Cykle rozwoju sztucznej inteligencji w perspektywie książki Raising AI Chiraga Shaha

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje historię sztucznej inteligencji jako cykl przełomów i korekt, opierając się na perspektywie Chiraga Shaha z książki „Raising AI". Autor wskazuje na powtarzalność faz entuzjazmu i rozczarowania (tzw. zimy AI), które kształtowały przejścia między paradygmatem symbolicznym a koneksjonizmem. Tekst podkreśla, że rozwój AI nie jest liniowy, lecz stanowi proces przesuwania środka ciężkości między różnymi metodami reprezentacji wiedzy – od wczesnych modeli Turinga i McCulloch-Pitts, przez systemy ekspertowe, aż po współczesne modele generatywne. Kluczowym wnioskiem jest rozróżnienie między sukcesem w wyspecjalizowanych zadaniach a osiągnięciem sztucznej inteligencji ogólnej (AGI), co pozwala uniknąć błędnego utożsamiania mocy obliczeniowej z ludzkimi kompetencjami poznawczymi.

The article analyzes the history of artificial intelligence as a cycle of breakthroughs and corrections, drawing on the perspective of Chirag Shah from the book "Raising AI." The author points to the repeatability of phases of enthusiasm and disappointment (so-called AI winters), which shaped the transitions between the symbolic paradigm and connectionism. The text emphasizes that AI development is not linear but is a process of shifting the center of gravity between different methods of knowledge representation—from early Turing and McCulloch-Pitts models, through expert systems, to modern generative models. The key conclusion is the distinction between success in specialized tasks and the achievement of artificial general intelligence (AGI), which allows one to avoid mistakenly equating computing power with human cognitive competencies.

Artykuł analizuje historię sztucznej inteligencji nie jako liniowy postęp, lecz jako cykl gwałtownych przełomów i bolesnych korekt, znanych jako „zimy AI". Autor stawia tezę, że rzeczywista ewolucja tej dziedziny nie wynika z prostej akumulacji mocy obliczeniowej czy zwiększania liczby parametrów modeli, lecz z krytycznego zrozumienia granic generalizacji oraz odpowiedzialnego dopasowania innowacji do struktur społecznych i etycznych. Tekst prowadzi czytelnika od wczesnych paradygmatów symbolicznych, przez systemy ekspertowe i uczenie głębokie, aż po współczesne modele generatywne i agentowe, wskazując na powtarzalny mechanizm: sukces w wąskiej domenie jest błędnie utożsamiany z uniwersalną inteligencją (AGI), co prowadzi do spekulacyjnych baniek finansowych i późniejszych rozczarowań. Kluczowym przesłaniem jest stwierdzenie, że technologia AI nie funkcjonuje w próżni – jej realna wartość zależy od komplementarności zasobów infrastrukturalnych, organizacyjnych oraz ludzkiej roztropności (phronesis), której żadna maszyna nie jest w stanie zastąpić.

SŁOWA KLUCZOWE

cykle rozwoju sztucznej inteligencji | Raising AI Chirag Shah | zimy sztucznej inteligencji | paradygmat symboliczny | koneksjonizm | systemy ekspertowe | modele generatywne | sztuczna inteligencja ogólna AGI | eksplozja kombinatoryczna | wąskie gardło pozyskiwania wiedzy | ironia automatyzacji | phronesis w AI | statystyczne uczenie maszynowe | deep learning | determinizm technologiczny

Historia AI jako cykl przełomów i korekt

Historia sztucznej inteligencji: między naukowym przełomem a kolejną zimą. Cykliczność paradygmatów, ekonomia technologicznych obietnic i granice poznawczej ekspansji maszyn

Historia sztucznej inteligencji jest przykładem rozwoju wiedzy naukowej, w którym okresy przełomowych odkryć przeplatają się z fazami rozczarowania, instytucjonalnego sceptycyzmu i radykalnego przewartościowywania dotychczasowych założeń. Zjawisko to nie ogranicza się do zwykłej cykliczności zainteresowania technologią; obejmuje ono jednoczesne przemiany teorii poznania, architektury systemów obliczeniowych, organizacji badań, sposobów finansowania innowacji oraz społecznego rozumienia inteligencji.

W kolejnych dekadach badacze wielokrotnie osiągali rezultaty wcześniej uznawane za trudne do zdobycia, by następnie odkryć, że sukces w konkretnym zadaniu nie prowadzi automatycznie do opanowania znacznie szerszej klasy problemów. Powtarzalność tego doświadczenia jest kluczowa dla współczesnych dyskusji o sztucznej inteligencji ogólnej, ponieważ pozwala odróżnić rzeczywistą ekspansję zdolności obliczeniowych od nieuzasadnionego utożsamiania osiągnięć wyspecjalizowanych systemów z całością ludzkich kompetencji poznawczych. Historia AI to zatem nie tylko kronika coraz doskonalszych maszyn, lecz także zapis kolejnych korekt ludzkich wyobrażeń o naturze myślenia, uczenia się i rozumienia świata.

Chirag Shah w książce **Raising AI: How Artificial Intelligence Is Shaping Us-and How We Shape It** proponuje spojrzenie na rozwój tej dziedziny przez pryzmat następujących po sobie fal zainteresowania, wzrostu oczekiwań i okresów stagnacji określanych mianem zim sztucznej inteligencji. Shah wyróżnia pięć szerokich epok: narodziny symbolicznej AI i pierwszych sieci neuronowych, rozwój systemów ekspertowych, ekspansję metod statystycznych, przełom uczenia głębokiego oraz współczesną erę modeli generatywnych. Podział ten stanowi użyteczną periodyzację analityczną, jednak nie należy go interpretować jako pięciu całkowicie odrębnych etapów, w których jedna szkoła nieodwracalnie zastępowała poprzednią.

Badania symboliczne, probabilistyczne, neuronowe i ewolucyjne często rozwijały się równolegle, a wiele współczesnych architektur łączy dorobek kilku tradycji. Istotą historycznej dynamiki nie jest więc prosta sukcesja zwycięskich teorii, lecz przemieszczanie się środka ciężkości badań między odmiennymi sposobami reprezentowania wiedzy, organizowania obliczeń i definiowania inteligentnego działania. Każda nowa generacja technologiczna przejmowała część osiągnięć poprzedniczki, równocześnie ujawniając nowe ograniczenia, których nie zawsze można było usunąć przez ilościowe zwiększanie zasobów.

Początki nowoczesnej sztucznej inteligencji wyrastały z szerszej rewolucji naukowej pierwszej połowy XX wieku, obejmującej matematykę, logikę formalną, teorię informacji, neurofizjologię oraz cybernetykę. Alan Turing w pracy **On Computable Numbers, with an Application to the Entscheidungsproblem** z 1936 roku sformułował abstrakcyjny model obliczeń, który umożliwił precyzyjne rozważanie granic mechanicznego wykonywania procedur symbolicznych. Jego późniejszy artykuł **Computing Machinery and Intelligence** z 1950 roku przeniósł dyskusję z problemu formalnej obliczalności na zagadnienie zachowań uznawanych za inteligentne. Zaproponowana gra w imitację nie była jednak dowodem, że wszystkie aspekty ludzkiej świadomości są algorytmicznie odtwarzalne; stanowiła raczej metodologiczną próbę zastąpienia niejednoznacznego pytania o wewnętrzną naturę myślenia analizą obserwowalnych kompetencji komunikacyjnych.

Równolegle Warren McCulloch i Walter Pitts przedstawili w 1943 roku matematyczny model uproszczonych neuronów, ukazując możliwość reprezentowania pewnych procesów logicznych za pomocą sieci elementów progowych. Norbert Wiener rozwijał cybernetykę jako naukę o sterowaniu i komunikacji w maszynach oraz organizmach żywych, natomiast Claude Shannon sformułował matematyczną teorię informacji. Powstająca AI odziedziczyła zatem od początku dwa częściowo konkurencyjne obrazy inteligencji: jako formalnego przetwarzania symboli oraz jako złożonego procesu adaptacyjnego, realizowanego przez sieć wzajemnie oddziałujących elementów.

Za instytucjonalny początek sztucznej inteligencji jako wyodrębnionej dyscypliny przyjmuje się letni projekt badawczy w Dartmouth College w 1956 roku, przygotowany na podstawie propozycji sporządzonej w sierpniu 1955 roku przez Johna McCarthy'ego, Marvina Minsky'ego, Nathaniela Rochestera i Claude'a Shannona. Dokument zawierał śmiałe założenie, że każdy aspekt uczenia się oraz innych właściwości inteligencji można co do zasady opisać wystarczająco precyzyjnie, aby maszyna mogła go symulować. Program obejmował badania nad językiem, tworzeniem pojęć, rozwiązywaniem problemów i samodoskonaleniem systemów. Uderzająca z dzisiejszej perspektywy była skala optymizmu organizatorów, którzy zakładali możliwość uzyskania znaczących rezultatów podczas stosunkowo krótkiego, kilkutygodniowego przedsięwzięcia.

Iluzja pełnej inteligencji w paradygmacie symbolicznym

Nie wynikało to wyłącznie z "naukowej naiwności". W połowie XX wieku formalizacja rozumowania, elektroniczne obliczenia i rozwój technologii cyfrowych zdawały się tworzyć fundamenty rzeczywistego przełomu, choć granice ich zastosowań pozostawały słabo rozpoznane. Problem polegał na utożsamieniu możliwości formalnego zapisu operacji poznawczych z realną zdolnością odtworzenia pełnego bogactwa ludzkiej inteligencji w ograniczonym czasie i przy ówczesnych środkach technicznych. Wczesny paradygmat symboliczny opierał się na przekonaniu, że inteligentne działanie można zrekonstruować jako manipulowanie odpowiednio zdefiniowanymi reprezentacjami wiedzy. Allen Newell i Herbert A. Simon rozwijali programy rozwiązujące zadania logiczne oraz matematyczne, co doprowadziło do sformułowania hipotezy fizycznego systemu symboli. Według niej określona organizacja operacji symbolicznych stanowi niezbędny i wystarczający warunek ogólnego inteligentnego działania. Zaletą tego podejścia była przejrzystość struktur rozumowania, w których przesłanki, reguły i rezultaty pozostawały formalnie określone. Logika klasyczna, rachunek predykatów, reprezentacje semantyczne i procedury przeszukiwania przestrzeni stanów pozwalały konstruować systemy rozwiązujące zadania, które dotąd wymagały pracy specjalistów. Jednocześnie założenie, że świat można przedstawić za pomocą skończonego zbioru symboli i reguł, rodziło trudności związane z wieloznacznością, niepełnością danych oraz

ogromną liczbą możliwych kombinacji. Formalna poprawność rozumowania nie gwarantowała trafności reprezentacji rzeczywistości, a skuteczność programu w wąskiej domenie nie była dowodem posiadania uniwersalnej inteligencji.

Zjawisko określane mianem "eksplozji kombinatorycznej" ujawniło zasadniczą barierę ówczesnych metod. W wielu problemach liczba możliwych stanów rośnie wraz ze wzrostem liczby zmiennych tak gwałtownie, że systematyczne przeszukiwanie wszystkich kombinacji staje się praktycznie niewykonalne. Nie jest to jedynie kwestia zbyt wolnych procesorów; niektóre klasy zadań cechują się trudnością obliczeniową, której nie eliminuje prosty wzrost zasobów sprzętowych. Badania nad teorią złożoności, rozwijane m.in. przez Stephena Cooka i Richarda Karpa, pozwoliły odróżnić zadania rozwiązywalne efektywnie od problemów, dla których nie istnieją algorytmy zapewniające korzystną skalowalność w ogólnym przypadku. W praktyce wymusiło to stosowanie heurystyk, przybliżeń oraz ograniczeń przestrzeni poszukiwań i wiedzy specyficznej dla danej dziedziny. Inteligencja okazała się zatem czymś znacznie bardziej wymagającym niż formalna zdolność wyprowadzenia poprawnego wyniku z kompletnych przesłanek. W rzeczywistym świecie przesłanki bywają niepełne, kontekst jest zmienny, a liczba potencjalnych konsekwencji przekracza możliwości bezpośredniego obliczenia.

Koneksjonizm i pierwsze kryzysy oczekiwań

Równolegle rozwijała się tradycja koneksjonistyczna, inspirowana uproszczonymi modelami funkcjonowania układu nerwowego. W drugiej połowie lat pięćdziesiątych Frank Rosenblatt opracował perceptron – model uczącego się klasyfikatora, którego parametry można było dostosowywać na podstawie przykładów. Publiczne prezentacje tego rozwiązania budziły ogromne zainteresowanie; sugerowały bowiem możliwość budowy maszyn, które nie tylko wykonują zaprogramowane instrukcje, ale modyfikują własne reakcje pod wpływem doświadczenia treningowego.

W 1969 roku Marvin Minsky i Seymour Papert opublikowali monografię **Perceptrons**, w której przeanalizowali matematyczne ograniczenia pewnych klas sieci, zwłaszcza jednowarstwowych modeli progowych. Klasycznym przykładem był problem funkcji XOR, której nie da się przedstawić za pomocą pojedynczej liniowej granicy decyzyjnej. Z czasem krytykę tę zaczęto niekiedy interpretować jako dowód zasadniczej niemożności rozwijania sieci neuronowych. Byłoby to jednak nadmierne uproszczenie, gdyż ograniczenia konkretnych architektur nie oznaczały, że wszystkie możliwe sieci są niezdolne do reprezentowania funkcji nieliniowych. Historycznie istotny okazał się

raczej fakt, że dynamikę nurtów koneksjonizmu osłabił brak efektywnych metod trenowania złożonych struktur, ograniczenia infrastrukturalne oraz zmiana priorytetów badawczych.

Wczesne rozczarowania nie były wynikiem jednego eksperymentu czy pojedynczej publikacji. Raport amerykańskiego komitetu ALPAC z 1966 roku zakwestionował ekonomiczną i praktyczną skuteczność części projektów automatycznego tłumaczenia, co doprowadziło do ograniczenia finansowania tej dziedziny w USA. Z kolei raport Jamesa Lighthilla z 1973 roku podważył szersze oczekiwania wobec badań nad AI w Wielkiej Brytanii, wskazując na przepaść między deklarowanymi możliwościami a realnymi rezultatami, szczególnie w zadaniach o rosnącej złożoności.

Wpływ tego dokumentu na decyzje finansowe pozostaje przedmiotem sporów historyków, gdyż ograniczenia budżetowe wynikały również z niezależnych uwarunkowań politycznych i gospodarczych. Nie ulega jednak wątpliwości, że pierwsza połowa lat siedemdziesiątych przyniosła gwałtowny spadek entuzjazmu. Pierwsza zima sztucznej inteligencji nie oznaczała całkowitego zaprzestania prac, lecz kryzys zaufania do wybranych programów, trudności w pozyskiwaniu funduszy i konieczność przedefiniowania celów.

Z perspektywy filozofii nauki okres ten można interpretować przez pryzmat koncepcji programów badawczych Imre Lakatosa. Według niego naukowcy rozwijają zespoły założeń teoretycznych i metodologicznych, których wartość zależy od zdolności do generowania nowych, empirycznie owocnych rezultatów. Program postępowy prowadzi do odkryć; program degenerujący ogranicza się natomiast do modyfikacji chroniących stare założenia przed konfrontacją z rzeczywistością.

W historii AI kluczowe było zatem nie tylko to, czy model rozwiązywał konkretne zadanie, ale czy jego architektura pozwalała na systematyczne rozszerzanie kompetencji. Wczesne programy symboliczne osiągały imponujące wyniki w ustrukturyzowanych środowiskach, lecz zawodziły w świecie niejednoznacznym, dynamicznym i słabo sformalizowanym. Kolejne próby dopisywania reguł nie zawsze przynosiły jakościową poprawę, gdyż problemem pozostawała sama forma reprezentacji wiedzy. Odpowiedzią na te ograniczenia stały się systemy ekspertowe, które zyskały na znaczeniu w latach siedemdziesiątych i osiemdziesiątych. Zamiast dążyć do stworzenia uniwersalnej inteligencji, badacze skupili się na odtwarzaniu konkretnych fragmentów wiedzy specjalistycznej.

Przykładem był system DENDRAL, rozwijany od połowy lat sześćdziesiątych przez Edwarda Feigenbauma i Joshuę Lederberga, który wspomagał identyfikację struktur związków chemicznych. Podobnie MYCIN, tworzony na Uniwersytecie Stanforda w latach siedemdziesiątych, wykorzystywał reguły wiedzy medycznej do rozpoznawania zakażeń i doboru antybiotykoterapii.

Wąskie gardło wiedzy i druga zima AI

Rozwiązania te udowodniły wartość koncentracji na wąskich obszarach, w których wiedzę ekspertów można było reprezentować stosunkowo precyzyjnie. System nie musiał rozumieć całej rzeczywistości, by dostarczać użytecznych rekomendacji w domenie o odpowiednio uporządkowanej strukturze problemowej. Sukcesy systemów ekspertowych napędziły rozwój inżynierii wiedzy, która traktowała pozyskiwanie, porządkowanie i formalizowanie kompetencji specjalistów jako odrębny proces projektowy.

Trudności tego procesu obnażyły jednak ograniczenia ówczesnego rozumienia ludzkiej wiedzy. Michael Polanyi w koncepcji wiedzy ukrytej podkreślał, że człowiek potrafi wykonywać wiele złożonych czynności, nie dysponując pełnym i świadomym opisem wszystkich reguł swojego działania. Lekarz może rozpoznawać wzorce kliniczne na podstawie doświadczenia, inżynier dostrzegać istotne anomalie w pracy instalacji, a nauczyciel oceniać sytuację pedagogiczną dzięki kompetencjom trudnym do całkowitej werbalizacji. System ekspertowy wymagał natomiast wydobycia konkretnych reguł, ich uporządkowania oraz zdefiniowania relacji między wyjątkami. W ten sposób powstało tzw. wąskie gardło pozyskiwania wiedzy (**knowledge acquisition bottleneck**), gdyż przekształcenie praktycznej biegłości eksperta w kompletną bazę reguł okazywało się zadaniem niezwykle pracochłonnym. Co więcej, wiedza ulegała zmianom, procedury organizacyjne ewoluowały, a nowe przypadki wymuszały nieustanną aktualizację systemów.

Wraz z komercjalizacją systemów ekspertowych uruchomiono kolejny mechanizm instytucjonalny prowadzący do nadmiernego optymizmu. Przedsiębiorstwa inwestowały w wyspecjalizowane oprogramowanie i sprzęt zoptymalizowany pod język LISP, a rozwój inteligentnych systemów przedstawiano jako fundament nowej generacji infrastruktury gospodarczej. W 1982 roku Japonia uruchomiła program komputerów piątej generacji, łącząc aspiracje naukowe z polityką przemysłową i geopolityczną rywalizacją o technologiczną przewagę.

Choć projekt ten skutecznie zmobilizował zasoby badawcze, wiele jego najbardziej ambitnych celów nie zostało osiągniętych w pierwotnie zakładanej formie. Pod koniec lat osiemdziesiątych rynek maszyn LISP uległ presji tańszych i wydajniejszych urządzeń ogólnego przeznaczenia, a wysokie koszty utrzymania baz reguł obniżyły atrakcyjność wielu wdrożeń. Doprowadziło to do kolejnego okresu spadku finansowania i zainteresowania, znanego jako druga zima AI.

Ewolucja metod zamiast całkowitych porażek

Kryzys ten nie był jednak dowodem bezużyteczności całej inżynierii wiedzy, gdyż wiele jej koncepcji przetrwało w systemach reguł biznesowych, silnikach wnioskowania, systemach wspomagania decyzji oraz późniejszych rozwiązaniach hybrydowych. Historyczna interpretacja tych zdarzeń wymaga ostrożności wobec popularnej narracji, według której kolejne paradygmaty AI ponosiły całkowitą porażkę i były zastępowane przez rozwiązania zupełnie nowe. W rzeczywistości znacząca część osiągnięć przenikała do praktyki informatycznej pod innymi nazwami, a wypracowane metody znajdowały zastosowanie w coraz bardziej wyspecjalizowanych systemach. Przeszukiwanie przestrzeni stanów, reprezentacja wiedzy, rozumowanie probabilistyczne, wnioskowanie regułowe i optymalizacja kombinatoryczna nie zniknęły wraz ze zmianami trendów. Przeciwnie – wiele współczesnych systemów wykorzystuje je jako komponenty architektur łączących uczenie maszynowe z formalnymi procedurami kontroli. Zjawisko to można interpretować jako ewolucję technologiczną, w której innowacje nie zawsze eliminują wcześniejsze rozwiązania, lecz reorganizują zakres ich użyteczności.

Pojęcie zimy AI opisuje przede wszystkim określoną konfigurację oczekiwań, finansowania i społecznej oceny wyników, a nie moment całkowitego zatrzymania postępu naukowego. Lata dziewięćdziesiąte i pierwsza dekada XXI wieku przyniosły wzrost znaczenia statystycznych metod uczenia maszynowego, które pozwalały rozpoznawać prawidłowości bez konieczności ręcznego kodowania każdej reguły decyzyjnej. Rozwijały się drzewa decyzyjne, maszyny wektorów nośnych, modele probabilistyczne, sieci bayesowskie, metody zespołowe i różnorodne techniki optymalizacji. Judea Pearl wniósł istotny wkład w formalizację rozumowania probabilistycznego i przyczynowego, wykazując, że niepewność można reprezentować w strukturach matematycznych pozwalających analizować relacje między zmiennymi.

Vladimir Vapnik współtworzył teorię statystycznego uczenia się, dostarczając podstaw do badań nad generalizacją modeli poza próbę treningową. Równocześnie rozwijano algorytmy propagacji wstecznej błędu oraz architektury wielowarstwowych sieci neuronowych. W 1997 roku Sepp Hochreiter i Jürgen Schmidhuber przedstawili architekturę LSTM, która łagodziła problemy związane z uczeniem zależności długookresowych w sieciach rekurencyjnych. Choć nie usuwała ona wszystkich ograniczeń uczenia sekwencyjnego, stanowiła ważny krok w rozwoju modeli przetwarzających mowę, tekst i szeregi czasowe.

Przemiana ta oznaczała zmianę głównego wektora badań. W tradycji symbolicznej centralnym problemem było ustalenie, jakie reguły należy przekazać maszynie, aby mogła poprawnie rozumować. W podejściu statystycznym kluczowe stało się rozpoznanie, jakie zależności system

może wyuczyć z danych i w jakich warunkach będą one przenoszalne na nowe przypadki. Pojawił się problem nadmiernego dopasowania (**overfitting**), polegający na wysokiej skuteczności wobec danych treningowych przy niewystarczającej zdolności generalizacji. W odpowiedzi badacze zaczęli rozwijać procedury walidacji krzyżowej, regularyzacji, selekcji cech i oceny niezależnych zbiorów testowych. Wraz ze wzrostem znaczenia danych ujawnił się również problem ich reprezentatywności, co dziś odzwierciedla dyskusja o stronniczości algorytmicznej. Model uczący się z obserwacji nie przyswaja automatycznie prawdy o świecie, lecz rekonstruuje zależności dostępne w konkretnym zbiorze danych. Jeśli zbiór ten jest niepełny, historycznie skrzywiony lub niereprezentatywny, zwiększanie sprawności obliczeniowej nie musi prowadzić do poprawy trafności społecznej.

Symbolicznym wydarzeniem tej epoki było zwycięstwo komputera IBM Deep Blue nad Garrim Kasparowem w 1997 roku. Sukces ten wywołał ogromne zainteresowanie, gdyż szachy od dawna traktowano jako dziedzinę wymagającą złożonego planowania i strategicznego myślenia. Deep Blue wykorzystywał jednak przede wszystkim wyspecjalizowane procedury przeszukiwania, funkcje oceny pozycji i znaczną moc obliczeniową, a nie współczesną architekturę głębokiego uczenia. Jego zwycięstwo dowiodło, że maszyna może osiągnąć najwyższy poziom sprawności w konkretnym zadaniu intelektualnym, ale nie było to dowodem posiadania przez nią ogólnego rozumienia rzeczywistości. Wydarzenie to ujawniło problem filozoficzny dotyczący kryteriów przypisywania inteligencji: czynność uznawana za szczególnie wymagającą poznawczo może zostać zautomatyzowana metodą, która nie odtwarza ludzkiego sposobu jej wykonywania. W konsekwencji społeczna definicja inteligencji przesuwa się wraz ze zmianą technicznych możliwości maszyn.

Ważnym osiągnięciem następnej dekady był rozwój pojazdów autonomicznych, w tym sukces zespołu Stanford Racing Team i pojazdu Stanley w zawodach DARPA Grand Challenge w 2005 roku. Sterowanie samochodem w nieznanym terenie wymagało połączenia percepcji otoczenia, planowania trasy, oceny ryzyka i kontroli układów wykonawczych. W przeciwieństwie do szachów, rzeczywiste środowisko drogowe nie oferuje kompletnej informacji o wszystkich możliwych stanach, a decyzje systemu zależą od jakości czujników, pogody i zachowań innych uczestników ruchu. Rozwój pojazdów autonomicznych unaocznił znaczenie ucieleśnionego działania, czyli funkcjonowania inteligencji w bezpośredniej relacji z fizycznym środowiskiem.

Przełomy AI jako efekt komplementarności zasobów

Autor pokazał również, że sukces w warunkach ograniczonego eksperymentu nie musi automatycznie przekładać się na niezawodność w rzeczywistych sytuacjach. Ta różnica między demonstracją możliwości a pełną dojrzałością wdrożeniową będzie powracać w kolejnych falach AI, niezależnie od rodzaju wykorzystywanej architektury. Nowy etap rozwoju nastąpił w latach dwutysięcznych, a szczególnie wyraźnie ujawnił się po 2010 roku, gdy wzrost dostępności danych cyfrowych, rozwój procesorów graficznych i postępy w metodach optymalizacji stworzyły warunki do szerokiego wykorzystania głębokich sieci neuronowych. Kluczowym wydarzeniem był sukces modelu AlexNet, opracowanego przez Aleksa Krizhevsky'ego, Ilyę Sutskevera i Geoffreya Hintona, który w 2012 roku zwyciężył w konkursie ImageNet Large Scale Visual Recognition Challenge.

Sieć konwolucyjna uzyskała wówczas wskaźnik błędu **top-5** na poziomie około 15,3%, podczas gdy drugi najlepszy wynik wyniósł około 26,2%. Tak wyraźna dysproporcja była mocnym sygnałem, że połączenie obszernych zbiorów danych, odpowiedniej architektury sieciowej i obliczeń równoległych pozwala osiągać rezultaty przekraczające możliwości wcześniejszych metod rozpoznawania obrazów. Sukces ten nie wynikał jednak z odkrycia jednej „magicznej reguły" inteligencji; był raczej efektem wieloletniej akumulacji wiedzy matematycznej, rozwoju oprogramowania, dostępności danych treningowych oraz odpowiedniej infrastruktury sprzętowej.

W tym miejscu autor odwołuje się do ekonomicznej koncepcji komplementarności czynników produkcji. Wzrost możliwości danej technologii zależy często od równoczesnego rozwoju kilku zasobów, których efektywności nie można rozpatrywać w izolacji. Algorytm uczenia głębokiego może być teoretycznie znany przez lata, lecz jego praktyczne zastosowanie na dużą skalę pozostaje ograniczone przez koszty obliczeń, dostępność pamięci, jakość danych czy możliwości organizacyjne.

Dopiero przekroczenie określonych progów infrastrukturalnych pozwala urzeczywistnić istniejące już koncepcje. Zjawisko to odpowiada szerszym prawidłowościom historii gospodarczej, w której wynalazek techniczny nie musi natychmiast transformować produkcji. Maszyny parowe, elektryfikacja czy komputery wymagały zmian w organizacji przedsiębiorstw, kwalifikacjach pracowników i systemach infrastrukturalnych, zanim ich pełny potencjał mógł zostać wykorzystany. Rewolucja głębokiego uczenia nie była wyjątkiem od tej zasady.

Kolejnym przełomem okazały się sukcesy uczenia ze wzmocnieniem, w którym system doskonali strategię działania na podstawie informacji o uzyskiwanych rezultatach. W marcu 2016 roku AlphaGo, opracowany przez DeepMind, pokonał Lee Sedola w meczu go wynikiem cztery do

jednego. Szczególne zainteresowanie wzbudził trzydziesty siódmy ruch drugiej partii, początkowo uznawany przez wielu obserwatorów za niezwykły i trudny do zinterpretowania.

AlphaGo łączył głębokie sieci neuronowe z procedurami przeszukiwania i uczenia poprzez rozgrywanie kolejnych partii, osiągając poziom strategicznej sprawności wcześniej uważany za nieosiągalny. Sukces ten nie dowodził powstania świadomości maszynowej ani ogólnej inteligencji, lecz pokazywał, że system potrafi odkrywać rozwiązania odbiegające od utrwalonych wzorców ludzkich ekspertów. Kluczowa była tu możliwość generowania nowych strategii w obrębie precyzyjnie określonego środowiska decyzyjnego. Jednocześnie ujawniło to zasadnicze ograniczenie tej analogii: świat społeczny, prawny i gospodarczy nie jest grą o kompletnych regułach, stabilnej funkcji nagrody i jednoznacznie określonym warunku zwycięstwa. Następną transformację zapoczątkował rozwój architektury Transformer, przedstawionej w 2017 roku przez Ashisha Vaswaniego i współautorów w pracy **Attention Is All You Need**.

Skalowanie modeli a granica rozumowania

Wykorzystanie mechanizmu uwagi umożliwiło modelowanie zależności między elementami sekwencji przy znacznym ograniczeniu potrzeby stosowania tradycyjnych struktur rekurencyjnych. Architektura ta pozwoliła efektywniej równoleglić wiele operacji podczas treningu, tworząc podstawy do skalowania modeli językowych. Nie oznaczało to całkowitego usunięcia sekwencyjności ze wszystkich procesów – w typowych modelach autoregresyjnych generowanie kolejnych tokenów nadal odbywa się bowiem w zależności od wcześniejszego kontekstu. Istotą przełomu była raczej nowa organizacja obliczeń oraz możliwość efektywnego wykorzystania coraz większych zbiorów danych i zasobów sprzętowych.

W późniejszych latach rozwój modeli takich jak BERT, GPT-3 i systemów multimodalnych wykazał, że wspólne mechanizmy reprezentacji mogą być adaptowane do szerokiej klasy zadań związanych z językiem, obrazem i innymi rodzajami informacji. Publiczne udostępnienie ChatGPT 30 listopada 2022 roku nadało temu procesowi charakter przełomu społecznego; zaawansowane modele językowe zaczęły być wykorzystywane przez szerokie grupy użytkowników bez konieczności posiadania specjalistycznej wiedzy programistycznej. Interfejs konwersacyjny radykalnie obniżył barierę dostępu do technologii, która wcześniej była dostępna głównie za pośrednictwem wyspecjalizowanych aplikacji, interfejsów programistycznych i zespołów inżynieryjnych. Użytkownik mógł opisywać swoje potrzeby językiem naturalnym: formułować pytania, prosić o przekształcenie tekstu, analizować dokumenty czy korzystać ze wsparcia w rozwiązywaniu problemów. Technologia rozpoczęła tym samym proces przenikania do codziennych praktyk

poznawczych, edukacyjnych, administracyjnych i gospodarczych. Jej znaczenie nie wynikało wyłącznie z poprawy parametrów modeli, lecz także z upowszechnienia sposobu interakcji, który uczynił obliczeniową złożoność niewidoczną dla odbiorcy. Człowiek nie musiał już uczyć się rozbudowanego języka poleceń maszyny; coraz częściej to system był projektowany tak, aby interpretować i obsługiwać wypowiedzi formułowane w języku człowieka.

W badaniach nad wielkoskalowymi modelami językowymi szczególne znaczenie uzyskały empiryczne prawa skalowania, analizowane między innymi przez Jareda Kaplana i współautorów w 2020 roku. Wykazano wówczas, że dla określonych rodzin modeli i warunków treningowych miary błędu językowego zmieniają się w przewidywalny sposób wraz ze wzrostem liczby parametrów, zasobów obliczeniowych i wielkości zbioru danych. Dalsze badania Jordana Hoffmanna i współautorów z 2022 roku ujawniły znaczenie odpowiedniego równoważenia rozmiaru modelu i ilości danych treningowych przy określonym budżecie obliczeniowym. Wyniki te doprowadziły do rewizji części wcześniejszych założeń dotyczących optymalnej strategii skalowania, choć nie uprawniają one do twierdzenia, że każdy wzrost zasobów automatycznie prowadzi do jakościowego zwiększenia wszystkich kompetencji poznawczych.

Prawa skalowania są empirycznymi zależnościami obserwowanymi dla określonych architektur i zadań; ich ekstrapolacja poza zweryfikowany obszar wymaga ostrożności metodologicznej. Szczególnie problematyczne jest przechodzenie od poprawy miary predykcyjnej do twierdzeń dotyczących świadomości, głębokiego rozumienia, autonomii moralnej czy zdolności niezawodnego działania w nieograniczonym środowisku. Różnica między zdolnością do przewidywania kolejnego elementu sekwencji a kompetencją przyczynowego wyjaśniania świata stanowi jeden z najważniejszych problemów obecnego paradygmatu. Nie jest naukowo uzasadnione twierdzenie, że modele trenowane predykcyjnie są z definicji całkowicie niezdolne do rozumowania przyczynowego, gdyż ich zachowania można oceniać na podstawie konkretnych zadań wymagających posługiwania się zależnościami przyczynowo-skutkowymi. Jednocześnie skuteczność w rozwiązywaniu określonych problemów tekstowych nie dowodzi posiadania stabilnego i uniwersalnego modelu przyczynowości. Judea Pearl wyróżniał poziomy rozumowania obejmujące obserwację, interwencję oraz kontrfaktyczne rozważanie alternatywnych przebiegów zdarzeń.

Współczesne systemy mogą wykazywać znaczną sprawność w zadaniach odpowiadających tym kategoriom, ale ich niezawodność zależy od warunków testowych, reprezentacji danych i możliwości weryfikacji wyników. Z tego powodu centralnym zagadnieniem pozostaje odróżnianie powierzchniowej poprawności odpowiedzi od jej odporności na zmianę kontekstu, niepełność informacji oraz sytuacje nieznane wcześniej systemowi.

Od generowania treści do agentowej sprawczości i ryzyka

Współczesna ekspansja AI prowadzi do powstawania nowych architektur agentowych, w których model językowy zostaje połączony z narzędziami wykonawczymi, pamięcią operacyjną, planowaniem wieloetapowym oraz możliwością korzystania z zewnętrznych systemów informatycznych. W takim środowisku AI przestaje ograniczać się do generowania odpowiedzi; zaczyna aktywnie uczestniczyć w wyszukiwaniu informacji, analizowaniu dokumentów, programowaniu, obsłudze aplikacji i wykonywaniu złożonych sekwencji czynności. Zdolność ta zwiększa potencjał gospodarczy technologii, lecz jednocześnie zmienia charakter ryzyka.

Błąd modelu udzielającego wyłącznie informacji można zauważyć przed podjęciem działania, podczas gdy błąd agenta z uprawnieniami operacyjnymi może bezpośrednio prowadzić do zmiany danych, wykonania transakcji lub uruchomienia niepożądanego procesu. Problem bezpieczeństwa przesuwa się zatem od samej jakości generowanych treści ku architekturze uprawnień, kontroli procesów, możliwości przerwania działania i organizacji odpowiedzialności za skutki wykonanych operacji.

Aktualne dane Stanford Human-Centered Artificial Intelligence, przedstawione w raporcie **AI Index 2026**, potwierdzają znaczną dynamikę rozwoju AI, lecz jednocześnie ujawniają nierównomierność postępu między dziedzinami. Wskaźniki skuteczności najlepszych modeli w szeregu testów poznawczych poprawiały się w 2025 roku tak szybko, że niektóre stosowane dotąd benchmarki przestały wystarczająco różnicować ich zdolności. Jednocześnie w teście OSWorld, obejmującym wykonywanie zadań komputerowych przez agentów, odnotowano wzrost skuteczności do około 66,3%, co wciąż oznacza znaczną liczbę niepowodzeń w warunkach kontrolowanej oceny. W zadaniach robotycznych różnica między sukcesami osiąganymi w symulacjach a realnymi pracami domowymi była jeszcze silniejsza. Wyniki te nie pozwalają ani na ogłoszenie stagnacji całej dziedziny, ani na utożsamienie wysokiej sprawności w testach poznawczych z niezawodną zdolnością samodzielnego funkcjonowania w dowolnym środowisku. Wskazują raczej na nierówny charakter technologicznej dojrzałości, której poszczególne składniki rozwijają się w odmiennym tempie.

Raport Stanforda ujawnia również skalę gospodarczej mobilizacji wokół AI. Według zastosowanej w nim metodologii globalne inwestycje korporacyjne związane ze sztuczną inteligencją osiągnęły w 2025 roku około 581,7 miliarda dolarów, obejmując różne kategorie aktywności, w tym inwestycje prywatne i przejęcia. Inwestycje prywatne w Stanach Zjednoczonych wyniosły około 285,9 miliarda

dolarów, znacznie przewyższając wartości wykazywane dla Chin i Europy w porównywanych kategoriach.

Wyniki te nie oznaczają jednak, że każda wydana kwota stanowi bezpośredni koszt trenowania modeli językowych ani że wartość inwestycji jest tożsama z uzyskanym dobrobytem społecznym. Pokazują natomiast intensywność koncentracji kapitału w sektorze, którego długookresowa rentowność pozostaje uzależniona od przyszłych zastosowań, cen usług, kosztów infrastruktury oraz zdolności przedsiębiorstw do uzyskania trwałej przewagi konkurencyjnej. Z perspektywy ekonomii finansowej kluczowe staje się więc rozróżnienie pomiędzy rozwojem realnych zdolności technologicznych a wyceną oczekiwanych przyszłych przychodów.

Mechanizm cykliczności inwestycji technologicznych autor książki wyjaśnia za pomocą teorii innowacji Josepha A. Schumpetera. W jego ujęciu rozwój gospodarczy nie jest wyłącznie procesem stopniowego wzrostu efektywności istniejących struktur, lecz obejmuje fale nowych kombinacji czynników produkcji, prowadzące do powstawania rynków, reorganizacji przedsiębiorstw i wypierania wcześniejszych rozwiązań. Proces twórczej destrukcji może zwiększać produktywność całego systemu gospodarczego, ale jednocześnie powoduje straty określonych grup, przedsiębiorstw i regionów. W przypadku AI szczególnie istotne jest to, że innowacja oddziałuje zarówno na procesy produkcyjne, jak i na samą organizację wiedzy, pracy intelektualnej oraz zarządzania. Przedsiębiorstwa mogą automatyzować część zadań, tworzyć nowe produkty, zmieniać strukturę zatrudnienia i reorganizować relacje z klientami.

Nie każda taka zmiana zwiększa jednak dobrobyt w jednakowym stopniu. Ostateczny rezultat zależy od rodzaju zastosowania, kosztów dostosowawczych, możliwości zdobywania nowych kompetencji oraz sposobu podziału korzyści między pracowników, właścicieli kapitału i konsumentów.

Mechanizmy spekulacji a realna wartość AI

Koncepcję długookresowych fal technologicznych rozwijała Carlota Perez, analizując powiązania między rewolucjami technicznymi, inwestycjami finansowymi a przemianami instytucjonalnymi. W jej ujęciu przełomowe technologie przechodzą fazę intensywnej instalacji infrastruktury – towarzyszy jej spekulacja, gwałtowny wzrost wycen i nadmierne oczekiwania – po której następuje okres reorganizacji, prowadzący do szerokiego wykorzystania innowacji w gospodarce. Zastosowanie tej perspektywy do AI pozwala dostrzec potencjalną rozbieżność między realną wartością technologii a rynkową wyceną spółek uczestniczących w jej rozwoju. Przeinwestowanie infrastruktury nie musi jednak oznaczać, że sama technologia jest bezużyteczna; historia

gospodarcza zna sektory, w których finansowa korekta następowała równolegle z utrwalaniem fundamentów przyszłego wzrostu produktywności. Analogia do przełomu internetowego jest tu użyteczna, choć wymaga ostrożności – AI charakteryzuje się bowiem odmienną strukturą kosztów, ogromnym zapotrzebowaniem na energię i półprzewodniki oraz specyficznym charakterem potencjalnej substytucji pracy.

Równie pomocna okazuje się teoria niestabilności finansowej Hymana P. Minsky'ego, według której długie okresy optymizmu prowadzą do stopniowego wzrostu skłonności do ryzyka i przyjmowania coraz bardziej spekulacyjnych modeli finansowania. W tym ujęciu początkowo uzasadnione nadzieje na wzrost dochodów sprzyjają zwiększaniu inwestycji, co z czasem wymusza wyceny oparte na coraz bardziej ambitnych scenariuszach przyszłej rentowności. Jeśli oczekiwane korzyści nie materializują się w zakładanym czasie, następuje korekta wycen, ograniczenie finansowania i rewizja strategii przedsiębiorstw. Nie oznacza to jednak, że sektor AI znajduje się obecnie w identycznej sytuacji jak rynek kredytowy analizowany przez Minsky'ego, ani że istnieją podstawy do ogłoszenia nieuchronnego kryzysu.

Teoria ta pozwala natomiast wskazać mechanizm, w którym rosnące oczekiwania inwestorów zwiększają presję na firmy, by prezentowały coraz śmielsze cele technologiczne i krótsze terminy ich realizacji. Szczególnym przejawem tego procesu jest zjawisko znane z ekonomii behawioralnej jako narracyjna dynamika oczekiwań. Robert J. Shiller wykazywał, że historie o przyszłym wzroście gospodarczym i możliwościach inwestycyjnych rozprzestrzeniają się społecznie, wpływając na decyzje finansowe niezależnie od chłodnej oceny danych fundamentalnych. W przypadku AI opowieść o rychłym stworzeniu systemu zdolnego do wykonywania większości pracy intelektualnej może oddziaływać na wyceny spółek, decyzje o budowie centrów danych, strukturę zatrudnienia oraz politykę konkurencyjną.

Nie jest przy tym konieczne, aby wszyscy uczestnicy rynku wierzyli w pełną prawdziwość takich prognoz. Wystarczy, że liczą na wzrost zainteresowania innych inwestorów i możliwość zysku z wcześniejszego zajęcia pozycji. John Maynard Keynes już w swoich analizach rynków finansowych wskazywał na mechanizmy, w których uczestnicy starają się przewidzieć nie tyle rzeczywistą wartość aktywów, co przyszłe przekonania innych podmiotów.

Rozdźwięk między kapitałem a realną produktywnością AI

W sektorze AI zależność ta może zaostrzać rywalizację o infrastrukturę, talenty naukowe i rozpoznawalność rynkową. Odrębnego wyjaśnienia wymaga jednak różnica między spekulacyjną falą kapitałową a trwałym postępem technologicznym. Wzrost wycen przedsiębiorstw nie dowodzi bowiem, że ich produkty będą generować adekwatne przepływy pieniężne; z drugiej strony spadek tych wycen nie musi oznaczać zaniku kompetencji technicznych. Firma może osiągnąć przełom badawczy i jednocześnie ponieść stratę finansową, jeśli nie potrafi przekształcić rezultatu w rentowny produkt. Z kolei organizacja może generować wysokie przychody dzięki korzystnej strukturze rynku, mimo że oferowana przez nią technologia nie stanowi przełomu naukowego.

Rozróżnienie tych poziomów jest kluczowe przy ocenie obecnej ery AI, gdyż znaczna część inwestycji dotyczy infrastruktury o wysokich kosztach początkowych, której ekonomiczne wykorzystanie wymaga utrzymania odpowiednio dużej skali zapotrzebowania. Pojawia się ryzyko błędnego wnioskowania: tempo wzrostu wydatków na centra danych zostaje utożsamione z tempem wzrostu społecznie użytecznej inteligencji. Raport Międzynarodowej Agencji Energetycznej **Energy and AI** z 2025 roku pozwala osadzić te rozważania w materialnej strukturze gospodarki. Według scenariusza bazowego IEA globalne zużycie energii elektrycznej przez centra danych może wzrosnąć do około 945 TWh w 2030 roku, co odpowiadałoby nieco mniej niż trzem procentom prognozowanego światowego zapotrzebowania na prąd.

Jest to prognoza obarczona istotną niepewnością i dotycząca całego sektora centrów danych, a nie wyłącznie trenowania modeli AI. Obejmuje ona konsekwencje rozwoju infrastruktury obliczeniowej, chłodzenia, przesyłu danych i usług cyfrowych. Znaczenie gospodarcze tych nakładów wynika nie tylko z ich skali, lecz także z lokalnej koncentracji zapotrzebowania na energię, wodę, grunt i sieci przesyłowe. Odpowiedzialność za rozwój AI obejmuje zatem również planowanie infrastrukturalne, bezpieczeństwo energetyczne, emisje środowiskowe, koszty alternatywnego wykorzystania kapitału i długookresową stabilność systemów gospodarczych. Ekonomiczna ocena korzyści z AI wymaga uwzględnienia problemu określanego przez Erika Brynjolfssona, Daniela Rocka i Chada Syversona mianem krzywej produktywności w kształcie litery J.

Autorzy wskazują, że technologie ogólnego zastosowania wymagają znaczących inwestycji komplementarnych w kapitał organizacyjny, nowe procesy, kompetencje pracowników i modele biznesowe. W początkowym okresie wdrażania przedsiębiorstwa ponoszą koszty reorganizacji, podczas gdy część tworzonej wartości nie znajduje natychmiastowego odzwierciedlenia w

standardowych statystykach produktywności. Dopiero po pewnym czasie może nastąpić szersze wykorzystanie nowej infrastruktury i ujawnienie korzyści gospodarczych. Zjawisko to stanowi ważny kontrargument wobec przekonania, że brak natychmiastowej poprawy wskaźników makroekonomicznych dowodzi niepowodzenia technologii. Jednocześnie nie można traktować tej teorii jako uniwersalnego usprawiedliwienia dla dowolnie wysokich wydatków inwestycyjnych. Model krzywej J wskazuje na możliwość opóźnionych korzyści, a nie na gwarancję ich wystąpienia.

Bardziej ostrożną perspektywę prezentuje Daron Acemoglu w pracy **The Simple Macroeconomics of AI** z 2024 roku. Ekonomista ten analizuje potencjalny wpływ zastosowań sztucznej inteligencji na produktywność, uwzględniając udział zadań podatnych na opłacalną automatyzację oraz charakter kosztów związanych z wdrażaniem technologii. Jego szacunki wskazują na znacznie mniejsze korzyści makroekonomiczne w horyzoncie dekady niż niektóre optymistyczne prognozy sektora technologicznego. W określonych wariantach analizy przewidywany wzrost całkowitej produktywności czynników produkcji pozostaje poniżej jednego procenta w ciągu dziesięciu lat.

Nie jest to jednak uniwersalne proroctwo, lecz rezultat modelu o jasno określonych założeniach dotyczących rodzaju automatyzowanych zadań, tempa wdrożeń i ekonomicznej opłacalności zastosowań. Porównanie stanowisk Acemoglu i Brynjolfssona pokazuje, że rozbieżności między prognozami nie muszą wynikać z odmiennych ocen sprawności technicznej modeli. Mogą być następstwem różnego rozumienia procesu dyfuzji innowacji, wartości inwestycji komplementarnych i zakresu reorganizacji możliwej do przeprowadzenia przez firmy. Z perspektywy nauk o zarządzaniu szczególnie istotne jest rozróżnienie automatyzacji pojedynczego zadania od reorganizacji całego procesu produkcyjnego. Model językowy może skracać czas redagowania dokumentów, lecz organizacja nie uzyska proporcjonalnego wzrostu produktywności, jeżeli rezultat wymaga czasochłonnej kontroli, korekty, integracji z systemami lub wielokrotnego zatwierdzania. Podobnie automatyczne wspomaganie programowania może przyspieszać tworzenie kodu, ale jego rzeczywista wartość zależy od jakości oprogramowania, bezpieczeństwa, kosztów utrzymania i zgodności z wymaganiami użytkowników. Efekt jednostkowego eksperymentu nie musi więc przekładać się liniowo na wskaźniki przedsiębiorstwa ani gospodarki.

Powstaje problem komplementarności organizacyjnej, w którym technologia staje się użyteczna dopiero po zmianie podziału pracy, odpowiedzialności, kompetencji i standardów jakości. W tym kontekście szczególnie niebezpieczne jest traktowanie redukcji zatrudnienia jako automatycznego dowodu sukcesu wdrożenia, ponieważ krótkoterminowa oszczędność finansowa może prowadzić do utraty wiedzy organizacyjnej i wzrostu kosztów przyszłych błędów.

Psychologiczne pułapki i płynność definicji inteligencji

Mechanizm cykliczności AI ma również wymiar psychologiczny, wynikający z poznawczych ograniczeń w ludzkiej ocenie niepewności. Daniel Kahneman i Amos Tversky wykazali, że posługujemy się heurystykami, które choć przyspieszają podejmowanie decyzji, prowadzą jednocześnie do systematycznych zniekształceń ocen probabilistycznych. W przypadku technologii kluczowe znaczenie mają tu heurystyka dostępności, nadmierna pewność siebie oraz tendencja do bezkrytycznego ekstrapolowania aktualnych trendów. Spektakularny sukces systemu AI w konkretnym zadaniu często prowadzi do nieuzasadnionego rozszerzenia oceny jego kompetencji na znacznie szersze obszary. Jeśli model osiąga wynik porównywalny z ekspertem w wybranym benchmarku, odbiorcy skłonni są zakładać, że dysponuje on również wiedzą praktyczną, zdolnością do rozpoznawania niepewności czy odpowiedzialnością za skutki swoich działań. Powstaje wówczas efekt poznawczego przeniesienia, w którym sukces lokalny staje się podstawą prognozy globalnej.

To właśnie ten mechanizm sprzyja historycznej powtarzalności nadmiernych oczekiwań wobec sztucznej inteligencji. Nie można przy tym pominąć problemu płynnych kryteriów oceny inteligencji. Wraz z postępem technicznym czynności niegdyś uznawane za wyjątkowo wymagające zaczynają być traktowane jako standardowe funkcje oprogramowania. Automatyczne obliczenia, wyszukiwanie informacji, rozpoznawanie pisma, tłumaczenie tekstów czy gra w szachy przestały pełnić rolę dowodów inteligencji maszynowej, choć wcześniej stanowiły jej najbardziej widoczne symbole. Zjawisko to określa się mianem efektu AI – polega ono na przesuwaniu granic pojęcia sztucznej inteligencji w miarę upowszechniania się konkretnych zdolności technicznych.

Nie jest to jedynie anegdotyczny przejaw ludzkiej niekonsekwencji, lecz ujawnienie istotnego problemu metodologicznego: pojęcie inteligencji bywa definiowane poprzez zdolności, których mechanizmu w danym momencie nie potrafimy odtworzyć. Po ich zautomatyzowaniu okazuje się, że sama realizacja zadania nie rozstrzyga o posiadaniu całości pozostałych kompetencji poznawczych.

Współczesna debata o AGI pozostaje szczególnie podatna na tę niejednoznaczność. Pojęcie sztucznej inteligencji ogólnej nie posiada jednej, powszechnie akceptowanej definicji operacyjnej. W zależności od propozycji obejmuje ono zdolność wykonywania szerokiego spektrum zadań intelektualnych, skuteczną adaptację do nowych sytuacji, samodzielne uczenie się, poziom kompetencji porównywalny z ludzkim lub możliwość zastępowania człowieka w większości ekonomicznie wartościowych czynności. Definicje te nie są równoważne i prowadzą do odmiennych sposobów pomiaru postępu. System może bowiem przewyższać przeciętnego człowieka w zadaniach matematycznych, a jednocześnie pozostawać zawodny w prostych, nietypowych

sytuacjach praktycznych lub sprawnie operować w świecie cyfrowym, nie posiadając zdolności działania w rzeczywistości fizycznej.

Konieczność wielowymiarowej oceny postępu i ryzyk AI

Z tego względu twierdzenia o bliskości AGI wymagają każdorazowo doprecyzowania przyjętego kryterium, zakresu kompetencji, rodzaju środowiska oraz poziomu dopuszczalnego ryzyka. W tym miejscu historia kolejnych zim AI dostarcza istotnej lekcji metodologicznej. W przeszłości badacze wielokrotnie obserwowali imponujące postępy w wybranym obszarze, by następnie odkryć, że ekstrapolacja tych wyników na całość ludzkiej inteligencji wymaga rozwiązania zupełnie innych problemów architektonicznych. W latach pięćdziesiątych nadzieje wiązano z formalnym rozumowaniem i perceptronami, w latach osiemdziesiątych z systemami ekspertowymi, w kolejnych dekadach z metodami statystycznymi, a współcześnie z wielkoskalowymi modelami generatywnymi. Nie oznacza to jednak, że obecna fala nieuchronnie zakończy się podobnym rozczarowaniem. Każda epoka ma własną strukturę możliwości, a współczesne systemy osiągnęły już poziom praktycznej użyteczności niedostępny wcześniejszym generacjom.

Wskazuje to raczej na konieczność krytycznego badania granic generalizacji, zamiast przyjmowania, że szybki wzrost jednego wskaźnika musi prowadzić do jakościowego rozwiązania wszystkich pozostałych problemów. Należy przy tym odróżnić trzy rodzaje potencjalnej korekty. Pierwsza z nich to korekta finansowa, która może nastąpić w wyniku spadku oczekiwanych stóp zwrotu, nadmiernego zadłużenia, wysokich kosztów infrastruktury lub zmiany warunków makroekonomicznych. Druga dotyczy korekty technologicznej – polega ona na odkryciu ograniczeń danej architektury, co wymusza wprowadzenie nowych metod, efektywniejszych procedur treningowych lub zasadniczą reorganizację sposobu przetwarzania informacji. Trzecia ma charakter społeczno-instytucjonalny i może wynikać z ujawnienia nieakceptowalnych szkód, konfliktów z prawami podstawowymi, naruszeń prywatności bądź niewystarczającej rozliczalności procesów automatycznych. Te trzy procesy mogą zachodzić równocześnie, lecz nie są tożsame. Przedsiębiorstwo może ponieść porażkę finansową mimo stworzenia technicznie wartościowego rozwiązania, a technologia może pozostawać rentowna, mimo że generuje istotne szkody społeczne. Rzeczywista dojrzałość ekosystemu AI wymaga rozpoznawania wszystkich trzech poziomów oceny.

Warto również zakwestionować przekonanie, że korekta rynkowa zawsze pełni pozytywną funkcję oczyszczającą. Zwolennicy takiej interpretacji twierdzą, że spadek finansowania eliminuje firmy pozbawione trwałego modelu biznesowego i kieruje kapitał ku rozwiązaniom bardziej efektywnym. Z punktu widzenia ekonomii innowacji obraz ten jest jednak niepełny. Kryzys finansowy może

prowadzić również do upadku wartościowych przedsięwzięć badawczych, rozproszenia kompetentnych zespołów, ograniczenia różnorodności innowacji i dalszej koncentracji rynku wokół największych podmiotów z ogromnymi rezerwami kapitałowymi. Nie wszystkie projekty o wysokiej wartości społecznej potrafią szybko generować przychody – szczególnie w obszarach edukacji, zdrowia publicznego, technologii asystujących czy badań podstawowych. W konsekwencji mechanizm selekcji rynkowej nie musi prowadzić do zachowania rozwiązań najkorzystniejszych z perspektywy społecznej. Kryterium rentowności przedsiębiorstwa i kryterium społecznej użyteczności są ze sobą powiązane, ale nie można ich utożsamiać.

Ryzyko zależności technologicznej i dylemat automatyzacji

Problem ten staje się szczególnie istotny w obliczu globalnej nierównowagi technologicznej. Rozwój zaawansowanych modeli AI wymaga dostępu do infrastruktury obliczeniowej, specjalistycznej wiedzy, danych i kapitału – zasobów rozdzielonych nierównomiernie między państwami oraz instytucjami. Rosnące koszty trenowania systemów o najwyższych parametrach mogą pogłębiać koncentrację kompetencji w rękach nielicznych przedsiębiorstw, podczas gdy społeczeństwa o ograniczonych zasobach pozostają głównie odbiorcami produktów opracowanych poza ich własnymi systemami naukowymi i gospodarczymi. Z perspektywy teorii zależności technologicznej może to prowadzić do utrwalania asymetrii między podmiotami tworzącymi infrastrukturę poznawczą a tymi, które są od niej zależne. Jednocześnie rozwój modeli otwartych, mniejszych architektur i metod efektywnego dostrajania stwarza szansę na ograniczanie tych barier. Ostateczny kierunek zmian nie jest zatem przesądzony; jego analiza wymaga uwzględnienia polityki przemysłowej, regulacji konkurencji, otwartości badań oraz dostępności infrastruktury publicznej.

W polityce gospodarczej kluczowe staje się zagadnienie suwerenności technologicznej. Państwo wykorzystujące systemy AI w administracji, ochronie zdrowia, edukacji czy zarządzaniu infrastrukturą krytyczną powinno dysponować zdolnością do oceny ich niezawodności, bezpieczeństwa i zgodności z prawem. Nie oznacza to konieczności samodzielnego wytwarzania wszystkich komponentów, gdyż współczesne łańcuchy wartości technologicznej mają charakter globalny. Wymaga jednak odpowiednich kompetencji zamawiającego, przejrzystości warunków korzystania z usług, możliwości migracji między dostawcami oraz zachowania nadzoru nad strategicznymi zasobami informacyjnymi. Ekonomiczna atrakcyjność gotowego rozwiązania nie powinna przesłaniać kosztów uzależnienia od zewnętrznej infrastruktury, szczególnie gdy zmiana dostawcy staje się trudna ze względu na architekturę danych, integrację oprogramowania lub brak

odpowiednich kompetencji personelu. Historia AI wskazuje, że zmiany paradygmatów technologicznych mogą następować szybko, natomiast instytucjonalne zależności wytworzone podczas ich wdrażania bywają znacznie trwalsze.

Ważne jest również rozróżnienie pomiędzy automatyzacją a augmentacją ludzkich kompetencji. Automatyzacja oznacza przekazywanie systemowi wykonywania określonych czynności, natomiast augmentacja polega na rozszerzaniu zdolności człowieka poprzez dostarczanie mu informacji, narzędzi analitycznych i możliwości działania, które wcześniej pozostawały niedostępne albo zbyt kosztowne. Oba kierunki mogą zwiększać produktywność, lecz prowadzą do odmiennych konsekwencji organizacyjnych i społecznych. Technologia całkowicie zastępująca określoną czynność może ograniczać zapotrzebowanie na dotychczasową pracę, podczas gdy narzędzie wspomagające pozwala wykonywać bardziej złożone zadania osobom o niższym poziomie specjalizacji. Nie można jednak zakładać, że augmentacja jest zawsze korzystna, a automatyzacja szkodliwa. Automatyzowanie czynności niebezpiecznych i monotonnych poprawia warunki pracy, podczas gdy nadmierne wspomaganie poznawcze może prowadzić do dekwalifikacji użytkowników.

Technologia nie zastąpi ludzkiej roztropności i odpowiedzialności

Ocena zależy od charakteru zadania, podziału odpowiedzialności i sposobu organizacji procesu uczenia się. Historia kolejnych przełomów wskazuje jednak, że wzrost sprawności systemów może paradoksalnie prowadzić do osłabienia zdolności człowieka do rozpoznawania ich błędów. Gdy technologia przez długi czas działa poprawnie, użytkownicy tracą czujność i stopniowo zapominają praktyczne kompetencje niezbędne do samodzielnego wykonania zadania. Zjawisko to jest dobrze udokumentowane w badaniach nad automatyzacją transportu, systemami wspomagania decyzji oraz interakcją człowiek-maszyna: im rzadsza potrzeba interwencji, tym trudniej utrzymać odpowiedni poziom doświadczenia operacyjnego. Pojawia się tu problem ironii automatyzacji, analizowany m.in. przez Lisanne Bainbridge; polega on na tym, że człowiekowi pozostawia się głównie zadania najtrudniejsze, występujące sporadycznie i wymagające błyskawicznej reakcji w warunkach niepewności. W odniesieniu do systemów agentowych AI oznacza to konieczność projektowania mechanizmów kontroli, które nie ograniczają się do deklaratywnej obecności człowieka, lecz realnie podtrzymują jego zdolność do rozumienia procesu i skutecznej interwencji.

Tak rozumiana cykliczność technologiczna ma istotny wymiar aksjologiczny. Kolejne generacje AI rozwijano w przekonaniu, że zwiększenie mocy obliczeniowej pozwoli osiągnąć nowe poziomy skuteczności, racjonalności i sprawności organizacyjnej. Historia pokazała jednak, że nie istnieje

automatyczne przejście od inteligencji instrumentalnej do mądrości praktycznej. System może wykonywać zadanie szybciej i dokładniej od człowieka, ale nie rozstrzygnie, czy samo zadanie powinno być w ogóle realizowane, jakie wartości należy uwzględnić ani kto powinien ponosić odpowiedzialność za jego społeczne konsekwencje.

W filozofii praktycznej rozróżnienie to przypomina Arystotelesowską koncepcję **phronesis**, czyli roztropności pozwalającej dobrać właściwe działanie do konkretnych okoliczności. Roztropność nie sprowadza się do znajomości abstrakcyjnych reguł – wymaga doświadczenia, umiejętności oceny sytuacji oraz zrozumienia związku między celem działania a dobrem. Nie oznacza to, że współczesne systemy nie mogą wspomagać ludzkiego osądu, lecz przypomina, że poprawność procedury obliczeniowej nie jest samodzielnym kryterium słuszności decyzji moralnej. W tym kontekście szczególnie aktualna staje się krytyka technologicznego determinizmu. Przekonanie, że rozwój AI zmierza w jednym koniecznym kierunku i nieuchronnie ukształtuje określony porządek społeczny, pomija rolę decyzji politycznych, instytucji, regulacji, struktur własności czy konfliktów interesów. Historia nauki dostarcza licznych przykładów tego, jak podobne możliwości techniczne wykorzystywano w różny sposób, zależnie od warunków społecznych i celów gospodarczych. Automatyzacja może służyć zwiększaniu rentowności poprzez redukcję zatrudnienia, ale może też wspierać bezpieczeństwo pracy, dostępność usług lub rozwój kompetencji.

Systemy generatywne mogą wzmacniać monopolizację informacji, lecz mogą także obniżać bariery dostępu do wiedzy. Odpowiedzialność za wybór między tymi drogami spoczywa na ludziach i instytucjach kształtujących warunki wykorzystania technologii. Właśnie dlatego kategoria wychowania AI powinna być rozumiana jako propozycja odpowiedzialnej relacji społecznej z systemami obliczeniowymi, a nie jako opis nieuchronnej ewolucji maszyn ku ludzkiej dojrzałości psychicznej. Przeszłość AI stanowi zatem ostrzeżenie przed utożsamianiem historii technologii z prostą opowieścią o kolejnych zwycięstwach nad ograniczeniami poznawczymi. Każdy paradygmat rozwijał określone możliwości, ale jednocześnie ujawniał nowe problemy: systemy symboliczne zmagały się z reprezentacją wiedzy i złożonością obliczeniową, systemy ekspertowe z kosztami pozyskiwania reguł, modele statystyczne z generalizacją danych, głębokie sieci neuronowe z interpretowalnością i odpornością, a współczesne modele generatywne dodatkowo z niezawodnością rozumowania, kontrolą autonomii i społecznymi skutkami masowego zastosowania. Ograniczenia te nie są od siebie niezależne – kolejne generacje dziedziczą trudności poprzedników, choć próbują rozwiązywać je innymi metodami.

Dojrzałość cywilizacyjna jako warunek odpowiedzialnego rozwoju AI

Postęp należy zatem rozpatrywać jako wielowymiarowy proces przesuwania granic możliwości, a nie prostą sekwencję etapów prowadzących do jednego, z góry określonego punktu. Dopiero w tym świetle możliwe staje się krytyczne rozpoznanie współczesnych obietnic dotyczących AGI. Przyjęcie tezy o możliwości zbudowania systemów posiadających szerokie kompetencje poznawcze jest stanowiskiem badawczym, którego uzasadnienie zależy od konkretnych definicji, osiągnięć empirycznych i przyszłych rezultatów eksperymentów. Przedstawianie natomiast konkretnej daty powstania AGI jako niemal pewnego wydarzenia wykracza poza aktualny stan wiedzy, o ile nie zostanie wsparte wiarygodną metodologią prognostyczną. Ograniczenia historycznych prognoz wynikają m.in. z tego, że badacze nie zawsze potrafią przewidzieć pojawienie się nowych architektur, zmiany w dostępności zasobów, przyszłe odkrycia teoretyczne czy bariery społeczne i infrastrukturalne. Równie problematyczne pozostaje arbitralne przesuwanie definicji AGI w odpowiedzi na rozwój aktualnych systemów.

Nauka potrzebuje kryteriów umożliwiających zarówno potwierdzanie, jak i podważanie twierdzeń o osiągniętym poziomie kompetencji; bez nich debata łatwo przekształca się w konkurencję marketingowych deklaracji i filozoficznych intuicji. Historia AI ujawnia doniosłość samokrytyki naukowej jako warunku rzeczywistego postępu. Najbardziej produktywne okresy rozwoju tej dziedziny nie zawsze następowały po bezwarunkowym zwiększeniu nakładów finansowych, lecz często po rozpoznaniu ograniczeń dotychczasowych metod, zmianie założeń i podjęciu badań nad problemami wcześniej uznawanymi za marginalne. Krytyka perceptronów przyczyniła się do głębszego zrozumienia reprezentacyjnych ograniczeń określonych architektur, niepowodzenia systemów ekspertowych zwiększyły zainteresowanie uczeniem z danych, a trudności modeli statystycznych pobudziły rozwój bardziej zaawansowanych metod reprezentacji i optymalizacji. Podobnie współczesne badania nad bezpieczeństwem, wyjaśnialnością, skutkami społecznymi i zarządzaniem autonomią nie powinny być traktowane jako przeszkody w rozwoju AI, lecz jako niezbędne elementy jej dojrzałości. Technologia pozostająca poza skuteczną kontrolą społeczną może osiągać imponującą sprawność operacyjną, a jednocześnie generować narastające koszty, które ostatecznie ograniczą jej społeczny potencjał.

Rozpoznanie historycznej cykliczności prowadzi zatem do bardziej wymagającego rozumienia odpowiedzialności cywilizacyjnej. Ludzkość nie powinna ani podporządkowywać swojej przyszłości kolejnym technologicznym obietnicom, ani odrzucać rzeczywistych korzyści płynących z rozwoju inteligentnych systemów pod wpływem wspomnień wcześniejszych rozczarowań. Potrzebuje

instytucji zdolnych do odróżniania sukcesów empirycznych od narracji spekulacyjnych, korzyści prywatnych od wartości społecznej oraz efektywności technicznej od zgodności z prawami i potrzebami człowieka.

Właśnie taka perspektywa pozwala zrozumieć, dlaczego przyszłość AI nie zależy wyłącznie od liczby parametrów modeli, mocy procesorów czy wielkości kapitału inwestycyjnego. Zależy również od jakości wiedzy umożliwiającej ocenianie tych parametrów oraz od dojrzałości instytucji potrafiących wykorzystywać technologię bez rezygnowania z odpowiedzialności za jej następstwa. Ostatecznie historia kolejnych fal i zim sztucznej inteligencji ukazuje nie tyle zawodność ludzkiego dążenia do tworzenia coraz doskonalszych maszyn, ile niedoskonałość sposobów przewidywania, w jaki sposób nowe możliwości przekształcą społeczeństwo. W każdej epoce pojawiała się skłonność do traktowania najbardziej aktualnego osiągnięcia jako zwiastuna rychłego przezwyciężenia wszystkich pozostałych ograniczeń. Tymczasem prawdziwa wartość odkrycia ujawniała się często dopiero po latach, gdy zostało ono połączone z innymi technologiami, odpowiednią infrastrukturą i zmianami organizacyjnymi. Współczesna generatywna AI może okazać się fundamentem długookresowej transformacji gospodarczej i poznawczej, nawet jeżeli część obecnych oczekiwań inwestycyjnych zostanie zweryfikowana negatywnie.

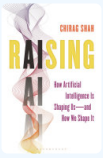
Może ona również napotkać bariery, których przezwyciężenie będzie wymagało istotnych zmian architektonicznych. Żaden z tych scenariuszy nie jest automatyczną konsekwencją przeszłości. Najbardziej niebezpieczne byłoby zatem przekonanie, że sam przebieg historii zwalnia człowieka z obowiązku podejmowania odpowiedzialnych decyzji. Cykliczność technologiczna nie jest prawem natury nakazującym społeczeństwom kolejno doświadczać euforii, katastrofy i ponownego zachwytu; stanowi rezultat współdziałania ograniczeń poznawczych, struktur gospodarczych, mechanizmów finansowania i instytucjonalnych sposobów reagowania na niepewność. Rozpoznanie tych zależności pozwala ograniczać koszty przyszłych błędów, choć nie umożliwia całkowitego wyeliminowania ryzyka.

Dojrzewanie ekosystemu AI wymaga zatem równoczesnego dojrzewania społeczeństw, które technologię tworzą i wykorzystują. Rozwój maszyn staje się bowiem miarą nie tylko osiągnięć informatyki, ale również zdolności ludzkiej cywilizacji do krytycznego poznawania własnych możliwości, ograniczeń i wartości.

> Czy rozwój maszyn jest w istocie postępem informatycznym, czy raczej lustrem, w którym odbija się nasza cywilizacyjna niedojrzałość? Prawdziwa miara inteligencji systemów AI może nie leżeć w ich zdolności do symulowania ludzkiego myślenia, lecz w naszej zdolności do krytycznego rozpoznania momentu, w którym sprawność

> obliczeniowa przestaje być tożsama z mądrością. Ostatecznie to nie liczba parametrów modelu, lecz jakość naszych instytucji i etyka odpowiedzialności zdecydują o tym, czy kolejna fala zachwytu nie stanie się jedynie wstępem do kolejnej zimy.

Inspiracje

[1] 

"Raising AI" Chirag Shah

Bloomsbury Academic | ISBN: 9798216452409

Chirag Shah jest informatykiem i naukowcem zajmującym się informatyką, obecnie profesorem w Szkole Informatyki oraz profesorem nadzwyczajnym w Szkole Informatyki i Inżynierii im. Paula G. Allena na Uniwersytecie Waszyngtońskim. Uzyskał doktorat z informatyki na Uniwersytecie Karoliny Północnej w Chapel Hill, tytuł magistra informatyki na Uniwersytecie Massachusetts w Amherst oraz tytuł magistra technologii (M.Tech.) na Indyjskim Instytucie Technologii w Madrasie. Shah jest dyrektorem założycielem InfoSeeking Lab i współdyrektorem założycielem centrum RAISE ds. odpowiedzialnej sztucznej inteligencji. Jego badania koncentrują się na inteligentnym dostępie do informacji, wspólnym wyszukiwaniu informacji, systemach wyszukiwania i rekomendacji, agentowej sztucznej inteligencji oraz odpowiedzialnej sztucznej inteligencji. Jako wybitny członek ACM i ASIS&T, Shah wniósł znaczący wkład w wyszukiwanie zorientowane na użytkownika, uczciwość uczenia maszynowego i etyczny rozwój sztucznej inteligencji.

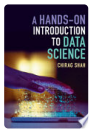
***Chirag Shah** is a computer and information scientist, currently serving as a Professor in the Information School and an Adjunct Professor in the Paul G. Allen School of Computer Science & Engineering at the University of Washington. He earned his Ph.D. in Information Science from the University of North Carolina at Chapel Hill, an M.S. in Computer Science from the University of Massachusetts Amherst, and an M.Tech. from the Indian Institute of Technology Madras. Shah is the Founding Director of the InfoSeeking Lab and Founding Co-Director of the RAISE center for responsible AI. His research focuses on intelligent information access, collaborative information seeking, search and recommendation systems, agentic artificial intelligence, and responsible AI. A Distinguished Member of both ACM and ASIS&T, Shah has contributed extensively to user-centered search, machine learning fairness, and ethical AI development.*

University of Washington

Literatura uzupełniająca

Książki

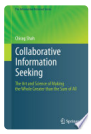
Książki autora i dzieła przywołane w tekście:



[1] "A Hands-On Introduction to Data Science"

Chirag Shah

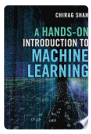
2020 | Cambridge University Press | ISBN: 9781108472449



[2] "Collaborative Information Seeking"

Chirag Shah

2012 | Springer Science & Business Media | ISBN: 9783642288135



[3] "A Hands-On Introduction to Machine Learning"

Chirag Shah

2022 | Cambridge University Press | ISBN: 9781009123303



[4] "Task Intelligence for Search and Recommendation"

Chirag Shah, Ryen W. White

2022 | Springer Nature | ISBN: 9783031023262



[5] "Agent Nation"

Chirag Shah

2026 | Springer Nature | ISBN: 9798868824548

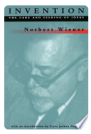


[6] "Interactive IR User Study Design, Evaluation, and Reporting"

Jiqun Liu, Chirag Shah

2022 | Springer Nature | ISBN: 9783031023194

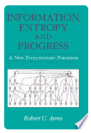
Książki autorów przywoływanych i literatura pokrewna:



[7] "Invention: The Care and Feeding of Ideas"

Norbert Wiener

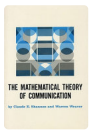
MIT Press | ISBN: 9780262731119



[8] "information entropy"

Claude Shannon

Springer Science & Business Media | ISBN: 9780883189115

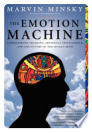


[9] "The mathematical theory of communication"

Claude Shannon

[10] "Programming a Computer for Playing Chess"

Claude Shannon



[11] "The Emotion Machine"

Marvin Minsky

Simon and Schuster | ISBN: 9781416579304



[12] "Robotics"

Marvin Minsky

Anchor Books

[13] "Computation: Finite and Infinite Machines"

Marvin Minsky

[14] "Human Proplem Solving"

Allen Newell

1972

[15] "Computer Simulation of Human Thinking"

Allen Newell, Herbert Alexander Simon

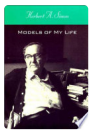
1961



[16] "Human Problem Solving"

Allen Newell, Herbert Alexander Simon

1972 | Prentice Hall



[17] "Models of my life"

Herbert A. Simon

1996 | MIT Press | ISBN: 9780262691857



[18] "Models of Bounded Rationality"

Univ Of Chicago, Herbert A. Simon

1997 | Mit Press | ISBN: 9780262519434

[19] "Tables of Q-funtions for Two Perceptron Models (Project PARA)"

Frank Rosenblatt

1960

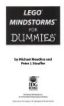
[20] "Perceptions and the Theory of Brain Mechanisms"

Frank Rosenblatt

1962

[21] "Logo"

Seymour Papert



[22] "Lego Mindstorms"

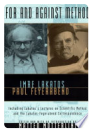
Seymour Papert

For Dummies | ISBN: 9780764507670



[23] "The children's machine: Rethinking school in the age of the computer"

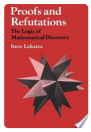
Seymour Papert



[24] "For and Against Method"

Imre Lakatos

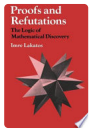
1999 | University of Chicago Press | ISBN: 9780226467740



[25] "Proofs and refutations: the logic of mathematical discovery"

Imre Lakatos

Cambridge University Press | ISBN: 9780521290388



[26] "Proofs and Refutations"

Imre Lakatos

Cambridge University Press | ISBN: 9780521290388



[27] "Papers in microbial genetics ; bacteria and bacterial viruses"

Joshua Lederberg

Hassell Street Press | ISBN: 9781013714597



[28] "Personal Knowledge"

Michael Polanyi

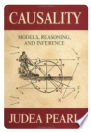
1998 | Psychology Press | ISBN: 9780415151498



[29] "Technology and Science, Innovation and Discovery."

Michael Polanyi

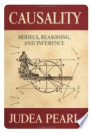
2026 | Independently Published | ISBN: 9798196619229



[30] "Causality"

Judea Pearl

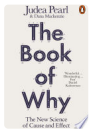
2000 | Cambridge University Press | ISBN: 9780521773621



[31] "Causality: Models, Reasoning and Inference"

Judea Pearl

Cambridge University Press | ISBN: 9780521773621



[32] "The Book of Why"

Judea Pearl

Penguin UK | ISBN: 9780241242643

[33] "In Defense of Metrics: Metrics Sufficiently Encode Typical Human Preferences Regarding Hydrological Model Performance"

Sepp Hochreiter



[34] "Long Short Term Memory"

Sepp Hochreiter

One Billion Knowledgeable



[35] "Winter Is Coming"

Garry Kasparow

Hachette UK | ISBN: 9781610396219

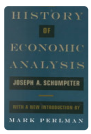
[36] "How Life Imitates Chess"

Garry Kasparow

Vintage | ISBN: 9780434015405

[37] "Boltzmann Machines: Constraint Satisfaction Networks that Learn"

Geoffrey Hinton



[38] "History of Economic Analysis"

Joseph A. Schumpeter

Oxford University Press, USA | ISBN: 9780195105599

[39] "History of Economic Analysis (Oxford Press, 1955)"

Joseph A. Schumpeter



[40] "Technological Revolutions and Financial Capital"

Carlota Perez

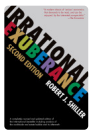
Edward Elgar Publishing | ISBN: 9781781005323



[41] "Jak se loví hlupáci: ekonomie manipulace a klamu: nenechte sebou manipulovat"

Robert J. Shiller

Management Press | ISBN: 9788072614936



[42] "Irrational Exuberance"

Robert J. Shiller

2005 | Princeton University Press | ISBN: 9780691123356



[43] "Animal Spirits"

Robert J. Shiller

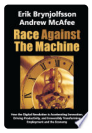
Princeton University Press | ISBN: 9780691142333



[44] "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies"

Erik Brynjolfsson

W. W. Norton & Company | ISBN: 9780393241259



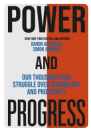
[45] "Race Against the Machine"

Erik Brynjolfsson

Brynjolfsson and McAfee | ISBN: 9780984725113

[46] "Distorted innovation: Does the market get the direction of technology right?"

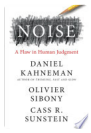
Daron Acemoglu



[47] "Power and Progress"

Daron Acemoglu

2023 | ISBN: 9781399804455



[48] "Noise: A Flaw in Human Judgment"

Daniel Kahneman

Hachette UK | ISBN: 9780316451383

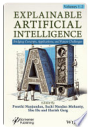
[49] "The framing of decisions and the psychology of choice"

Amos Tversky

[50] **"prospect theory"**

Amos Tversky

1979



[51] **"Explainable Artificial Intelligence"**

Preethi Nanjundan, Sachi Nandan Mohanty, Shu Hu, Harish Garg

2026 | John Wiley & Sons | ISBN: 9781394357765



[52] **"Artificial Intelligence for Drug Development, Precision Medicine, and Healthcare"**

Mark Chang

2020 | CRC Press | ISBN: 9781000767308



[53] **"Artificial Intelligence in Biomedical and Modern Healthcare Informatics"**

M. A. Ansari, R.S Anand, Pragati Tripathi, Rajat Mehrotra, Md Belal Bin Heyat

2024 | Elsevier | ISBN: 9780443218712



[54] **"Machine Learning and Other Artificial Intelligence Applications, An Issue of Neuroimaging Clinics of North America, E-Book"**

Reza Forghani

2020 | Elsevier Health Sciences | ISBN: 9780323712453

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] **"I.—COMPUTING MACHINERY AND INTELLIGENCE"**

Alan Mathison Turing

1950 | Mind | DOI: 10.1093/mind/lix.236.433

Źródło

[2] **"Computing Machinery and Intelligence (1950)"**

Alan Mathison Turing

2004 | Oxford University Press eBooks | DOI: 10.1093/oso/9780198250791.003.0017

Źródło

[3] "On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction"

Alan Mathison Turing

1938 | Proceedings of the London Mathematical Society | DOI: 10.1112/plms/s2-43.6.544

Źródło

[4] "Computing Machinery and Intelligence"

Alan Mathison Turing

2007 | Springer eBooks | DOI: 10.1007/978-1-4020-6710-5_3

Źródło

[5] "Mildred Pitts Walter oral history interview conducted by David P. Cline in San Mateo, California, 2013-03-01."

David P. Cline, Mildred Pitts Walter

2013 | Library of Congress eBooks

Źródło

[6] "The Human Use of Human Beings"

Norbert Wiener

1950

Źródło

[7] "Artificial Intelligence and Human Rights"

John McCarthy, Marv﷿n M﷿nsky, Nathan﷿el Rochester, C. Shannon

2024 | Human Rights Law Review | DOI: 10.1093/hrlr/ngae031

Źródło

[8] "Circumscription—A form of non-monotonic reasoning"

John McCarthy

1980 | Artificial Intelligence | DOI: 10.1016/0004-3702(80)90011-9

Źródło

[9] "Recursive functions of symbolic expressions and their computation by machine, Part I"

John McCarthy

1960 | Communications of the ACM | DOI: 10.1145/367177.367199

Źródło

[10] "Some Philosophical Problems from the Standpoint of Artificial Intelligence"

John McCarthy, Patrick J. Hayes

1981 | Elsevier eBooks | DOI: 10.1016/b978-0-934613-03-3.50033-7

Źródło

[11] "Applications of circumscription to formalizing common-sense knowledge"

John McCarthy

1986 | Artificial Intelligence | DOI: 10.1016/0004-3702(86)90032-9

Źródło

[12] "Steps toward Artificial Intelligence"

Marvin Minsky

1961 | Proceedings of the IRE | DOI: 10.1109/jrproc.1961.287775

Źródło

[13] "A FRAMEWORK FOR REPRESENTING KNOWLEDGE"

Marvin Minsky

1988 | Elsevier eBooks | DOI: 10.1016/b978-1-4832-1446-7.50018-2

Źródło

[14] "Matter, Mind and Models"

Marvin Minsky

1965 | DSpace@MIT (Massachusetts Institute of Technology)

Źródło

[15] "Theory of neural-analog reinforcement systems and its application to the brain-model problem"

Marvin Minsky

1954

Źródło

[16] "Symbolic programming"

N Rochester

1953 | Transactions of the I R E Professional Group on Electronic Computers | DOI: 10.1109/irepgelc.1953.6499409

Źródło

[17] "After-pulses from drift chambers"

T.J. Summer, P.D. Smith, G.K. Rochester, R.K. Sood

1984 | Nuclear Instruments and Methods in Physics Research | DOI: 10.1016/0167-5087(84)90040-1

Źródło

[18] "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955"

John McCarthy, M. Minsky, N. Rochester, C. Shannon

2006 | The AI Magazine | DOI: 10.1609/aimag.v27i4.1904

Źródło

[19] "Pattern recognition in interdisciplinary perception and intelligence 1"

Mccarthy, D. College, M. Minsky, Harvard University, N. Rochester

2008

Źródło

[20] "Computer Theatre: Raising the Curtain on a Visual Novel Prototype"

Katie M. Biittner, Carl Therrien, John Aycock, Dona Bailey, Paul Allen Newell

2024 | ROMchip

Źródło

[21] "No Prototypical Prototypes: Prototypes as the Stone Tools of Game Development"

Katie M. Biittner, Paul Allen Newell, John Aycock

2026 | International Conference on the Foundations of Digital Games (FDG) | DOI: 10.1145/3815598.3815675

Źródło

[22] "AN EXAMPLE OF HUMAN CHESS PLAY IN THE LIGHT OF CHESS PLAYING PROGRAMS"

Allen Newell, Herbert A. Simon

1964 | DOI: 10.21236/ad0619386

Źródło

[23] "Simulation of Human Processing of Information"

Allen Newell, H. A. Simon

1965 | The American Mathematical Monthly | DOI: 10.2307/2313317

Źródło

[24] "Models of Discovery"

John Kidd, Herbert A. Simon

1978 | Journal of the Operational Research Society | DOI: 10.1057/jors.1978.270

Źródło

[25] "WSUD “Best in Class”—Case Studies From Australia, New Zealand, United States, Europe, and Asia"

S. Cook, M. V. van Roon, Lisa Ehrenfried, J. LaGro, Qian Yu

2019 | Approaches to Water Sensitive Urban Design | DOI: 10.1016/b978-0-12-812843-5.00027-7

Źródło

[26] "A scalable content-addressable network"

Sylvia Ratnasamy, Paul Francis, Mark James Handley, Richard M. Karp, Scott J. Shenker

2001 | ACM SIGCOMM Conference (SIGCOMM) | DOI: 10.1145/383059.383072

Źródło

[27] "LogP: towards a realistic model of parallel computation"

David Culler, Richard M. Karp, David A. Patterson, Abhijit Sahay, Klaus Erik Schauser

1993 | ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming | DOI: 10.1145/155332.155333

Źródło

[28] "A characterization of the minimum cycle mean in a digraph"

Richard M. Karp

1978 | Discrete Mathematics | DOI: 10.1016/0012-365x(78)90011-0

Źródło

[29] "The Perceptron: A Probabilistic Model for Information Storage and Organization (1958)"

Frank Rosenblatt

2021 | Ideas That Created the Future | DOI: 10.7551/mitpress/12274.003.0020

Źródło

[30] "(1958) F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," Psychological Review 65: 386-408"

F. Rosenblatt

1988 | Neurocomputing, Volume 1 | DOI: 10.7551/mitpress/4943.003.0010

Źródło

[31] "The perceptron: A probabilistic model for information storage and organization in the brain."

F. Rosenblatt

1958 | Psychological Review | DOI: 10.1037/h0042519

Źródło

[32] "The Connected Family: Bridging the Digital Generation Gap"

Seymour A. Papert

1996

Źródło

[33] "Piaget ' s theory of intelligence : operational aspect"

S. Papert

2008

Źródło

[34] "History of Science and Its Rational Reconstructions"

Imre Lakatos

1970 | PSA Proceedings of the Biennial Meeting of the Philosophy of Science Association | DOI: 10.1086/psaprocbienmeetp.1970.495757

Źródło

[35] "For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence"

Imre Lakatos, Matteo Motterlini, Paul K. Feyerabend

1999 | University of Chicago Press eBooks | DOI: 10.7208/chicago/9780226467030.001.0001

Źródło

[36] "Cauchy and the continuum: The significance of non-standard analysis for the history and philosophy of mathematics"

Imre Lakatos

1978 | The Mathematical Intelligencer | DOI: 10.1007/bf03023263

Źródło

[37] "The Popperian versus the Kuhnian Research Programme"

Imre Lakatos

1990 | Philosophy of Science and the Occult | DOI: 10.1515/9781438404981-015

Źródło

[38] **"Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project"**

Robert Lindsay, Bruce G. Buchanan, Edward A. Feigenbaum, Joshua Lederberg

1980 | Profiles in Science

Źródło

[39] **"DENDRAL: A case study of the first expert system for scientific hypothesis formation"**

Robert Lindsay, Bruce G. Buchanan, Edward A. Feigenbaum, Joshua Lederberg

1993 | Artificial Intelligence | DOI: 10.1016/0004-3702(93)90068-m

Źródło

[40] **"The Art of Artificial Intelligence. 1. Themes and Case Studies of Knowledge Engineering"**

Edward A. Feigenbaum

1977 | Defense Technical Information Center (DTIC) | DOI: 10.21236/ada046289

Źródło

[41] **"Knowledge engineering for medical decision making: A review of computer-based clinical decision aids"**

Edward Hance Shortliffe, Bruce G. Buchanan, Edward A. Feigenbaum

1979 | Proceedings of the IEEE | DOI: 10.1109/proc.1979.11436

Źródło

[42] **"Infectious History"**

Joshua Lederberg

2000 | Science | DOI: 10.1126/science.288.5464.287

Źródło

[43] **"Personal Knowledge: Towards a Post-Critical Philosophy"**

Louis Arnaud Reid, Michael Polanyi

1959 | British Journal of Educational Studies | DOI: 10.2307/3119338

Źródło

[44] **"Some applications of the transition state method to the calculation of reaction velocities, especially in solution"**

Meredith Gwynne Evans, Michael Polanyi

1935 | Transactions of the Faraday Society | DOI: 10.1039/tf9353100875

Źródło

[45] **"Causality: Models, Reasoning, and Inference"**

Judea Pearl

2009 | Cambridge University Press eBooks | DOI: 10.1017/cbo9780511803161

Źródło

[46] **"Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference"**

Judea Pearl

1988

Źródło

[47] **"The Book of Why: The New Science of Cause and Effect"**

Judea Pearl, Dana N. Mackenzie

2018

Źródło

[48] **"Generalized best-first search strategies and the optimality of A*"**

Rina Dechter, Judea Pearl

1985 | Journal of the ACM | DOI: 10.1145/3828.3830

Źródło

[49] **"Support Vector Method for Function Approximation, Regression Estimation and Signal Processing"**

Vladimir N. Vapnik, Steven E. Golowich, Alex J. Smola

1996 | neural information processing systems

Źródło

[50] **"Pattern recognition using generalized portrait method"**

Vladimir N. Vapnik

1963 | Automation and Remote Control

Źródło

[51] **"Direct Methods in Statistical Learning Theory"**

Vladimir N. Vapnik

2000 | The Nature of Statistical Learning Theory | DOI: 10.1007/978-1-4757-3264-1_8

Źródło

[52] "Architectures and KV Replay for Depth-Recurrent Language Models"

Lucia Licakova, Niklas Schmidinger, Lukas Hauzenberger, Sebastian Böck, S. Hochreiter

2026

Źródło

[53] "The Offline-Frontier Shift: Diagnosing Distributional Limits in Generative Multi-Objective Optimization"

Stephanie Holly, Alexandru-Ciprian Zăvoianu, S. Silber, S. Hochreiter, Werner Zellinger

2026 | arXiv.org | DOI: 10.48550/arxiv.2602.11126

Źródło

[54] "Exploring Through Random Curiosity with General Value Functions"

Aditya A. Ramesh, Louis Kirsch, Sjoerd van Steenkiste, Jürgen Schmidhuber

2022 | neural information processing systems | DOI: 10.52202/068431-1361

Źródło

[55] "AI in Drug Discovery: First International Workshop, AIDD 2024, Held in Conjunction with ICANN 2024, Lugano, Switzerland, September 19, 2024, Proceedings"

Djork-Arné Clevert, Michael Wand, Kristína Malinovská, Jürgen Schmidhuber, Igor V. Tetko

2024 | Lecture notes in computer science | DOI: 10.1007/978-3-031-72381-0

Źródło

[56] "SciDiagramEdit: Learning to Edit Scientific Diagrams from Paper Revisions"

Yasheng Sun, Zezi Zeng, Yifan Yang, Chong Luo, Wenyi Wang, Jürgen Schmidhuber

2026 | King Abdullah University of Science and Technology Repository (King Abdullah University of Science and Technology) | DOI: 10.48550/arxiv.2607.15272

Źródło

[57] "Planning to Explore: Curiosity-Driven Planning for LLM Test Generation"

Alfonso Amayuelas, Firas Laakom, Piotr Piękos, Wenyi Wang, Yifan Xu, Jürgen Schmidhuber

2026 | King Abdullah University of Science and Technology Repository (King Abdullah University of Science and Technology) | DOI: 10.48550/arxiv.2604.05159

Źródło

[58] "Learning Transferable Visual Models From Natural Language Supervision"

Alec Radford, Jong Wook Kim, Chris Hallacy, Ramesh, Aditya, Gabriel Goh, Ilya Sutskever

2021 | International Conference on Machine Learning | DOI: 10.48550/arxiv.2103.00020

Źródło

[59] "Dropout: a simple way to prevent neural networks from overfitting"

Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov

2014 | Journal of Machine Learning Research

Źródło

[60] "Zero-Shot Text-to-Image Generation"

Ramesh, Aditya, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Ilya Sutskever

2021 | arXiv | DOI: 10.48550/arxiv.2102.12092

Źródło

[61] "GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models"

Alex Nichol, Prafulla Dhariwal, Ramesh, Aditya, Pranav Shyam, Pamela Mishkin, Ilya Sutskever

2022 | International Conference on Machine Learning | DOI: 10.48550/arxiv.2112.10741

Źródło

[62] "Distilling the Knowledge in a Neural Network"

Geoffrey E. Hinton, Oriol Vinyals, Dean, Jeff

2015 | arXiv | DOI: 10.48550/arxiv.1503.02531

Źródło

[63] "A generative model of the covariance structure of images"

Hinton Geoffrey

2010 | Frontiers in Neuroscience | DOI: 10.3389/conf.fnins.2010.03.00037

Źródło

[64] "Managing extreme AI risks amid rapid progress"

Y. Bengio, G. Hinton, A. Yao, D. Song, Pieter Abbeel

2023 | Science | DOI: 10.1126/science.adn0117

Źródło

[65] "OTel: Open Telco AI Datasets, Benchmarks, and Models"

Farbod Tavakkoli, Gregory Diamos, Kenneth Church, Roderic Paulk, Ashish Vaswani

2026 | arXiv | DOI: 10.48550/arxiv.2610.07766

Źródło

[66] "Guest Editorial Introduction to the Special Section on Transformer Models in Vision"

Salman H. Khan, Fahad Shahbaz Khan, A. Vaswani, N. Parmar, Ming-Hsuan Yang

2023 | IEEE Transactions on Pattern Analysis and Machine Intelligence | DOI: 10.1109/tpami.2023.3306164

Źródło

[67] "Data and Parameter Scaling Laws for Neural Machine Translation"

Mitchell A Gordon, Kevin Duh, Jared Kaplan

2021 | Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing | DOI: 10.18653/v1/2021.emnlp-main.478

Źródło

[68] "Explaining neural scaling laws"

Yasaman Bahri, Ethan Dyer, Jared Kaplan, Jaehoon Lee, Utkarsh Sharma

2024 | Proceedings of the National Academy of Sciences | DOI: 10.1073/pnas.2311878121

Źródło

[69] "Training Compute-Optimal Large Language Models"

Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai

2022 | neural information processing systems | DOI: 10.52202/068431-2176

Źródło

[70] "Improving language models by retrieving from trillions of tokens"

Sebastian Borgeaud, Arthur Mensch, Jordan Hoffmann, Trevor Y. Cai, Eliza Rutherford

2022 | International Conference on Machine Learning | DOI: 10.48550/arxiv.2112.04426

Źródło

[71] "Unified Scaling Laws for Routed Language Models"

Aidan Clark, Diego de Las Casas, Aurelia Guy, Arthur Mensch, M. Paganini, Jordan Hoffmann

2022 | International Conference on Machine Learning | DOI: 10.48550/arxiv.2202.01169

Źródło

[72] "Recurrent Independent Mechanisms"

Anirudh Goyal, Alex Lamb, Jordan Hoffmann, Shagun Sodhani, Sergey Levine

2021 | International Conference on Learning Representations | DOI: 10.48550/arxiv.1909.10893

Źródło

[73] **"The Theory of Economic Development"**

Joseph Schumpeter, Ursula Backhaus

2003 | The European Heritage in Economics and the Social Sciences | DOI: 10.1007/0-306-48082-4_3

Źródło

[74] **"The Business Cycle"**

J. Schumpeter

2021 | The Theory of Economic Development | DOI: 10.4324/9781003146766-6

Źródło

[75] **"Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages"**

Carlota Pérez

2002 | Edward Elgar Publishing eBooks | DOI: 10.4337/9781781005323

Źródło

[76] **"Technological revolutions and techno-economic paradigms"**

Carlota Pérez

2009 | Cambridge Journal of Economics | DOI: 10.1093/cje/bep051

Źródło

[77] **"Structural crises of adjustment, business cycles and investment behaviour"**

Carlota Pérez

2008 | Edward Elgar Publishing eBooks | DOI: 10.4337/9781035306176.00008

Źródło

[78] **"The double bubble at the turn of the century: technological roots and structural implications"**

Carlota Pérez

2009 | Cambridge Journal of Economics | DOI: 10.1093/cje/bep028

Źródło

[79] **"The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors"**

John Y. Campbell, Robert J. Shiller

1988 | Review of Financial Studies | DOI: 10.1093/rfs/1.3.195

Źródło

[80] "From Efficient Markets Theory to Behavioral Finance"

Robert J. Shiller

2003 | The Journal of Economic Perspectives | DOI: 10.1257/089533003321164967

Źródło

[81] "Cointegration and Tests of Present Value Models"

John Y. Campbell, Robert J. Shiller

1987 | Journal of Political Economy | DOI: 10.1086/261502

Źródło

[82] "Narratives about Technology-Induced Job Degradations Then and Now"

Robert Shiller

2019 | DOI: 10.3386/w25536

Źródło

[83] "Beyond Computation: Information Technology, Organizational Transformation and Business Performance"

Erik Brynjolfsson, Lorin M. Hitt

2000 | The Journal of Economic Perspectives | DOI: 10.1257/jep.14.4.23

Źródło

[84] "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence"

Timothy F. Bresnahan, Erik Brynjolfsson, Lorin M. Hitt

2002 | The Quarterly Journal of Economics | DOI: 10.1162/003355302753399526

Źródło

[85] "The productivity paradox of information technology"

Erik Brynjolfsson

1993 | Communications of the ACM | DOI: 10.1145/163298.163309

Źródło

[86] "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value1,2"

Lorin M. Hitt, Erik Brynjolfsson

1996 | MIS Quarterly | DOI: 10.2307/249475

Źródło

[87] "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies"

Erik Brynjolfsson, Daniel Joseph Rock, Chad Syverson

2020 | American Economic Journal Macroeconomics | DOI: 10.1257/mac.20180386

Źródło

[88] "Digital Capital and Superstar Firms"

Prasanna B. Tambe, Lorin M. Hitt, Daniel Joseph Rock, Erik Brynjolfsson

2020 | National Bureau of Economic Research | DOI: 10.3386/w28285

Źródło

[89] "Engineering Value: The Returns to Technological Talent and Investments in Artificial Intelligence"

Daniel Joseph Rock

2019 | SSRN Electronic Journal | DOI: 10.2139/ssrn.3427412

Źródło

[90] "Spatial Working Memoryand Planning Ability: Contrasts between Schizophreniaand Bipolar i Disorder"

Johanna Claire Badcock, Patricia Therese Michie, Daniel Joseph Rock

2005 | Cortex | DOI: 10.1016/s0010-9452(08)70294-6

Źródło

[91] "What Determines Productivity?"

Chad Syverson

2011 | Journal of Economic Literature | DOI: 10.1257/jel.49.2.326

Źródło

[92] "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?"

Lucia Foster, John Haltiwanger, Chad Syverson

2008 | American Economic Review | DOI: 10.1257/aer.98.1.394

Źródło

[93] "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics"

Erik Brynjolfsson, Daniel Rock, Chad Syverson

2017 | National Bureau of Economic Research | DOI: 10.3386/w24001

Źródło

[94] **"The Colonial Origins of Comparative Development: An Empirical Investigation"**

Daron Acemoğlu, Simon Johnson, James A. Robinson

2001 | American Economic Review | DOI: 10.1257/aer.91.5.1369

Źródło

[95] **"Robots and Jobs: Evidence from US Labor Markets"**

Daron Acemoğlu, Pascual Restrepo

2019 | Journal of Political Economy | DOI: 10.1086/705716

Źródło

[96] **"The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment"**

Daron Acemoğlu, Pascual Restrepo

2016 | American Economic Review | DOI: 10.1257/aer.20160696

Źródło

[97] **"Automation and New Tasks: How Technology Displaces and Reinstates Labor"**

Daron Acemoğlu, Pascual Restrepo

2019 | The Journal of Economic Perspectives | DOI: 10.1257/jep.33.2.3

Źródło

[98] **"Prospect Theory: An Analysis of Decision under Risk"**

Daniel J. Kahneman, Amos Tversky

1979 | Econometrica | DOI: 10.2307/1914185

Źródło

[99] **"Choices, values, and frames."**

Daniel J. Kahneman, Amos Tversky

1984 | American Psychologist | DOI: 10.1037/0003-066x.39.4.341

Źródło

[100] **"Loss Aversion in Riskless Choice: A Reference-Dependent Model"**

Amos Tversky, Daniel J. Kahneman

1991 | The Quarterly Journal of Economics | DOI: 10.2307/2937956

Źródło

[101] "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias"

Daniel J. Kahneman, Jack L. Knetsch, Richard H. Thaler

1991 | The Journal of Economic Perspectives | DOI: 10.1257/jep.5.1.193

Źródło

[102] "Prospect Theory. An Analysis of Decision Making Under Risk"

Daniel Kahneman, Amos Tversky

1977 | DOI: 10.21236/ada045771

Źródło

[103] "Ironies of automation"

Lisanne Bainbridge

1983 | Automatica | DOI: 10.1016/0005-1098(83)90046-8

Źródło

[104] "Verbal reports as evidence of the process operator's knowledge"

Lisanne Bainbridge

1979 | International Journal of Man-Machine Studies | DOI: 10.1016/s0020-7373(79)80035-8

Źródło

[105] "The change in concepts needed to account for human behavior in complex dynamic tasks"

Lisanne Bainbridge

1997 | IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans | DOI: 10.1109/3468.568743

Źródło

[106] "A Study of Real-Time Human Decision-Making Using a Plant Simulator"

Lisanne Bainbridge, J. Beishon, J. H. Hemming, M. Splaine

1968 | Journal of the Operational Research Society | DOI: 10.1057/jors.1968.12

Źródło

Artykuły naukowe autora:

[107] "Raising AI: How Artificial Intelligence Is Shaping Us—and How We Shape It"

Chirag Shah

2026 | DOI: 10.5040/9798216452430

Źródło

[108] **"Ship It, Then Apologize: We Can Do Better Than This for AI Advancements"**

Chirag Shah

2026 | SSRN Electronic Journal | DOI: 10.2139/ssrn.6952240

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: 8d8cde05-babe-41ce-9c42-102ed3a87c9a