

PERSPECTIVA

LEGNICKIE STUDIA TEOLOGICZNO-HISTORYCZNE

Rok XXIV 2025 nr 2 (47) s. 162-163

JAROSŁAW SEMPRYK*

Reiner Mühlhoff,
The Ethics of AI – Power, Critique, Responsibility,
Bristol University Press 2025, ss. 232

Autor wprowadza koncepcję „mocy świadomej” etyki sztucznej (AI) inteligencji, podkreślając, że problematyka etyczna AI to nie tylko kwestie technologiczne, ale przede wszystkim struktury władzy. Zauważa, że AI nie funkcjonuje w próżni – jest osadzona w instytucjach, procesach społecznych i ekonomicznych, które same generują nierówności. Wstęp zawiera także metodyczne uzasadnienie krytycznego podejścia filozoficznego, wywodzącego się z tradycji myślicieli, takich jak Marks, Nietzsche czy przedstawiciele poststrukturalizmu. Autor zaprasza czytelników do wspólnej refleksji nad tym, jak etyka AI może przyczyniać się do systemowych zmian, a nie tylko naprawiania symptomów problemu.

Struktura książki składa się z trzech części. Pierwsza część, zatytułowana *Władza sztucznej inteligencji*, definiuje, co dokładnie rozumiemy przez „sztuczną inteligencję”, pokazując, że nie jest to jednolita kategoria. Autor analizuje historyczne i technologiczne korzenie AI, w tym aspekty uczenia maszynowego i systemów predykcyjnych. Omawia także koncepcję *Human Aided AI* (AI wspieranej ludzkimi działaniami), podkreślając jej hybrydowy charakter oraz współdziałanie ludzi i maszyn. Wskazuje, jak interfejsy i projektowanie wpływają na decyzje użytkowników i ich relacje z technologią. Następnie rozważa, jaką realną władzę mają projektanci systemów AI poprzez swoje decyzje projektowe – nie tylko technologiczne, ale także etyczne i polityczne. Wprowadza pojęcie „cyfrowego przeciw-Oświecenia”, które oznacza przesunięcie od idei racjonalnej autonomii

* Jarosław Sempryk – doktor nauk ekonomicznych w zakresie zarządzania (Uniwersytet Ekonomiczny we Wrocławiu); absolwent Uniwersytetu Wrocławskiego (filozofia); tłumacz przysięgły języka angielskiego; e-mail: sempryk@pro.onet.pl.

użytkownika ku bardziej manipulacyjnym i projektowanym interakcjom. Ostatni podrozdział pierwszej części pokazuje, że systemy AI nie tylko analizują dane, ale również kształtują sposób bycia, pragnienia i decyzje użytkowników, co autor określa jako „subiektywizację technologiczną”.

Druga część książki, *Władza przewidywania*, obejmuje trzy podrozdziały: *Systemy KI jako maszyny predykcyjne*, *Prywatność predykcyjna* oraz *Kultura przewidywania – etyka i epistemologia*. Pierwszy podrozdział analizuje AI jako narzędzie predykcyjne, które prognozuje ludzkie zachowania, potrzeby i ryzyka. Autor zwraca uwagę, że przewidywanie nie jest neutralne – kształtują je algorytmy, dane i kontekst społeczny, co nadaje im formę władzy. Drugi podrozdział poświęcony jest prywatności predykcyjnej – nie tylko gromadzeniu danych, ale także potencjalnym prognozom dotyczącym przyszłych zachowań jednostki. Mühlhoff wskazuje na ryzyka związane z decyzjami ekonomicznymi, ubezpieczeniowymi czy policyjnymi, wynikającymi z algorytmicznych przewidywań. Trzeci podrozdział bada „kulturę przewidywania”, w której społeczeństwo coraz bardziej ufa predykcjom algorytmicznym, traktując je jako podstawę decyzji i wiedzy. Autor analizuje, jakie wartości epistemologiczne są promowane przez tę kulturę i jakie konsekwencje etyczne niesie podporządkowanie się predykcyjnym maszynom.

Trzecia część, *Władza kontroli*, koncentruje się na cybernetycznych korzeniach AI oraz koncepcjach samoregulacji, sprzężeń zwrotnych i sterowania w nowoczesnych systemach. Mühlhoff omawia brak przejrzystości (*opacity*) w uczeniu maszynowym i predykcyjnej analizie danych, wskazując, że złożoność modeli i „czarne skrzynki” ograniczają odpowiedzialność i utrudniają kontrolę. Analizuje uprzedzenia (*bias*) w systemach AI, które często odzwierciedlają strukturalne nierówności społeczne – rasowe, klasowe czy płciowe. Autor podkreśla, że odpowiedzialność za AI powinna być kolektywna i systemowa, uwzględniając instytucje, regulacje, polityki i wspólnoty jako podmioty odpowiedzialności.

W zakończeniu książki Mühlhoff przedstawia swój manifest etyki AI światowej władzy, postulując głębokie reformy w projektowaniu technologii i instytucjach regulacyjnych. Celem jest, aby AI służyła sprawiedliwości, równości i wspólnemu dobru.

Książka *The Ethics of AI – Power, Critique, Responsibility* jest godna polecenia wszystkim, którzy chcą zrozumieć, jak sztuczna inteligencja wpływa na społeczeństwo, władzę i etykę oraz jakie wyzwania etyczne i odpowiedzialność się z tym wiążą.