

MICHAŁ KLEIBER *

Etyka w sztucznej inteligencji – działania UNESCO

Niedawno otworzona w Instytucie Podstawowych Problemów Techniki PAN prestiżowa Katedra UNESCO dotycząca zarządzania etyką w sztucznej inteligencji kolejny raz wskazuje na niezwykle ważny problem ściśle związany z wykorzystywaniem AI. Zaczniemy jednak od wyjaśnienia, co oznacza termin Katedra UNESCO. Odnosi się on do naukowych zespołów wyróżnionych w skomplikowanej procedurze selekcji przez jedną z głównych organizacji ONZ-u w ramach programu o nazwie UNITWIN/UNESCO Chairs. UNITWIN jest skrótem angielskiego wyrażenia University Twinning and Networking, oznaczającego partnerstwo uniwersyteckie i tworzenie sieci kontaktów, a chairs to szefowie powoływanych katedr. Celem tego programu jest wspieranie badań, współpracy instytucji badawczych z różnych krajów oraz kształcenie w obszarach działalności UNESCO, czyli w edukacji, nauce, kulturze i w zrównoważonym rozwoju, a także wspieranie tych obszarów w społeczeństwach całego świata, ze szczególnym uwzględnieniem krajów słabszych gospodarczo. Od Katedr UNESCO i Sieci UNITWIN, złożonej ze współpracujących ze sobą jednostek badawczych z różnych państw, oczekuje się prowadzenia innowacyjnych badań i tworzenia wiedzy potrzebnej do mierzenia się z aktualnymi, stojącymi przed nami wyzwaniami, w tym do realizacji Milenijnych Celów Rozwoju ONZ. Instytucje w tym programie współpracują z organizacjami pozarządowymi, fundacjami oraz organizacjami sektora publicznego i prywatnego, odgrywając w ten sposób ważną rolę w obszarze szkolnictwa wyższego i nauki poszczególnych państw. Istotną rolę w działaniu tych instytucji mają Narodowe Komitety ds. UNESCO, promujące i ułatwiające realizowanie programów oraz wspomagające ocenę uzyskiwanych efektów. Program został zainicjowany przez UNESCO w 1992 r. i obejmuje badania w zakresie edukacji, nauk ścisłych, przyrodniczych, społecznych i humanistycznych oraz szeroko rozumianych obszarów kultury i komunikacji społecznej. Na świecie w 120 krajach działa obecnie prawie tysiąc Katedr UNESCO i 45 Sieci UNITWIN, a liczba badaczy ściśle zaangażowanych w tę działalność wynosi kilkanaście tysięcy. Ponieważ program ma wielodyscyplinarny charakter, jest on jednym z najbardziej szerokich programów

* Prof. dr hab. Michał Kleiber (michal.kleiber@pan.pl), członek rzeczywisty PAN, Instytut Podstawowych Problemów Techniki PAN, Warszawa. Tekst ukazał się w nrze 36 (2025) kwartalnika „Wszystko co Najważniejsze”

Organizacji. Jak dotąd projekty Katedry UNESCO i Sieci UNITWIN okazały się przydatne w ustanawianiu nowych programów nauczania, generowaniu nowych innowacyjnych pomysłów oraz ułatwianiu wzbogacania istniejących programów edukacyjnych przy jednoczesnym poszanowaniu różnorodności kulturowej. Znacznie ponad połowa instytucji zaangażowanych w tę działalność mieści się obecnie w Europie i Północnej Ameryce, ale szeroka działalność promująca znaczenie tej inicjatywy zmierza do zmiany tej sytuacji. Polska nie była do niedawna bardzo aktywna w staraniach o przyznanie naszym instytucjom statusu uczestnika programu, ale ta sytuacja zmieniła się na korzyść w ostatnich latach. Obecnie istnieje u nas 12 Katedr UNESCO o nazwach wskazujących na reprezentowaną przez nie tematykę. Są to Katedry:

- Trwałego Rozwoju na Uniwersytecie Warszawskim,
- Kobiet, Społeczeństwa i Rozwoju na Uniwersytecie Warszawskim,
- im. Janusza Korczaka Pedagogiki Społecznej w Akademii Pedagogiki Społecznej w Warszawie,
- Dziedzictwa Kulturowego i Studiów Miejskich na Uniwersytecie Ekonomicznym w Krakowie,
- Międzynarodowej Promocji Technologii i Edukacji na AGH w Krakowie,
- Badań Instytucjonalnych i Szkolnictwa Wyższego na Uniwersytecie im. Adama Mickiewicza w Poznaniu,
- Całocyciowego Poradnictwa na Uniwersytecie Wrocławskim,
- Prawa i Ochrony Dóbr Kultury na Uniwersytecie Opolskim,
- Ekohydrologii i Ekologii Stosowanej na Uniwersytecie Łódzkim,
- Zrównoważonego Rozwoju Społecznego na Uniwersytecie Szczecińskim,
- Procesów Publicznego i Globalnego Zarządzania Niematerialnym Dziedzictwem Kulturowym na Uniwersytecie Warszawskim,
- Etycznego Zarządzania Sztuczną Inteligencją w Instytucie Podstawowych Problemów PAN w Warszawie.

Dla całości obrazu funkcjonowania Katedr UNESCO w Polsce wymienimy jeszcze cztery Katedry, które funkcjonowały w naszym kraju w ubiegłych latach:

- Praw Człowieka i Pokoju na Uniwersytecie Mikołaja Kopernika w Toruniu,
- Jakości Nauki i Nauczania w Szkolnictwie Wyższym na Uniwersytecie Warmińsko-Mazurskim w Olsztynie,
- Badań nad Przekładem i Komunikacją Międzykulturową na Uniwersytecie Jagiellońskim,
- Edukacji o Holokauście na Uniwersytecie Jagiellońskim.

Z kolei Sieci UNITWIN/UNESCO mające swoich reprezentantów w Polsce to aktualnie

- UNITWIN Network for Underwater Archaeology (Sieć Archeologii Podwodnej), (2012, Uniwersytet Warszawski, Uniwersytet Mikołaja Kopernika w Toruniu),

- UNITWIN Network on Life designing interventions (counseling, guidance, education) for decent work and sustainable development (Sieć projektowania interwencji (doradztwo, poradnictwo, edukacja) na rzecz godnej pracy i zrównoważonego rozwoju), (2017, Uniwersytet Wrocławski),
- UNITWIN Network on Culture in Emergencies (Sieć Kultury w Sytuacjach Nadzwyczajnych), (2024, Uniwersytet im. Adama Mickiewicza w Poznaniu).

Przejdźmy teraz do wskazanej na wstępie problematyki będącej tematem nowo otwartej Katedry, czyli do spraw wykorzystywania sztucznej inteligencji w kontekście konieczności przestrzegania zasad etycznych. Przez etykę w AI rozumiemy zasady funkcjonowania AI zgodnego z powszechnymi ludzkimi wartościami i moralnością, czyli odbywającego się w sposób pod każdym względem korzystny dla społeczeństwa. Rozumiemy przez to rozwijanie, realizację wdrożeń i użytkowanie AI zgodne z wartościami i zasadami odwołującymi się do powszechnie akceptowanych standardów społecznego dobra i zła. Innymi słowy, etyka AI to obszar badań zajmujących się moralnymi, społecznymi i filozoficznymi konsekwencjami związanymi z rozwojem i wykorzystaniem systemów sztucznej inteligencji. Jest interdyscyplinarnym obszarem łączącym badania z zakresu filozofii, informatyki, prawa, socjologii i nauk politycznych, kluczowych dla zapewnienia, że rozwój technologii AI nie spowoduje niezamierzonych, szkodliwych skutków. Musimy być zawsze świadomi, że olbrzymie tempo rozwoju AI rodzi pytania nie tylko technologiczne, ale też moralne. Szczególnie wtedy, gdy algorytmy podejmują autonomiczne decyzje wpływające na nasze dzisiejsze życie, a już niebawem być może na całą przyszłość cywilizacji. Właśnie dlatego potrzebujemy jasnych zasad, które pomogą określić, co AI może, a czego nie powinna robić. Etyka w świecie maszyn to nie fantazja, lecz konieczność, jeśli chcemy, by technologia wspierała człowieka, a nie mu zagrażała. Główne obszary, które są szczególnie istotne w tej dziedzinie to:

- dobrostan i godność człowieka – systemy sztucznej inteligencji powinny zawsze nadawać priorytet dobrostanowi, bezpieczeństwu i szacunkowi dla ludzi, przyczyniając się do spokoju społecznego i rozwoju gospodarczego;
- odpowiedzialność – możliwość ustalenia, kto ponosi konsekwencje działania systemów AI podejmujących decyzje w sposób autonomiczny bez bezpośredniej ingerencji człowieka z naciskiem, aby ostateczne decyzje w kwestiach etycznych zawsze należały do ludzi;
- moralność w autonomicznych systemach – systemy AI wykorzystywane w celach militarnych, generujących cyberataki czy działające w autonomicznych pojazdach, wymagają ustalenia zasad działania w sytuacjach kryzysowych, np. jak postąpić w przypadku zagrożenia życia ludzi;
- ułatwienia ludzkiej kontroli nad wdrożeniami AI poprzez ich przejrzystość i wyjaśnialność;

- dbałość o prywatność i bezpieczeństwo danych, AI przetwarza bowiem ogromne zbiory danych, w tym dane osobowe i wrażliwe, co stwarza ryzyko nieuprawnionego do nich dostępu;
- eliminowanie uprzedzeń i dyskryminacji, czyli promocja społecznej inkluzji. Algorytmy oparte na uczeniu maszynowym mogą być podatne na uprzedzenia, jeśli dane używane do ich trenowania zawierają ukrytą stronniczość. Takie systemy mogą nieświadomie prowadzić do dyskryminacji w obszarach takich jak zatrudnienie czy wymiar sprawiedliwości.

Na świecie podejmowane są różne inicjatywy mające na celu opracowanie wytycznych i regulacji dotyczących odpowiedzialnego rozwoju sztucznej inteligencji. Wśród nich są na przykład:

- Wytyczne Unii Europejskiej (2019) – Komisja Europejska opublikowała dokument zawierający zasady takie, jak: przejrzystość, sprawiedliwość, prywatność i odpowiedzialność w kontekście AI. Komisja Europejska zaproponowała również AI Act.
- Rekomendacje UNESCO dotyczące etyki w AI (2021).
- Inicjatywa OpenAI – Organizacja non-profit OpenAI rozwija sztuczną inteligencję zgodnie z wartościami etycznymi, kładąc nacisk na transparentność i odpowiedzialność.
- Asilomar AI Principles (2017) – Zbiór 23 zasad opracowanych na konferencji w Asilomar w USA, które dotyczą zarówno rozwoju AI, jak i zapobiegania jej złośliwemu wykorzystaniu.

Podkreślić należy, że różne osoby i instytucje związane z wykorzystywaniem AI mają do odegrania różne role w etyce tej technologii. Programiści i badacze na etapie tworzenia systemów sztucznej inteligencji, decydenci i organy regulacyjne poprzez ustanawianie przepisów i regulacji dotyczących wdrażania systemów, liderzy biznesowi poprzez troskę, aby ich organizacje korzystały z tej technologii w sposób pozytywny dla społeczeństwa, społeczne organizacje obywatelskie promując znaczenie etyki w AI, instytucje akademickie i badawcze opracowując wytyczne w zakresie etyki. Wszyscy działając na rzecz zasad dbających o przejrzystość, równość szans i bezpieczeństwo użytkowników.

W kontekście powyższych uwag rola nowo powołanej w Polsce Katedry UNESCO jest jasna – badania dotyczące etycznego wdrażania AI i dotyczące jej regulacje przesyłają o dalszych losach tej technologii, a UNESCO ze względu na swój zasięg ma tu do odegrania znaczącą rolę. Katedry UNESCO dotyczące AI działają m.in. na uczelniach w Londynie, Pekinie, Madrycie, Paryżu, Austin i Lublanie, a problematyką AI zajmują się także liczne sieci UNITWIN. Powiedzmy na koniec dobitnie – etyczna AI to nie przeszkoda, ale fundament zaufania kluczowego dla społecznych korzyści dalszego jej rozwoju, z pewnością ważnego dla naszej wspólnej pomyślnej przyszłości.

Etyka w sztucznej inteligencji – działania UNESCO

Artykuł dotyczy sztucznej inteligencji w kontekście niezwykle ważnych problemów dotyczących etyki jej wykorzystywania. Omawiane są podejmowane na świecie próby regulacji w tym zakresie. W kontekście tego tematu przedstawiona jest także ważna część działań UNESCO w zakresie wspierania badań naukowych.

Słowa kluczowe: sieci UNITWIN i Katedry UNESCO, etyka w sztucznej inteligencji, globalne regulacje

Ethics in Artificial Intelligence – UNESCO Activities

This article examines artificial intelligence in the context of crucial ethical issues surrounding its applications. It discusses current global regulatory efforts in this area. It also highlights a significant portion of UNESCO's efforts to support scientific research.

Key words: UNITWIN networks and UNESCO Chairs, ethics in artificial intelligence, global regulations