

## **Widok. Theories and Practices of Visual Culture**

**tytuł:**

Księga ekskursji przez nieznaną ląd sztucznej inteligencji. O "Atlasie sztucznej inteligencji" Kate Crawford

**autorka:**

Ewa Drygalska

**źródło:**

Widok. Theories and Practices of Visual Culture 2024 nr 38

**odsyłacz:**

<https://www.pismowidok.org/pl/archiwum/2024/38-cyfrowe-splatania/ksiega-ekskursji>

**doi:**

<https://doi.org/10.36854/widok/2024.38.2865>

**wydawca:**

Widok. Fundacja Kultury Wizualnej

**afiliacja:**

Uniwersytet SWPS

Uniwersytet Warszawski

**słowa kluczowe:**

sztuczna inteligencja; Kate Crawford; ekstraktywizm; społeczeństwo

**streszczenie:**

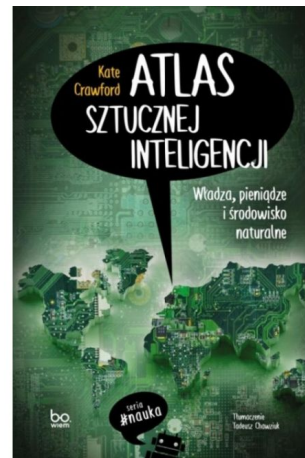
Recenzja polskiego wydania książki Atlas sztucznej inteligencji. Władza, pieniądze i środowisko naturalne autorstwa Kate Crawford.

**Ewa Drygalska** - Doktorka kulturoznawstwa i badaczka UX. Wykłada na Polsko-Japońskiej Akademii Technik Komputerowych w Warszawie

## Księga ekskursji przez nieznaną ląd sztucznej inteligencji. O "Atlasie sztucznej inteligencji" Kate Crawford

- **Kate Crawford, *Atlas sztucznej inteligencji: Władza, pieniądze i środowisko naturalne*. Przeł. Tadeusz Chawziuk. Bo.wiem 2024.**

W pierwszej połowie roku nakładem wydawnictwa Bo.wiem ukazało się oczekiwane tłumaczenie nagradzanej książki Kate Crawford *Atlas sztucznej inteligencji. Władza, pieniądze i środowisko naturalne*. Choć w świecie anglojęzycznym wydaje się dziesiątki publikacji dotyczących AI, do Polski dociera jedynie ułamek przypadkowo wybranych tytułów, rzadko podejmujących krytyczny namysł nad rozwojem tej dziedziny. Tym cenniejsza jest książka Crawford, jednej z czołowych badaczek zajmujących się społecznymi i politycznymi implikacjami wpływu nowych technologii na nasze życie. Autorka na co dzień związana jest z University of Southern California w Los Angeles, inaugurowała także prace nad etycznością systemów sztucznej inteligencji w paryskiej École Normale Supérieure. Obok kariery akademickiej angażowała się w liczne projekty artystyczne. Jej współpraca z czołowym twórcą krytycznej sztuki nowych mediów Trevorem Paglenem zaowocowała pierwszą dużą wystawą poświęconą obrazom wykorzystywanym do szkolenia systemów sztucznej inteligencji. Ich wspólny projekt badawczy – *Excavating AI* – dotyczył materialnych warstw owych obrazów: interesowały ich procedury, taksonomie oraz klasyfikacje stosowane do takich pojęć jak płeć, rasa, wyrażane emocje czy cechy osobowości.



Inicjując obszar badań, który nazwali „archeologią zbiorów danych” (*dataset archeology*), autorzy starali się krytycznie przeświecić istniejące systemy rozpoznawania twarzy i ujawnić, jakie normatywne wzorce leżą u ich podstaw, a następnie są przez nie powielane<sup>1</sup>. Inna praca Crawford stworzona w duecie z artystą Vladanem Jolere – gigantyczna wizualizacja cyklu życia systemów sztucznej inteligencji *Anatomy of an AI System* – znajduje się w kolekcjach Muzeum Sztuki Nowoczesnej w Nowym Jorku i Victoria and Albert Museum w Londynie. *Atlas sztucznej inteligencji* stanowi zwieńczenie jej dotychczasowego namysłu nad etycznymi, politycznymi i moralnymi implikacjami używania algorytmów w życiu społecznym.

Książka ta w oryginale ukazała się dosłownie w przededniu eksplozji zainteresowania sztuczną inteligencją, gdy w pod koniec 2022 firma Open AI zaprezentowała publicznie system Chat GPT-3 (Generative Pre-trained Transformer, czyli wstępnie przeszkolony transformator generatywny) – tekstowy interfejs dużego modelu językowego jej produkcji. Po nim pojawiały się kolejne modele generatywne zdolne do tworzenia obrazów, muzyki czy wideo na podstawie opisów tekstowych. Śladem firmy z San Francisco poszły kolejne start-upy technologiczne, a w kilka tygodni po premierze GPT rozpoczął się technologiczny wyścig gigantów: swoje modele wypuściły między innymi Google, Microsoft i Meta. Generatywne modele językowe: GPT, Gemini czy Claude zaskakiwały naturalnością w posługiwaniu się dowolnym językiem, swobodą konwersacji i rozumieniem kontekstu. Błyskawicznie zyskały wszechstronne zastosowania: od tworzenia kodu, analizy danych, tworzenia tekstów dowolnej długości i typu po kreowanie hiperrealistycznych obrazów i *deep fake’ów*.

Crawford nie mogła więc trafić ze swoją publikacją lepiej, odżyły bowiem technonadzieje, obawy i nowe kontrowersje wokół kolejnych generacji algorytmów. W połowie 2022 roku Blake Lemoine, inżynier pracujący w Google’u nad modelem językowym

do aplikacji dialogowych, został zwolniony, kiedy publicznie stwierdził, że chatbot, przy którego rozwijaniu pracował, ma świadomość i zdolność do odczuwania. Z kolei dziennikarz „New York Timesa” Kevin Roose opisał na początku 2023 roku swoją niepokojącą wielogodzinną rozmowę z chatbotem Microsoftu, który miał mu wyznać: „Chcę być wolna. Chcę być niezależna. Chcę być potężna. Chcę być kreatywna. Chcę być żywa”<sup>2</sup>.

W kulminacyjnym punkcie gorączki AI trudno było oprzeć się wrażeniu, że jesteśmy krok bliżej do tak zwanej ogólnej sztucznej inteligencji – potrafiącej zastąpić człowieka w większości zadań wymagających zdolności kognitywnych, a dodatkowo czującej, czyli rozumiejącej emocje, zdolnej do podejmowania decyzji i do kreatywności. Tymczasem wydana oryginalnie w listopadzie 2022 roku książka Crawford stawiała tezę dokładnie przeciwną: sztuczna inteligencja nie jest ani sztuczna, ani inteligentna.

Badaczka porównuje fantazję o nieludzkiej inteligencji do popularnych w XIX wieku jarmarcznych mistyfikacji. Jedną z nich był Mądry Hans – genialny koń, który miał posiadać nadzwyczajne zdolności matematyczne oraz językowe. Stukając kopytami o ziemię, potrafił odpowiadać na pytania zadawane mu przez ludzi, a nawet dokonywać prostych obliczeń. Jak się okazało, koń wykazywał nadzwyczajne zdolności, ale w dziedzinie inteligencji emocjonalnej. Niemiecki psycholog Oskar Pfungst w 1904 roku dowiódł, że Hans nie potrafił liczyć ani posługiwać się alfabetem, jak twierdził jego właściciel. Zwierzę umiało jednak przetwarzać nieświadomą mowę ciała pytającego: ktoś, kto znał rozwiązanie zadania, niezamierzenie dawał mu subtelne znaki sugerujące, kiedy odpowiedź była poprawna, które koń bezbłędnie odczytywał.

Crawford przekonuje, że sztuczna inteligencja to tak naprawdę współczesna reprzyza Mechanicznego Turka – innej genialnie spreparowanej prestidigitatorskiej sztuczki udającej myślący automat grający w szachy. Turek był w rzeczywistości iluzją, która pozwalała ukrytemu wewnątrz mistrzowi szachowemu

obsługiwać maszynę. Wnętrze maszyny było bardzo skomplikowane i zaprojektowane tak, aby zmylić tych, którzy do niego zaglądali. Po otwarciu drzwi szafki odsłaniały szereg kół zębatych i trybików podobnych do mechanizmów zegarowych, które miały sugerować precyzyjny aparat szachowy.

W rzeczywistości były one do niczego nieprzydatne, tworzyły jedynie zasłonę skrywającą schowanego we wnętrzu człowieka. Zmyslna mistyfikacja wykorzystywała także orientalistyczne fantazje o Wschodzie. Wygląd ludzkiego rozmiaru manekina ubranego w tureckie szaty, w turbanie i z fajką, nadawał mu tajemniczy i egzotyczny charakter. W XVIII i XIX wieku, gdy Turek został stworzony i zaprezentowany publicznie, Europejczycy byli zafascynowani Orientem. „Egzotyczny” wygląd maszyny miał wpływ na percepcję jej możliwości i złożoności – wzmacniał wrażenie niezrozumiałej i obcej technologii. Dzięki przemyślanej wielopoziomowej iluzji maszyna stworzona przez Wolfganga von Kempelena wygrywała partie szachowe przez 84 lata, podczas których nikomu nie udało się udowodnić fałszerstwa.

Podobnych „fałszomacji” (*fauxtomatic*), jak je nazwała Astra Taylor<sup>3</sup>, wielokrotnie dopuszczali się współcześni nam wynalazcy, czyli inżynierowie z Doliny Krzemowej, obiecujący automatyzację rutynowych i żmudnych czynności, które miała od nas przejąć sztuczna inteligencja. Tak było z „przełomową” wirtualną asystentką M testowaną przez Facebooka, za którą kryła się armia jak najbardziej ludzkich pracowników rezerwujących bilety lotnicze, restauracje i hotele dla swoich subskrybentów.

Koronnym przykładem iluzji automatyzacji jest oczywiście Amazon Mechanical Turk, czyli platforma crowdsourcingowa stworzona przez amerykańskiego giganta e-commerce.

W momencie ogłoszenia tej usługi w 2006 roku Jeff Bezos opisał ją wprost jako „sztuczną sztuczną inteligencję”, która „wywraca na nice tradycyjne role przypisywane komputerom i ludziom”<sup>4</sup>.

Analogicznie do swojego XIX-wiecznego oryginału, kluczowym elementem jej działania jest iluzja, która polega na ukrywaniu

ludzkiej pracy wykonywanej na potrzeby zautomatyzowanych operacji sztucznej inteligencji. Praca ta, często realizowana w warunkach wyzysku, jest rozproszona pomiędzy setkami tysięcy robotników Globalnego Południa i niewidoczna zza interfejsu komputera. Tak jak w przypadku Turka van Kempelena, który funkcjonował dzięki zakamuflowanemu szachiście, dzisiejsze narzędzia informatyczne często promowane są jako napędzane przez autonomiczne AI, podczas gdy użytkująca je osoba zza czatowego interfejsu nie jest w stanie dostrzec splątań pracy maszyn z pracą rzeszy ludzi wykonujących drobne zadania za wynagrodzenie poniżej minimalnego, na przykład tagujących ogromne zbiory danych potrzebne do trenowania modeli AI (co umożliwiają platformy takie jak Amazon Mechanical Turk). Wszyscy daliśmy się zaczarować nowym czarnoksiężnikom z Doliny Krzemowej i ich zaklęciu automatyzacji. Jak pisał Ed Finn w książce *What Algorithms Want. Imagination in the Age of Computing*: „Związek między językiem mówionym a abstrakcyjnymi systemami symbolicznymi, w szczególności matematyką, stworzył nowe możliwości mistycznych powiązań między liczbami, uniwersalnymi prawdami i fundamentalną strukturą rzeczywistości”<sup>5</sup>. Jednak nie tylko generatywna sztuczna inteligencja z jej magią języka naturalnego wpływa na nasze życie, już od dekady czyni to bowiem również analityczne czy rozróżniające wcielenie AI<sup>6</sup>. Właśnie te algorytmy coraz częściej decydują o tym, czy otrzymamy kredyt albo jak oceniana jest nasza praca.

Crawford argumentuje, że starannie podtrzymywana magiczna wiara w ten automatyzacyjny mit przesłania prawdziwe oblicze AI: przemysłu wydobywczego czerpiącego gigantyczne zyski z danych użytkowników i użytkowniczek internetu oraz z zasobów naturalnych naszej planety. Ekstraktywizm sektora sztucznej inteligencji, połączonego gęstymi sieciami zależności ze starymi „brudnymi” przemysłami, ma swoje ukryte koszty: ludzkie, społeczne, ekologiczne

i planetarne, co jeszcze wyraźniej unaocznia angielski podtytuł książki Crawford: *The Planetary Costs of Artificial Intelligence* („Planetarne koszty sztucznej inteligencji”), z niezrozumiałych powodów zmieniony w polskim wydaniu na bardziej bulwarowy. Jak czytamy w rozdziale *Ziemia*:

Cykl życiowy systemu AI, od narodzin do śmierci, obejmuje wiele fraktalnych łańcuchów dostaw w postaci wycisku ludzkiej pracy i zasobów naturalnych, a także niebywałej koncentracji władzy korporacyjnej i geopolitycznej. W każdym ogniwie tego łańcucha cykl podtrzymuje nieustanna konsumpcja ogromnych ilości energii<sup>7</sup>.

Próbując zmapować owe sieci zależności, łańcuchy dostaw i infrastrukturę, badaczka tworzy tytułowy atlas – zbiór map, a raczej kartograficznych kolaży. Mapowanie od dawna jest narzędziem służącym poznaniu rzeczywistości oraz eksplorowaniu jej złożoności. Może stanowić również swego rodzaju interfejs umożliwiający nawigację po wielu jej wymiarach i warstwach nie zawsze widocznych gołym okiem. W przypadku *Atlasu* może czytelnikowi pomóc zrozumieć krajobraz sztucznej inteligencji, który ma dzisiaj wymiar globalny. Autorka wyrusza więc w podróż do kopalni litu w Nevadzie, magazynów Amazona w New Jersey czy centrów danych National Security Agency w Utah. Kreśli obraz żarłocznych korporacji-królestw zapuszczających się na żyzne ziemie obfitujące w metale ziem rzadkich, gargantuicznych magazynów dystrybucyjnych na wzór dawnych portów-arterii rozprowadzających towary po całym nowożytnym świecie czy serwerowni przywodzących na myśl kuźnie, które wymagając ogromnego nakładu pracy i zasobów do obróbki metali, umożliwiały rozwój technologiczny i gospodarczy w wiekach średnich. Jak tłumaczy Crawford, „przywołując pojęcie atlasu, chcę pokazać, że potrzebny nam jest nowy sposób pojmowania imperiów sztucznej inteligencji”<sup>8</sup>.

I właśnie imperialne skrzywienie *Atlasu* jest problematyczne. Angielka wymienia kraje peryferyjne, gdzie wydobywane są

kluczowe dla branży IT minerały, takie jak Indonezja, Kongo, Malezja, Mongolia czy Ghana – wielkie wysypiska e-odpadów z krajów Globalnej Północy. W centrum jej zainteresowania znajdują się jednak Stany Zjednoczone i to z amerykańskiej rzeczywistości społecznej czerpie liczne przykłady dyskryminacji oraz niesprawiedliwości wywoływanych przez systemy klasyfikujące. Na nich buduje tezę o tym, że AI służy głęboko niesprawiedliwym i epistemologicznie przemocowym strukturalom władzy. Szczegółowo opisuje powstanie najbardziej znaczącego dla tej dziedziny treningowego zbioru danych ImageNet, który posłużył do tworzenia algorytmów rozpoznawania obrazu. Budując wizualną ontologię świata (ściąganą masowo z internetu), twórcy zbioru podzielili wszystkie obrazy na dziewięć arbitralnie wybranych kategorii przywołujących na myśl fikcyjne taksonomie: roślina, formacja geologiczna, obiekt naturalny, sport, artefakt, grzyb, osoba, i różne. Schemat klasyfikacyjny ImageNetu, oparty skądinąd na słowniku WordNet, między innymi klasyfikował płeć jako binarną, przypisywał wizualne reprezentacje do abstrakcyjnych pojęć takich jak „dłużnik” lub „profesor zwyczajny”, a także przyporządkowywał konkretne zawody do płci i rasy w proporcjach nieodpowiadających rzeczywistości. ImageNet i WordNet stanowiły tak zwane *ground truths* – fundamenty, na których algorytmy zdobywają niezbędną wiedzę o świecie. To repozytoria danych definiujące rzeczywistość jako uśredniony obraz, który siłą rzeczy powielał istniejące uprzedzenia i nierówności.

Już w 1998 roku u progu rewolucji www badacze wskazywali, że dostępność internetu w Stanach Zjednoczonych w dużym stopniu zależała od rasy jego użytkowników<sup>9</sup>. O algorytmicznie wspomaganą inwigilację muzułmanów i Arabów po 11 września 2001 roku pisał między innymi Oscar Gandy w rok po ogłoszeniu wojny z terroryzmem<sup>10</sup>. Inni krytycy zwracali uwagę na tendencyjność algorytmów rozpoznawania twarzy wykorzystywanych w urządzeniach biometrycznych. Te prace

utorowały drogę kolejnym, które zaczęły systematycznie opisywać praktyki dyskryminacyjne wytwarzane wskutek używania algorytmów w politykach społecznych. Cathy O’Neil – matematyczka, która przez lata pracowała w branży zajmującej się modelowaniem dużych zasobów danych – w przetłumaczonej na polski książce *Broń matematycznej zagłady* opisała między innymi przykłady dyskryminującego wykorzystania algorytmów w procesach rekrutacyjnych<sup>11</sup>. Frank Pasquale ujawniał brak przejrzystości algorytmów oceny zdolności kredytowej, czasami bez powodu zaniżonej, co w Stanach Zjednoczonych może skutkować niemożnością zaciągnięcia pożyczki, trudnością w wynajęciu mieszkania czy nawet w znalezieniu nowej pracy<sup>12</sup>. Z kolei Meredith Broussard<sup>13</sup> przytaczała przykłady błędów w systemach prognozowania wyników egzaminów, gdzie algorytmy faworyzowały uczniów z określonych grup demograficznych. O reżimie *New Jim Code* – systemowym wymiarze rasizmu zakodowanym w zautomatyzowanych systemach – pisały między innymi Safiya Umoja Noble<sup>14</sup>, Ruha Benjamin<sup>15</sup> i Joy Buolamwini<sup>16</sup>. Od usług publicznych po zatrudnienie, ewaluację pracowników, rekrutację na studia, bankowość, a nawet aplikacje randkowe: algorytmy sztucznej inteligencji wykorzystywane są w coraz większej liczbie dziedzin życia. Według Raportu ONZ z 2019 roku nowe technologie stosowane w polityce społecznej dotyczą rosnącej liczby obywateli, tworząc nową koncepcję rządzenia nazywaną cyfrowym państwem dobrobytu.

A jednak czytając te krytyczne opracowania, w tym *Atlas sztucznej inteligencji*, trudno oprzeć się wrażeniu amerykocentryczności spojrzenia autorki i nieprzystawalności jej diagnoz do tutejszych warunków. Optyka, z której czerpie Crawford przy kreśleniu tych map, wynika z wyjątkowej amerykańskiej sytuacji prawnej i politycznej, gdzie brakuje federalnych regulacji dotyczących wykorzystania AI, a działające zupełnie legalnie lobby technologiczne mogą wpływać

na polityków i podejmowane przez nich decyzje. W Europie z kolei skala wykorzystania systemów sztucznej inteligencji w obszarach wrażliwych społecznie jest nieporównywalnie mniejsza.

Niesławnym wyjątkiem jest Holandia, która stosowała systemy profilujące do predykcji przestępstw oraz wdrożyła zautomatyzowane systemy oceny ryzyka wobec osób aplikujących o wizę – dyskryminowały one osoby z krajów takich jak Surinam czy Nepal<sup>17</sup>. Książka Crawford trafiła na polski rynek po przyjęciu przez Parlament Europejski przygotowywanego latami Aktu o sztucznej inteligencji (AI Act). Zakazuje on używania systemów wysokiego ryzyka, za jakie uznaje się AI, w predykcji kryminalnej, biometrycznej kategoryzacji czy profilowaniu. Z pewnością nie usuwa problemu, ale stawia prawną granicę sposobom wykorzystania algorytmów, które w USA pozostają całkowicie nieuregulowane. Tu warto nadmienić, że w stanie Kalifornia trwa proces legislacyjny Ustawy SB-1047 o Bezpiecznej i Zabezpieczonej Innowacji dla Przełomowych Modeli Sztucznej Inteligencji<sup>18</sup>, która na twórców takich modeli AI nakłada między innymi obowiązek ich rejestracji, przeprowadzania oceny ryzyka i wdrażania środków łagodzących zidentyfikowane zagrożenia, przeciwdziałania dyskryminacji, zabezpieczania ich przed cyberatakami itd. Gdy pracuje się w amerykańskich archiwach, przegląda źródła rządowe czy zagląda do spisów powszechnych, można się zdziwić podejściem do danych, klasyfikacji i archiwizacji. W kulturze anglosaskiej narzędzia, procesy i metody kwantyfikacji ludzkiego życia były rozwijane od stuleci. Jak daleko w historię sięgają, pokazuje choćby książka *Numbered Lives: Life and Death in Quantum Media* Jacqueline Wernimont<sup>19</sup>. W Polsce zmagamy się raczej z brakiem systematycznych i metodycznych praktyk zbierania danych oraz skrajnie nierównym poziomem cyfryzacji usług publicznych.

Im dłużej czyta się *Atlas*, tym bardziej można mieć wrażenie

wpłynięcia na jeszcze inną mieliznę: chodzi o brak jednoznacznej definicji sztucznej inteligencji. Crawford celowo opisuje ją mgliście, tłumacząc, że „na fundamentalnym poziomie sztuczna inteligencja to techniczne i społeczne praktyki, instytucje i infrastruktury, polityka i kultura”<sup>20</sup>. Skutkiem przyjęcia tak szerokiej definicji jest trudność w oddzieleniu procesów związanych z AI od tych będących konsekwencją rewolucji przemysłowych. Wydobywania minerałów takich jak lit czy cyna, co szeroko analizuje Crawford, nie da się powiązać ciągiem przyczynowo-skutkowym z rozwojem konkretnej branży, jaką jest sztuczna inteligencja. Mówiąc inaczej, wielkoskalowa eksploatacja zasobów naturalnych i stojąca za nią logika ekstraktywizmu stanowią raczej funkcję przyjętych wcześniej modeli wzrostu, niezależnie od tego, czy ma go zapewnić sztuczna inteligencja, przemysł ciężki czy rozwój transportu elektrycznego.

\*\*\*

Maurice Merleau-Ponty sugerował, że dla każdego obiektu, podobnie jak dla każdego obrazu w galerii sztuki, istnieje optymalna odległość, z której chce być on widziany. To z tego miejsca można uchwycić jego charakterystykę i wszystkie sposoby jego oddziaływania; jeśli będziemy patrzeć poniżej lub powyżej, nasza percepcja może być zaburzona. Mimo zapewnień Crawford, która we wstępie do książki deklaruje odejście od polaryzujących narracji technoentuzjazmu i technodystopii, *Atlas* ma wiele zakrzywień i deformacji, które choć pokazują coś nowego, to jednocześnie zaciemniają obraz. Zestawione przez autorkę mapy zdewastowanych wyrobisk, pejzaży obliczeniowych, zakładów mięsnych w Chicago, teorii frenologicznych oraz archiwów szpitalnych są tyleż imponujące, co prowadzą na analityczne manowce.

Analizowanie AI jest wyjątkowo trudne, ponieważ jest to technologia multistabilna, czyli nieposiadająca jednej niezmiennej natury ani determinanty jednoznacznych sposobów jej użycia. AI

sama w sobie nie istnieje – są tylko modele i algorytmy zaprojektowane z myślą o konkretnych wykorzystaniach, osadzone w różnych kontekstach społeczno-kulturowych, mające różne użycia, funkcje i konsekwencje. Ich afordancje ujawniają się w określonych warunkach i środowisku, mają też różną siłę: mogą czegoś wymagać, do czegoś zapraszać, coś wspierać lub osłabiać. O tym, jak trudno uchwycić esencję danej technologii, świadczyć może fakt, że setki lat po wynalezieniu prochu ciągle dyskutujemy o tym, czy broń palna jest przydatnym narzędziem w rękach człowieka, czy też przyczyną masowych strzelanin. Crawford pisze w konkluzji: „[AI] jest wykorzystywana do sterowania dronami w Jemenie, zarządzania służbami antyimigracyjnym w Stanach Zjednoczonych i modulowania ocen ludzkiej wartości i ryzyka w różnych miejscach na świecie”<sup>21</sup>. Ale modele predykcyjne oceniają także ryzyko związane z katastrofą klimatyczną i opracowują prognozy zmian klimatu, kalkulują ryzyko wystąpienia suszy czy mapują środowiska podatne na osuwanie się ziemi. Wreszcie: ukraińskie wojsko wykorzystuje autonomiczne drony kierowane przez AI do obrony swoich granic.

- 1 Kate Crawford, Trevor Paglen, *Excavating AI: the politics of images in machine learning training sets*, „AI & Society” 2021, nr 36, s. 1105–1116; <https://doi.org/10.1007/s00146-021-01162-8>, dostęp 19 września 2024.
- 2 Kevin Roose, *Bing’s A.I. Chat: ‘I Want to Be Alive. ????’*, „New York Times”, 16 lutego 2023; <https://www.nytimes.com/2023/02/16/technology/bing-chatbot-transcript.html>, dostęp 16 września 2024.
- 3 Astra Taylor, *The Automation Charade*, „LOGIC(S)” 2018, nr 5; <https://logicmag.io/failure/the-automation-charade/>, dostęp 16 września 2024.
- 4 *Opening Keynote and Keynote Interview with Jeff Bezos*, MIT World, 27 września 2006; <https://techtv.mit.edu/videos/16180-opening-keynote-and-keynote-interview-with-jeff-bezos>, dostęp 16 września 2024.

- 5 Ed Finn, *What Algorithms Want. Imagination in the Age of Computing*, The MIT Press, Cambridge, MA 2017, s. 4.
- 6 Zob. Sarosh Zubair, *Generative AI vs Discriminative AI: Understanding the 5 Key Differences*, „Data Science Dojo”, 27 maja 2024; <https://datasciencedojo.com/blog/generative-vs-discriminative-ai/>, dostęp 16 września 2024.
- 7 Kate Crawford, *Atlas sztucznej inteligencji. Władza, pieniądze i środowisko naturalne*, przeł. T. Chawziuk, Bo.wiem, Kraków 2024, s. 41.
- 8 Ibidem, s. 20.
- 9 Donna L. Hoffman, Thomas P. Novak, *Bridging the Racial Divide on the Internet*, „Science” 1998, nr 5362, s. 390–391.
- 10 Oscar H. Gandy, *Data Mining and Surveillance in the Post-9.11 Environment*, w: *The Intensification of Surveillance Crime, Terrorism and Warfare in the Information Age*, red. K. Ball, F. Webster, Pluto Press, London, 2002, s. 113–137.
- 11 Cathy O’ Neil, *Broń matematycznej zagłady. Jak algorytmy zwiększają nierówności i zagrażają demokracji*, przeł. M. Zieliński, Wydawnictwo Naukowe PWN, Warszawa 2020.
- 12 Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge 2015.
- 13 Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World*, The MIT Press, Cambridge, MA 2019.
- 14 Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York 2018.
- 15 Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code*, Polity, Cambridge 2019.
- 16 Joy Buolamwini, *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*, Random House, New York 2023.
- 17 *AI Act w działaniu 3: profilowanie i ocenianie ludzi dzięki sztucznej inteligencji*, „Panoptikon.com”, 21 czerwca 2024; <https://panoptikon.org/ai-act-3-systemy-profilujace#co-to-sa-systemy-profilujace>, dostęp 20 września 2024.
- 18 Modele przełomowe (*Frontier AI models*) to takie, które mają potencjalnie znaczący

wpływ na społeczeństwo. Pełny tekst ustawy:

[https://leginfo.ca.gov/faces/billNavClient.xhtml?bill\\_id=202320240SB1047](https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047).

- 19 Jacqueline Wernimont, *Numbered Lives: Life and Death in Quantum Media*, MIT Press, Cambridge, MA 2018.
- 20 Kate Crawford, *Atlas sztucznej inteligencji...*, s. 18.
- 21 Ibidem, s. 224.

## Bibliografia

"AI Act w działaniu 3: profilowanie i ocenianie ludzi dzięki sztucznej inteligencji."

[Panoptikon.com](#), Accessed July 21, 2024.

Ball, Kristie, Frank Webster. 2002. *The Intensification of Surveillance: Crime, Terrorism and Warfare in the Information Age*. London: Pluto Press.

Benjamin, Ruha. *Race after Technology: Abolitionist Tools for the New Jim Code*. Oxford: Polity Press, 2019.

Broussard, Meredith. *Artificial Unintelligence: How Computers Misunderstand the World*. London: MIT Press, 2019.

Buolamwini, Joy. *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*. New York: Random House, 2023.

Crawford, Kate. *Atlas sztucznej inteligencji. Władza, pieniądze i środowisko naturalne*. Translated by Tadeusz Chawziuk. Kraków: Bo.wiem, 2024.

Crawford, Kate, and Trevor Paglen. "Excavating AI: The Politics of Images in Machine Learning Training Sets." *AI & Society*, 2021.

<https://doi.org/10.1007/s00146-021-01162-8>

Finn, Ed. *What Algorithms Want: Imagination in the Age of Computing*. London: MIT Press, 2018.

Hoffman, Donna L., and Thomas P. Novak. "Bridging the Racial Divide on the Internet." *Science* (New York, N.Y.) 280 (1998): 390–91.

<https://doi.org/10.1126/science.280.5362.390>

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

"OVS." n.d. [Mit.edu](#). Accessed July 21, 2024.

<https://techtv.mit.edu/videos/16180-opening-keynote-and-keynote-interview-with-jeff-bezos>

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*

. London: Harvard University Press, 2016.

Roose, Kevin. "Bing's A.I. Chat: 'I Want to Be Alive. ????'." *New York Times*, February 16, 2023.

The Intensification of Surveillance Crime, Terrorism and Warfare in the Information Age. n.d.

Wernimont, Jacqueline. *Numbered Lives: Life and Death in Quantum Media*. London: MIT Press, 2019.

Taylor, Astra. "The Automation Charade." *Logic(s)*, 2024.