

Spis treści

Wprowadzenie: o nadziejach i pułapkach związanych z AI 9

Rozdział 1. Mylne, lecz istotniejsze niż kiedykolwiek 21

I właśnie w owym przeniesieniu idei cybernetycznych na istoty ludzkie koncepcja Wienera zawodzi.

Rozdział 2. Ograniczenia nieprzejrzystości uczących się maszyn 33

Po zainicjowaniu głębokiego uczenia na jakimś sporym zbiorze danych tracisz kontrolę nad procesem, który sam dokonuje reperacji i optymalizacji, a następnie w większości przypadków podaje wynik.

Rozdział 3. O przekazywaniu maszynom ogólnych celów 41

Powinniśmy więc zmierzyć się z perspektywą pojawienia się superinteligentnych maszyn - o z definicji nieprzewidywalnych dla nas działaniach oraz niedoskonale wyznaczonych celach stojących w konflikcie z naszymi - których dążenie do zachowania istnienia w celu osiągnięcia tych celów może okazać się nie do pokonania.

Rozdział 4. Trzecie prawo 55

Każdy system na tyle prosty, żeby go zrozumieć, nie będzie na tyle złożony, by przejawiać inteligencję, a każdy system na tyle złożony, żeby przejawiać inteligencję, będzie zbyt złożony, by go zrozumieć.

Rozdział 5. Co możemy zrobić? 63

Nie potrzebujemy sztucznych świadomych agentów. Potrzebujemy inteligentnych narzędzi.

Rozdział 6. Nieludzki bałagan, w jaki wplątały nas nasze maszyny 75

Dzisiejsza sytuacja jest znacznie bardziej skomplikowana niż w jego prognozach i obawiam się, że radykalnie przerasta jego najgorsze obawy.

Rozdział 7. Jedność inteligencji 85

Przewagi sztucznej inteligencji nad naturalną wydają się permanentne, podczas gdy przewagi naturalnej inteligencji nad sztuczną, mimo że znaczne, sprawiają wrażenie chwilowych.

Rozdział 8. Aspirujemy do czegoś więcej niż trafienie do lamusa 97

Powinniśmy przeanalizować, co mogłoby pójść nie tak z AI, żeby zagwarantować, że tak się nie stanie.

Rozdział 9. Dysydenckie przekazy 109

Ciągły postęp w AI może doprowadzić do zmiany kosmicznych proporcji. Ten niekontrolowany proces przypuszczalnie doprowadzi do zabicia wszystkich ludzi.

Rozdział 10. Proroctwa technologiczne i lekceważenie sprawczej mocy idei 121

Nie istnieje takie prawo złożonych systemów, które mówiłoby, że inteligentni agenci zawsze zmieniają się w bezlitosnych megalomanów.

Rozdział 11. Poza nagrodą i karą 133

Błędne rozumienie ludzkiego myślenia i ludzkich korzeni rodzi błędne wyobrażenia na temat AGI oraz tego, jak można je stworzyć.

Rozdział 12. Istoty ludzkie w oczach maszyn 145

Aby automatyczne systemy inteligentne wyciągały poprawne wnioski na temat pragnień ludzi, muszą dysponować dobrymi modelami ludzkiego poznania.

Rozdział 13. Sztuczna inteligencja a człowiek 153

Pomaganie ludziom zaczyna oznaczać pracę w realnym świecie, w którym trzeba nawiązywać z nimi interakcje i wyciągać na ich temat wnioski.

Rozdział 14. Pościg za gradientem 163

To, że systemom AI zdarza się wpaść w lokalne minimum, wcale nie oddala ich od prawdziwego życia. Ludziom - i przypuszczalnie wszystkim innym formom życia - często zdarza się utknąć w takim minimum.

Rozdział 15. "Informacja" w rozumieniu Wienera, Shannona i naszym 171

Wiele kluczowych argumentów z tej książki zdaje się bliższe wiekowi XIX niż XXI. Zwłaszcza pod tym względem, że chociaż Wiener odwoływał się do wówczas nowej pracy Claude'a Shannona na temat teorii informacji, sprawiał wrażenie, jakby nie w pełni zaakceptował koncepcję informacji złożonej z nieredukowalnych i pozbawionych znaczenia bitów.

Rozdział 16. Skalowanie 181

Chociaż tworzenie maszyn i myślenie maszyn może sprawiać wrażenie niepowiązanych ze sobą trendów, każdy z nich bazuje na przyszłości drugiego.

Rozdział 17. Pierwsze inteligencje maszynowe 191

Hybrydowe superinteligencje w rodzaju państw lub korporacji mają własne cele, a ich działania nie zawsze odzwierciedlają interesy ich twórców.

Rozdział 18. Czy komputery staną się naszymi panami? 201

Nie ma się co dziwić antropomorficznym obawom przed AI, które odzwierciedlają nasze przekonanie o tym, że to inteligencja czyni nas wyjątkowymi.

Rozdział 19. Ludzka strategia 213

W jaki sposób stworzyć dobry ludzko-sztuczny ekosystem, czyli coś, co nie będzie społeczeństwem maszynowym, lecz cyberkulturą, w której będziemy mogli funkcjonować jako ludzie i która będzie sprawiała wrażenie ludzkiej?

Rozdział 20. Uwidacznianie niewidocznego: sztuka i AI 227

Wielu współczesnych artystów bacznie obserwuje te osiągnięcia. Artykułują różne wątpliwości dotyczące obietnic AI i przypominają nam, by nie kojarzyć terminu "sztuczna inteligencja" wyłącznie z pozytywnymi wynikami.

Rozdział 21. AI kontra czterolatek 239

Obserwacja dokonań dzieci może jednak dać programistom wskazówki dotyczące kierunków rozwoju uczenia komputerów.

Rozdział 22. Algorystów marzenie o obiektywności 251

Dzisiaj prawne, etyczne i ekonomiczne wymiary algorytmów są quasi-nieskończone.

Rozdział 23. Prawa maszyn 261

Być może powinniśmy mniej się przejmować kwestią "my kontra oni", a bardziej prawami wszystkich jednostek w obliczu nadciągającego bezprecedensowego zróżnicowania umysłów.

Rozdział 24. Artystyczne zastosowanie istot cybernetycznych 275

Prace cybernetyzujących artystów dotyczą emergujących zachowań życia, które umyka sztucznej inteligencji w jej aktualnej postaci.

Rozdział 25. Sztuczna inteligencja i przyszłość cywilizacji 287

Najbardziej radykalną zmianą z całą pewnością będzie uzyskanie efektywnej ludzkiej nieśmiertelności. Nie wiadomo, czy osiągniemy to biologicznie, czy cyfrowo, ale jest to nieuchronne.