnej, edukacji, administracji publicznej i branży gamingowej. W grach AI odpowiada za adaptacyjne zachowania NPC, moderację treści i spersonalizowane doświadczenia, ale kwestie etyczne, takie jak manipulacja graczami, uprzedzenia w algorytmach i ryzyka dot. prywatności, wymagają nadzoru. Firmy angażują teraz specjalistów ds. etyki AI, aby zapewnić sprawiedliwość, przejrzystość i odpowiedzialne korzystanie z AI w procesie tworzenia gier i interakcji z graczami, gwarantując, że decyzje wspierane algorytmicznie pozostaną uczciwe, przejrzyste i odpowiedzialne.

4 ŚCIEŻKA KARIERY

Edukacja

- Licencjat lub magisterium w dziedzinie AI.
- Informatyka i znajomość algorytmów.
- Filozofia, prawo lub etyka.

Kursy specjalistyczne i certyfikaty

- Etyka AI.

- Odpowiedzialna AI.
- AI i prawa człowieka.

5 KOMPETENCJE

Umiejętności miękkie

- Silne zdolności analityczne i krytyczne myślenie.
- Efektywna komunikacja i umiejętność współpracy.

Szkolenia

- Doświadczenie zbierane w firmach technologicznych, instytucjach badawczych lub organach regulacyjnych.
- Praktyki na stanowiskach związanych z etyką AI.

Podstawowe umiejętności twarde

- Rozumienie systemów AI – jak działają algorytmy i jak podejmują decyzje.
- Prywatność danych – ochrona danych osobowych i przestrzeganie przepisów o ochronie prywatności.
- Wykrywanie uprzedzeń – identyfikacja i naprawa niesprawiedliwych lub dyskryminacyjnych wyników AI.
- Prawo i etyka AI – znajomość przepisów, takich jak EU AI Act, i umiejętność ich stosowania.
- Wyjaśnialna AI – zapewnienie, że decyzje AI można jasno wytłumaczyć.
- Ocena ryzyka etycznego – ocena, czy system AI może wyrządzić krzywdę jednostkom lub społeczeństwu.
- Odpowiedzialne projektowanie – tworzenie AI, która jest sprawiedliwa, bezpieczna i przejrzysta.

Networking i rozwój zawodowy

- Dołączanie do organizacji i forów ds. etyki AI.
- Uczestnictwo w konferencjach i warsztatach.
- Znajomość odpowiednich ram i regulacji etycznych.

6 POKREWNE ZAWODY

Data Scientist

Kluczowa rola: przekształca surowe dane w informacje wspierające decyzje.

Buduje modele uczenia maszynowego i systemy na podstawie dużych zbiorów danych.

Analizuje trendy w danych, wykrywa wzorce i zapewnia dokładność modeli.
Współpracuje z ekspertami ds. etyki AI, aby dane były uczciwe i pozbawione uprzedzeń.

Monitoruje wpływ systemów opartych na danych i poprawia ich wydajność.

AI Researcher

Kluczowa rola: opracowuje nowe algorytmy i przesuwa granice sztucznej inteligencji.

Prowadzi badania nad deep learningiem, przetwarzaniem języka naturalnego i robotyką.

Projektuje eksperymentalne modele oraz testuje ich możliwości i ograniczenia.

Ścisłe współpracuje z etykami w celu przeciwdziałania ryzyku, takiemu jak uprzedzenia czy niewłaściwe wykorzystanie.

Publikuje wyniki badań, aby kształtować przyszłość rozwoju AI.

Technology Lawyer

Kluczowa rola: zapewnia zgodność systemów AI i produktów technologicznych z prawem.

Doradza firmom w sprawach dotyczących AI, takich jak RODO i AI Act.

Opracowuje i przegląda polityki prawne dotyczące wykorzystania danych, własności intelektualnej i odpowiedzialności.

Wspiera etyków AI w interpretacji przepisów i standardów etycznych.

Pomaga kształtować ramy prawne dla nowych technologii.

Konsultant etyczny

Kluczowa rola: doradza organizacjom w podejmowaniu decyzji opartych na wartościach.

Ocenia ryzyka etyczne pod kątem produktów, polityk i technologii.

Projektuje programy szkoleniowe i ramy decyzyjne w zakresie etyki.

Współpracuje z interesariuszami, aby praktyki AI były zgodne z wartościami społecznymi.

Zapewnia przejrzystość, sprawiedliwość i odpowiedzialność we wszystkich projektach organizacji.

7 PODSUMOWANIE

Główny specjalista ds. etyki AI to nie tylko rola techniczna – to stanowisko przywódcze, które zapewnia, że sztuczna inteligencja staje się odpowiedzialnym

narzędziem. W miarę jak AI ewoluuje, organizacje, które priorytetowo traktują etyczne zarządzanie, zyskują zaufanie, redukują ryzyko i zapewniają, że innowacje w polu ich oddziaływania respektują prawa człowieka oraz wartości społeczne. Etyczna AI to nie tylko kwestia techniczna. Chodzi o wspólną odpowiedzialność. Etycy odegrają kluczową rolę w kształtowaniu przyszłości, w której AI nie tylko wspiera potencjał człowieka, ale również gwarantuje sprawiedliwość, przejrzystość i odpowiedzialność.

8 KLUCZOWE TERMINY

- **Uprzedzenia** – tendencja modelu AI do faworyzowania lub dyskryminowania określonych grup na podstawie danych wejściowych. Jeden z kluczowych problemów etycznych w sztucznej inteligencji.
- **Wyjaśnialność** – zdolność algorytmu do jasnego przedstawienia, jak i dlaczego podjął konkretną decyzję. Podstawowa dla przejrzystości systemu i budowania zaufania użytkowników.
- **Odpowiedzialność** – pytanie, kto ponosi odpowiedzialność za decyzje podejmowane przez systemy AI? Projektanci, użytkownicy czy organizacje?

9 ŹRÓDŁA

- <https://www.oecd.org/going-digital/ai/principles/>
- <https://arxiv.org/abs/1702.08608>
- <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>
- <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- <https://www.governance.ai/>
- <https://standards.ieee.org/industry-connections/activities/ieee-global-initiative/>
- <https://harvardmagazine.com/sites/default/files/pdf/2019/01-pdfs/0119-44.pdf>
- <https://cyber.harvard.edu/topics/ethics-and-governance-ai>
- <https://professional.dce.harvard.edu/programs/ethics-of-ai>
- https://eitca.org/eitca-ai-artificial-intelligence-academy/?ch=c6ju43&gad_source=1&gad_campaignid=13286655760&gbraid=0AAAAABZqwiZqdpu_jyd8u7mibtov-C2P_&gclid=Cj0KCQjwmqPDBhCAARIsADorxIzQFcl-Hzf08J7DCeryuw-dHF2dd6oIf0H-xRsJlgYNL-z0EN3oAfAaAsArEALw_wcB
- <https://eetikakeskus.ut.ee/en/content/international-ma-philosophy-combines-philosophy-practice-and-todays-global-and-local>
- <https://initiatives.weforum.org/ai-governance-alliance/home>

Dofinansowane ze środków UE. Wyrażone poglądy i opinie są jedynie opiniami autora lub autorów i niekoniecznie odzwierciedlają poglądy i opinie Unii Europejskiej lub Fundacji Rozwoju Systemu Edukacji. Unia Europejska ani Fundacja Rozwoju Systemu Edukacji nie ponoszą za nie odpowiedzialności.

Wszystkie rezultaty opracowane w ramach projektu „Zawody przyszłości w przemyśle 5.0” są udostępniane na otwartej licencji Creative Commons Attribution – ShareAlike 4.0 International (CC BY-SA 4.0). Dozwolone jest ich nieodpłatne rozpowszechnianie, pod warunkiem wskazania autorstwa oraz źródła finansowania projektu.



Dofinansowane przez
Unię Europejską

