SZTUCZNA INTELIGENCJA W WYMIARZE SPRAWIEDLIWOŚCI

Patrycja Dolniak
Tomasz Kuźma
Andrzej Ludwiński
Konrad Wasik

Wolters Kluwer

SZTUCZNA INTELIGENCJA W WYMIARZE SPRAWIEDLIWOŚCI

Między prawem a algorytmami

Patrycja Dolniak
Tomasz Kuźma
Andrzej Ludwiński
Konrad Wasik

Wolters Kluwer

Warszawa 2024

Stan prawny na 10 sierpnia 2024 r.

Recenzenci
Dr hab. Andrzej Torbus, prof. Uniwersytetu Śląskiego
Dr Damian Flisak

Wydawca
Anna Kubuj-Kacperek

Redaktorka prowadząca
Joanna Cybulska

Opracowanie redakcyjne
Anna Kunz

Projekt okładek serii
Wojtek Janikowski

prawolubni

Ta książka jest wspólnym dziełem twórcy i wydawcy. Prosimy, byś przestrzegał przysługujących im praw. Książkę możesz udostępnić osobom bliskim lub osobiście znanym, ale nie publikuj jej w internecie. Jeśli cytujesz fragmenty, nie zmieniaj ich treści i koniecznie zaznacz, czyje to dzieło. A jeśli musisz skopiować część, rób to jedynie na użytek osobisty.

Szanujmy prawo i własność
Więcej na www.legalnakultura.pl
Polska Izba Książki

© Copyright by Wolters Kluwer Polska Sp. z o.o., 2024

ISBN 978-83-8358-925-1 ISBN PDF-a: 978-83-8390-101-5

Wolters Kluwer Polska Sp. z o.o
Dział Praw Autorskich
01-208 Warszawa, ul. Przyokopowa 33
tel. +48 728 313 462
e-mail: ksiazki@wolterskluwer.pl

księgarnia internetowa www.profinfo.pl

SPIS TREŚCI

Wykaz skrótów 11

Wprowadzenie 17

Rozdział 1
Labirynt AI 21
1. Historia AI – *rollercoaster* sukcesów i niepowodzeń 26
2. Czy maszyny mogą myśleć? 35
3. Gdy oczekiwania wyprzedzają rzeczywistość 44
4. AI – podmiot czy przedmiot? 46
5. Definiowanie niezdefiniowanego 50
6. Rewolucja AI – gdzie jest polski sędzia? 56

Rozdział 2
Prawne i instytucjonalne uwarunkowania AI w Polsce 63
1. Gra o przyszłość: polska strategia dla rozwoju sztucznej inteligencji 63
2. Grupa Robocza ds. Sztucznej Inteligencji (GRAI) 66
3. Ministerstwo Sprawiedliwości a sztuczna inteligencja 70

Rozdział 3
AI w służbie publicznej 73
1. Jak „Arbuz" weryfikuje sprzedawców? 73
2. Algorytmy na lekcji – B3 kontra Rzecznik Praw Obywatelskich 75
3. Kasa bez kasjera 76

Rozdział 4
AI na styku prawa i cyfryzacji 79
1. Jak innowacyjna technologia stała się polem bitwy o prywatność 79
2. Rzeczywistość i fikcja: *deepfake* 86
   2.1. Korzyści i zagrożenia 86
   2.2. Przykłady *deepfake* 89
   2.3. *Deepfake* w postępowaniu sądowym 92
3. *Digital twin* – cyfrowy bliźniak 95
4. Osobliwie obywatelskie preferencje wyborcze AI 103

Rozdział 5
Cywilistyczna wieża Babel: europejskie podejście do AI 107
1. Unia Europejska 107
   1.1. Europejski technologiczny krajobraz regulacyjny 109
   1.2. Unia algorytmów – Akt o sztucznej inteligencji (AI Act) w europejskim stylu 115
   1.3. Nowa dyrektywa w sprawie odpowiedzialności za produkty wadliwe (PLD) 127
   1.4. Dyrektywa w sprawie odpowiedzialności za sztuczną inteligencję (AILD) 131
   1.5. Odpowiedzialność za AI: regulacje kontra rzeczywistość 134
2. Rada Europy 156
   2.1. Europejska Karta Etyczna dotycząca AI w sądach 157
   2.2. Konwencja ramowa o AI, prawach człowieka, demokracji i praworządności 159
3. Systemy AI a prawo międzynarodowe 162
   3.1. Błędna diagnoza czerniaka – kazus I 162
      3.1.1. Jurysdykcja krajowa 163
      3.1.2. Prawo właściwe 168
   3.2. Zawyżenie wyceny nieruchomości – kazus II 175
      3.2.1. Jurysdykcja krajowa 175
      3.2.2. Prawo właściwe 177

Rozdział 6
E-Waga: prawo karne w nowej erze AI 181
1. Materialnoprawne koncepcje odpowiedzialności karnej AI i jej twórców 188

1.1. Czy i jak karać algorytmy? 188
1.2. *Bias* (stronniczość, uprzedzenia) w kontekście działania AI 197
1.2.1. Definicja dyskryminacji 199
1.2.2. Powody powstawania uprzedzeń 199
1.2.3. Uprzedzenia AI a praca sędziego i prokuratora 201
1.3. Kto odpowiada za algorytm? 204
2. Cyberbezpieczeństwo AI 207
3. AI w pracy organów ścigania 211
3.1. Nowoczesna prokuratura i policja 211
3.2. *Predictive policing* 214
3.2.1. Polska 216
3.2.2. Niemcy 216
3.2.3. Niderlandy 217
3.3. JusticeText 217

Rozdział 7
Sędzia i AI: sojusz czy starcie? 221
1. Decyzje zasilane algorytmami: inteligentne wsparcie dla sprawiedliwości 224
2. Narzędzia wyrokowe, narzędzia wspierające decyzje sądu 230
3. Jak AI zmienia pracę sędziego na świecie? 231
3.1. Anonimizacja dokumentów 232
3.2. Przeglądanie dokumentów 234
3.3. Automatyczne rozwiązywanie sporów 238
3.4. Przewidywanie wyników postępowania sądowego 239
3.5. Narzędzia wspomagania i podejmowania decyzji 240
3.6. Składanie pism drogą elektroniczną 245
3.7. Segregacja, przydzielanie i automatyzacja przepływu pracy 247
3.8. Narzędzia do przetwarzania języka naturalnego 251
3.9. Usługi informacyjne i pomocowe 253
3.10. Inne narzędzia 254
4. Gdy zanika granica między AI a niezawisłością sędziego 256
4.1. Chiny 256
4.2. Stany Zjednoczone Ameryki 262
4.3. Izrael 266
5. Brazylijska lekcja dla polskiego wymiaru sprawiedliwości 268
5.1. Justice 4.0 269

5.2. Krajowa platforma cyfrowa sądownictwa 270
5.3. Kodeks: zarządzanie danymi z wykorzystaniem AI 272
5.4. Jasne zasady tworzenia AI 273
5.5. Wspólna platforma i API – Sinapses 274
5.6. Victor, czyli AI w Sądzie Najwyższym 275
5.7. Socrates – rozpoznawanie spraw w blokach i cyfrowa bariera 277
6. Trybunał Sprawiedliwości Unii Europejskiej w epoce AI 278
7. Gdy AI wkracza do Międzynarodowego Trybunału Karnego 282
8. Predykcja rozstrzygnięć ETPC 284

Rozdział 8
Cyfrowa era ADR. Rola AI w mediacji i arbitrażu 287
1. Tajemnicze algorytmy 289
2. Kierować do mediacji czy nie kierować? Oto jest pytanie! 292
2.1. XAI4MA, czyli algorytmy według Włochów 294
3. Modele językowe a mediacja. Koncyliacja za pośrednictwem platform opartych na AI 297
3.1. LLMediator 298
3.2. British Columbia Civil Resolution Tribunal 299
3.3. SmartSettle 300
3.4. Amica 302
3.5. CoParenter 303
3.6. Rechtwijzer 303
3.7. Wyniki badań ankietowych 304
3.8. Podsumowanie 306
4. Sztuczna inteligencja jako wsparcie w trakcie mediacji 307
4.1. AI jako asystent mediatora 308
4.2. Analiza BATNA 310
4.3. Wyniki badań ankietowych 312
5. Emocjonujące emocje 313
5.1. Wyniki badań ankietowych 314
5.2. Zagrożenia 315
5.3. Szanse 318
6. Zamiast pytać „czy?” zapytajmy „jak?” 319
7. Algorytmy w arbitrażu 324
7.1. Sztuczna inteligencja jako narzędzie wspomagające arbitra 325

7.2. Zagrożenia związane z wykorzystaniem AI w arbitrażu 327
7.3. Sztuczna inteligencja w arbitrażu – wyniki badań ankietowych 329
7.4. ENOIK 330
7.5. Ultima Ratio 331
8. Wnioski końcowe 332

Co na to sama zainteresowana? Zakończenie 335

Bibliografia 339

Opinie o książce 349

O autorach 351

WYKAZ SKRÓTÓW

Akty prawne

AIA, AI Act	– rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z 13.06.2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Dz.Urz. UE L 2024/1689)
AIDP	– krajowa biała księga dotycząca strategii Chin w sprawie AI
AILD	– AI Liability Directive, projekt dyrektywy Parlamentu Europejskiego i Rady w sprawie dostosowania przepisów dotyczących pozaumownej odpowiedzialności cywilnej do sztucznej inteligencji (dyrektywa w sprawie odpowiedzialności za sztuczną inteligencję), https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52022PC0496 (dostęp: 26.06.2024)
biała księga w sprawie sztucznej inteligencji, biała księga	– biała księga w sprawie sztucznej inteligencji. Europejskie podejście do doskonałości i zaufania, https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52020DC0065 (dostęp: 26.06.2024)
CRA, unijny akt o cyberodporności	– projekt rozporządzenia Parlamentu Europejskiego i Rady w sprawie horyzontalnych wymogów cyberbezpieczeństwa w odniesieniu do produktów

z elementami cyfrowymi i zmieniającego rozporządzenie (UE) 2019/1020, https://eur-lex.europa.eu/legal-content/PL/TXT/HTML/?uri=CELEX:52022PC0454 (dostęp: 26.06.2024)

CSA – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2019/881 z 17.04.2019 r. w sprawie ENISA (Agencji Unii Europejskiej ds. Cyberbezpieczeństwa) oraz certyfikacji cyberbezpieczeństwa w zakresie technologii informacyjno-komunikacyjnych oraz uchylenia rozporządzenia (UE) nr 526/2013 (akt o cyberbezpieczeństwie) (Dz.Urz. UE L 151, s. 15)

Data Act – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2023/2854 z 13.12.2023 r. w sprawie zharmonizowanych przepisów dotyczących sprawiedliwego dostępu do danych i ich wykorzystywania oraz w sprawie zmiany rozporządzenia (UE) 2017/2394 i dyrektywy (UE) 2020/1828 (akt w sprawie danych) (Dz.Urz. UE L 2023/2854)

DGA – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/868 z 30.05.2022 r. w sprawie europejskiego zarządzania danymi i zmieniające rozporządzenie (UE) 2018/1724 (akt w sprawie zarządzania danymi) (Dz.Urz. UE L 152, s. 1)

DMA – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/1925 z 14.09.2022 r. w sprawie kontestowalnych i uczciwych rynków w sektorze cyfrowym oraz zmiany dyrektyw (UE) 2019/1937 i (UE) 2020/1828 (akt o rynkach cyfrowych), (Dz.Urz. UE L 265, s. 1)

DSA – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z 19.10.2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (akt o usługach cyfrowych) (Dz.Urz. UE L 277, s. 1)

Dyrektywa (UE) 2016/680 – dyrektywa Parlamentu Europejskiego i Rady (UE) 2016/680 z 27.04.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez właściwe organy do celów zapobiegania przestępczości, prowadzenia postępowań przygotowawczych, wykrywania i ścigania czynów

		zabronionych i wykonywania kar, w sprawie swobodnego przepływu takich danych oraz uchylająca decyzję ramową Rady 2008/977/WSiSW (Dz.Urz. UE L 119, s. 89)
Dyrektywa 85/374/EWG	–	dyrektywa Rady z 25.07.1985 r. w sprawie zbliżenia przepisów ustawowych, wykonawczych i administracyjnych Państw Członkowskich dotyczących odpowiedzialności za produkty wadliwe (85/374/EWG) (Dz.Urz. UE L 210, s. 29)
EIDAS	–	rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 910/2014 z 23.07.2014 r. w sprawie identyfikacji elektronicznej i usług zaufania w odniesieniu do transakcji elektronicznych na rynku wewnętrznym oraz uchylające dyrektywę 1999/93/WE (Dz.Urz. UE L 257, s. 73)
EIDAS 2.0	–	rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1183 z 11.04.2024 r. w sprawie zmiany rozporządzenia (UE) nr 910/2014 w odniesieniu do ustanowienia europejskich ram tożsamości cyfrowej PE/68/2023/REV/1 (Dz.Urz. UE L 2024/1183)
EKPC	–	Konwencja o ochronie praw człowieka i podstawowych wolności, sporządzona w Rzymie 4.11.1950 r., zmieniona następnie Protokołami nr 3, 5 i 8 oraz uzupełniona Protokołem nr 2 (Dz.U. z 1993 r. Nr 61, poz. 284 ze zm.)
GPSD, dyrektywa w sprawie ogólnego bezpieczeństwa produktów	–	dyrektywa 2001/95/WE Parlamentu Europejskiego i Rady z 3.12.2001 r. w sprawie ogólnego bezpieczeństwa produktów (Dz.Urz. UE L 11 z 2002 r., s. 4)
GPSR, rozporządzenie w sprawie ogólnego bezpieczeństwa produktów	–	rozporządzenie Parlamentu Europejskiego i Rady (UE) 2023/988 z 10.05.2023 r. w sprawie ogólnego bezpieczeństwa produktów, zmieniające rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1025/2012 i dyrektywę Parlamentu Europejskiego i Rady (UE) 2020/1828 oraz uchylające dyrektywę 2001/95/WE Parlamentu Europejskiego i Rady i dyrektywę Rady 87/357/EWG (Dz.Urz. UE L 135, s. 1)
k.c.	–	ustawa z 23.04.1964 r. – Kodeks cywilny (Dz.U. z 2024 r. poz. 1061 ze zm.)

k.k. – ustawa z 6.06.1997 r. – Kodeks karny (Dz.U. z 2024 r. poz. 17 ze zm.)

k.p.c. – ustawa z 17.11.1964 r. – Kodeks postępowania cywilnego (Dz.U. z 2023 r. poz. 1550 ze zm.)

k.p.k. – ustawa z 6.06.1997 r. – Kodeks postępowania karnego (Dz.U. z 2024 r. poz. 37 ze zm.)

NIS2, dyrektywa NIS2 – dyrektywa Parlamentu Europejskiego i Rady (UE) 2022/2555 z 14.12.2022 r. w sprawie środków na rzecz wysokiego wspólnego poziomu cyberbezpieczeństwa na terytorium Unii, zmieniająca rozporządzenie (UE) nr 910/2014 i dyrektywę (UE) 2018/1972 oraz uchylająca dyrektywę (UE) 2016/1148 (dyrektywa NIS 2) (Dz.Urz. UE L 333, s. 80)

PLD – projekt dyrektywy Parlamentu Europejskiego i Rady w sprawie odpowiedzialności za produkty wadliwe, https://single-market-economy.ec.europa.eu/document/download/3193da9a-cecb-44ad-9a9c-7b6b23220bcd_en?filename=COM_2022_495_1_EN_ACT_part1_v6.pdf (dostęp: 26.06.2024)

RODO – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27.04.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Dz.Urz. UE L 119, s. 1)

Inne skróty

ADR – *Alternative Dispute Resolution* (alternatywne metody rozwiązywania sporów)

AGI – *Artificial General Intelligence* (sztuczna inteligencja ogólna)

AI – *Artificial Intelligence*

ANI – *Artificial Narrow Intelligence* (wąska sztuczna inteligencja)

API – *application programming interface* (interfejs programowania aplikacji)

BATNA – *Best Alternative to Negotiated Agreement* (najlepsza alternatywa dla negocjowanego porozumienia)

CAHAI – Ad hoc Committee on Artificial Intelligence (Komitet *ad hoc* do spraw Sztucznej Inteligencji)
CAI – Committee on Artificial Intelligence (Komisja do spraw Sztucznej Inteligencji)
CAS – Crime Anticipation System (System Przewidywania Przestępczości)
CEPEJ – European Commission for the Efficiency of Justice (Europejska Komisja do spraw Efektywności Wymiaru Sprawiedliwości)
CNJ – Conselho Nacional de Justiça (Krajowa Rada Sprawiedliwości)
CRT – British Columbia Civil Resolution Tribunal
DNA – kwas deoksyrybonukleinowy
DSS – *Decision Support System* (system wspierania decyzji)
DT – *Digital Twin* (cyfrowy bliźniak)
ENISA – European Union Agency for Cybersecurity (Agencja Unii Europejskiej do spraw Cyberbezpieczeństwa)
ETPC – Europejski Trybunał Praw Człowieka
GPAI – Global Partnership on Artificial Intelligence (Globalne Partnerstwo na rzecz Sztucznej Inteligencji)
GRAI – Grupa Robocza do spraw Sztucznej Inteligencji przy Kancelarii Prezesa Rady Ministrów
IDSS – *Intelligent Decision Support System* (inteligentny system wspierania decyzji)
IMY – Integritetsskyddsmyndigheten (szwedzki organ ochrony danych)
IoT – *Internet of Things*, internet rzeczy
KE – Komisja Europejska
LLM – *Large Language Model* (duży model językowy)
ML – *machine learning* (uczenie maszynowe)
MTK – Międzynarodowy Trybunał Karny
NLP – *Natural Language Processing* (przetwarzanie języka naturalnego)
OCR – *Optical Character Recognition* (optyczne rozpoznawanie znaków)
OECD – Organisation for Economic Cooperation and Development (Organizacja Współpracy Gospodarczej i Rozwoju)
PDPJ – Plataforma Digital do Poder Judiciário Brasileiro (Brazylijska Cyfrowa Platforma Sądownictwa)

PE – Parlament Europejski
RPA – *Robotic Process Automation* (zrobotyzowana automatyzacja procesów)
SI – sztuczna inteligencja
STF – Supremo Tribunal Federal (Sąd Najwyższy Brazylii)
STJ – Superior Tribunal de Justiça (Najwyższy Trybunał Sprawiedliwości w Brazylii)
TSUE – Trybunał Sprawiedliwości Unii Europejskiej
UNESCO – United Nations Educational, Scientific and Cultural Organization (Organizacja Narodów Zjednoczonych do spraw Edukacji, Nauki i Kultury)
UOKiK – Urząd Ochrony Konkurencji i Konsumentów
VLOP – *Very Large Online Platforms* (bardzo duże platformy internetowe)
VLOSE – *Very Large Online Search Engines* (bardzo duże wyszukiwarki internetowe)

WPROWADZENIE

Publikacja obejmuje analizę wpływu sztucznej inteligencji (AI) na wymiar sprawiedliwości, poruszając zarówno aspekty technologiczne, jak i prawne oraz etyczne. Jako praktykujący sędziowie, prokuratorzy i przede wszystkim pasjonaci nowoczesnych technologii stawiamy tezę, że prawnicy muszą aktywnie uczestniczyć w kształtowaniu przyszłości polskiego wymiaru sprawiedliwości w erze cyfrowej. W dającej się przewidzieć przyszłości coraz więcej prawników będzie korzystało z technologii, której nie będą w stanie w pełni zrozumieć. Stawiamy także tezę, że sędziowie muszą być kompetentni cyfrowo. Co to znaczy być kompetentnym? Nie zakładamy, że należy uczyć się, jak działa ChatGPT, Copilot, Gemini czy Claude. Zakładamy jednak, że na sędziach spocznie ciężar odpowiedzialności w zakresie stworzenia pomostu między prawem i technologią.

W Europie ponad 90% stanowisk zawodowych wymaga podstawowego poziomu wiedzy cyfrowej, podobnie jak podstawowe umiejętności czytania, pisania i liczenia. Wykorzystanie technologii cyfrowych rozprzestrzenia się we wszystkich sektorach, od biznesu, poprzez transport, rolnictwo, aż po sądownictwo. Jednak około 42% Europejczyków nie ma podstawowych umiejętności cyfrowych, w tym 37% to osoby pracujące[1]. Należy pamiętać, że umiejętności cyfrowe dotyczą nie tylko tego, co jest przydatne w życiu osobistym, ale pomagają również w miejscach pracy.

Od 2014 r. Komisja Europejska monitoruje postępy cyfrowe państw członkowskich za pomocą sprawozdań z indeksu gospodarki cyfrowej i społeczeństwa cyfrowego (tzw. DESI – *Digital Economy and Society*

[1] https://digital-strategy.ec.europa.eu/en/policies/digital-skills (dostęp: 27.06.2024).

Index, indeks gospodarki cyfrowej i społeczeństwa cyfrowego). Sprawozdania DESI 2022 opierają się głównie na danych z 2021 r. i śledzą postępy poczynione w państwach członkowskich UE w dziedzinie technologii cyfrowych. W 2022 r. polski indeks gospodarki cyfrowej i społeczeństwa (około 40) był jednym z najniższych w Unii Europejskiej (średnia ponad 50) i wyprzedzał jedynie Grecję, Bułgarię i Rumunię[2]. Pandemia COVID-19 uwydatniła znaczenie umiejętności cyfrowych i stworzyła wyjątkową okazję do przyspieszenia cyfrowej i technologicznej rewolucji w uczeniu się przez całe życie. Wprowadzenie rozpraw zdalnych w czasie pandemii wstrząsnęło sądami i pokazało, że umiejętności cyfrowe są tak samo ważne jak stosowanie prawa.

Z kolei zgodnie z badaniami przeprowadzonymi w 2023 r. na Uniwersytecie Princeton dotyczącymi „ekspozycji" zawodów na AI, czyli zastępowania ich lub zwiększania ich zależności od AI, zawód sędziego znajdował się na 17. miejscu na 774 zawody objęte analizą[3]. Podkreślono również, że sektor usług prawnych znajduje się na pierwszym miejscu pod względem tego wpływu na analizowane branże[4].

Wobec opisanych wyzwań w rozdziale 1 wprowadzamy czytelnika w świat AI, omawiając jej historię, rozwój oraz podstawowe pojęcia, podkreślając, że choć idea sztucznej inteligencji nie jest nowa, to jej obecny rozwój i zastosowanie są bezprecedensowe. Rozdział ten omawia również różne rodzaje uczenia maszynowego, takie jak uczenie nadzorowane, nienadzorowane i przez wzmacnianie, oraz przedstawia kluczowe technologie, takie jak sieci neuronowe i uczenie głębokie. Zwracamy również uwagę na rosnące oczekiwania wobec AI, podkreślając, że jej rzeczywisty potencjał dopiero się ujawni. W rozdziale 2 skupiamy się na prawnym i instytucjonalnym polskim krajobrazie dotyczącym AI i wysiłkach podejmowanych w kształtowaniu polskiego podejścia do tej technologii.

---

[2] https://digital-strategy.ec.europa.eu/pl/policies/desi (dostęp: 27.06.2024).

[3] E. Felten, M. Raj, R. Seamans, *Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses*, „Strategic Management Journal" 2021/42, iss. 12, s. 1–23.

[4] E. Felten, M. Raj, R. Seamans, *How will Language Modelers like ChatGPT Affect Occupations and Industries?,* https://arxiv.org/ftp/arxiv/papers/2303/2303.01157.pdf (dostęp: 27.06.2024).

Rozdział 3 przedstawia przykłady wykorzystania nowoczesnych technologii w polskim sektorze publicznym, w tym automatyzację i robotyzację, co przynosi realne korzyści w postaci zwiększenia efektywności i oszczędności czasu. Z kolei rozdział 4 prezentuje kontrowersyjne zastosowania AI, porusza problem *deepfake'ów* oraz potencjał i zagrożenia związane z technologią cyfrowego bliźniaka. Następnie rozdział 5 rzuca światło na dynamicznie rozwijający się europejski krajobraz regulacyjny dotyczący AI, ze szczególnym uwzględnieniem AI Act, AILD, PLD oraz problematyki odpowiedzialności za szkody wyrządzone przez AI. W rozdziale 6 znajduje się omówienie koncepcji odpowiedzialności karnej za czyny popełnione przy użyciu AI, w tym trudności związanych z przypisaniem winy i odpowiedzialności w przypadku autonomicznych systemów. Poruszone zostają również kwestie cyberbezpieczeństwa AI. Analizujemy także wykorzystanie AI w pracy organów ścigania. Rozdział 7 omawia rolę AI we wspieraniu pracy sędziów oraz przedstawia przykłady narzędzi i systemów, które mogą pomóc sędziom w analizie spraw, wyszukiwaniu informacji, przygotowywaniu orzeczeń, a nawet przewidywaniu wyników spraw. Omawiamy również doświadczenia różnych krajów, takich jak Brazylia, Chiny, USA i Izrael w zakresie wdrażania AI w sądownictwie. Rozdział 8 jest poświęcony mediacji i arbitrażowi oraz obejmuje analizę możliwości wykorzystania AI w alternatywnych metodach rozwiązywania sporów dokonaną na podstawie przeprowadzonych na potrzeby publikacji badań ankietowych.

Sędziowie będą ponosić odpowiedzialność za bycie ostatecznym gwarantem tego, w jakim kierunku podąży polski wymiar sprawiedliwości. Uważamy bowiem, że przyszłość polskiego wymiaru sprawiedliwości będzie pisana w ciągu następnej dekady. Z niepokojem obserwujemy rosnącą przepaść między twórcami a użytkownikami nowoczesnych technologii, zwłaszcza w dziedzinie prawa. Z jednej strony mamy wizjonerów technologicznych, którzy nie zawsze rozumieją złożoność i niuanse systemu prawnego, z drugiej zaś konserwatywnych prawników, którzy często niechętnie przyswajają nowe narzędzia. Między nimi znajdują się decydenci, którzy, oczarowani blaskiem innowacji, nie zawsze dokonują wyborów przynoszących optymalne rezultaty dla tej drugiej grupy. Ta złożona dynamika rodzi rozwiązania, które nie zawsze spełniają oczekiwania i potrzeby wszystkich zainteresowanych stron.

W niniejszej publikacji stawiamy sobie za cel, aby wypełnić tę niepokojącą nas lukę. Sędziowie muszą uczestniczyć w procesie implementacji technologii, z której niewątpliwie przyjdzie im korzystać w ramach pełnionej służby. Doświadczenia z innych krajów, opisane w dalszej części publikacji, pokazują, że obraz zagranicznych wymiarów sprawiedliwości zmienia się dynamicznie. Technologiczny pociąg powoli odjeżdża, a od nas samych zależy, czy zajmiemy miejsce w którymkolwiek z wagonów. W tym duchu należy odczytywać, co oznacza kompetencja w tym przypadku. Zdajemy sobie sprawę, że jako prawnicy nigdy w pełni nie zrozumiemy technologii i towarzyszących jej dynamicznych zmian. Chcemy jednak, aby efektem naszej pracy było to, że kompetentny sędzia będzie umiał zadać odpowiednie pytania, aby w trakcie pełnionej służby mógł zrozumieć sztuczną inteligencję. Niniejsza publikacja jest pracą zbiorową, a co za tym idzie, czytelnik może odnieść wrażenie, że w niektórych fragmentach dochodzimy do odmiennych wniosków, a nawet sobie przeczymy. Nie jest to zabieg przypadkowy. Naszą pracą staramy się pobudzić ciekawość i skłonić do dyskusji. Jako praktykujący prawnicy i badacze hołdujemy wolności do eksplorowania obszarów prawnych i poruszania się w nieoczekiwanych kierunkach.

W książce posługujemy się zamiennie zwrotami „sztuczna inteligencja" (SI) oraz skrótem AI (*Artificial Intelligence*), ponieważ skrót ten ma wymiar uniwersalny.

Rozdział 1

LABIRYNT AI

Sztuczna inteligencja jako dyscyplina naukowa nie jest zjawiskiem nowym, a jej początki sięgają lat 40. XX wieku (podrozdział 1). Trudności związane z definiowaniem AI i brak jednolitej (podrozdział 2), uniwersalnej definicji powodują, że najlepiej rozumieć AI jako zdolność maszyny do naśladowania ludzkiej inteligencji (podrozdział 5). Rzeczywisty potencjał sztucznej inteligencji dopiero się ujawni (podrozdział 3), a takie kwestie, jak jej podmiotowość prawna, będą przedmiotem zainteresowania praktyków i teoretyków prawa (podrozdział 4). AI już dziś wpływa na rynek pracy, a w tym także na zawód sędziego (podrozdział 6).

Marvin Minsky, kognitywista i jeden z największych pionierów sztucznej inteligencji, ukuł termin „słowa-walizki"[1] dla określeń, które niosą ze sobą wiele różnych znaczeń, nawet jeśli mamy na myśli tylko jedno z nich. Używanie takich terminów zwiększa ryzyko nieporozumień. Na świecie nie wypracowano jednolitej definicji sztucznej inteligencji. Sztuczna inteligencja może być zdefiniowana na wiele sposobów. Należy uznać ją jako termin ogólny (tzw. *umbrella term*), obejmujący różnorodne technologie oparte na algorytmach, które rozwiązują skomplikowane zadania, wykonując funkcje, które wcześniej wymagały ludzkiego myślenia. Sztuczna inteligencja jest dyscypliną naukową, podobnie jak matematyka czy biologia. Pojęcie to obejmuje zbiór koncepcji, problemów i metod ich rozwiązywania. Według słownika PWN[2] sztuczna inteligencja to dziedzina nauki zajmująca się badaniem mechanizmów ludzkiej inteligencji oraz modelowaniem

[1] *Consciousness is a Big Suitcase, A Talk with Marvin Minsky,* https://www.edge.org/3rd_culture/minsky/ (dostęp: 27.06.2024).

[2] https://encyklopedia.pwn.pl/haslo/sztuczna-inteligencja;3983490.html (dostęp: 27.06.2024).

i konstruowaniem systemów, które są w stanie wspomagać lub zastępować inteligentne działania człowieka. Mimo że sztuczna inteligencja istnieje od dłuższego czasu, to niedawne postępy spowodowały, że projektanci AI są w stanie budować systemy zdolne do wykonywania tych skomplikowanych zadań. Za obecnym sukcesem AI stoi przełom w trzech dziedzinach.

Po pierwsze, AI potrzebuje wysokiej jakości bezstronnych **danych** i to w dużych ilościach. Badacze budujący sieci neuronowe wykorzystują duże zbiory danych, które powstały w miarę cyfryzacji społeczeństwa. Jeden z modeli językowych – Co-Pilot, wspierając programistów i podpowiadając dalsze linie kodu, czerpie dane z miliardów linii kodu udostępnianych w GitHub (serwis internetowy przeznaczony do projektów programistycznych). ChatGPT, Gemini, Co-pilot i inne duże modele językowe korzystają z miliardów stron internetowych i dokumentów tekstowych przechowywanych online. Narzędzia zamiany tekstu na obraz, takie jak Stable Diffusion, DALL-E i Midjourney, wykorzystują pary obraz–tekst z gotowych zestawów danych, takich jak LAION-5B (zbiór danych składający się z 5,85 miliarda par obraz–tekst, umożliwiający rozwój badań nad uczeniem maszynowym)[3]. Modele sztucznej inteligencji będą nadal ewoluować pod względem wyrafinowania i wpływu, w miarę jak cyfryzujemy coraz więcej dziedzin naszego życia i zapewniamy im alternatywne źródła danych, takie jak symulowane dane lub dane z ustawień gier, takich jak np. Minecraft.

Po drugie, AI potrzebuje **infrastruktury obliczeniowej** do skutecznego szkolenia. Komputery z roku na rok stają się coraz potężniejsze, zaś modele, które obecnie wymagają intensywnych wysiłków i obliczeń na dużą skalę, mogą w niedalekiej przyszłości być obsługiwane lokalnie. Na przykład Stable Diffusion może już działać na komputerach lokalnych, a nie w środowiskach chmurowych. Według prawa Moore'a (Gordon Moore – współzałożyciel Intela) pamięć i szybkość komputerów podwaja się co roku (zob. pkt 1 w niniejszym rozdziale). Prawo Moore'a to – ściślej rzecz biorąc – obserwacja, że liczba tranzystorów w układzie scalonym podwaja się co dwa lata przy minimalnym wzroście kosztów. Niestety korzystanie z AI wiąże się z potrzebą zużycia kolosalnych ilości energii. W jednym z badań szacuje się, że do 2027 r. serwery AI będą zużywać tyle samo

[3] https://laion.ai/ (dostęp: 25.06.2024).

energii co Argentyna czy Szwecja[4]. Z tego powodu firmy chcą wykorzystywać energię jądrową do wsparcia energochłonnych centrów danych, które odpowiadają za szkolenie i wdrażanie modeli sztucznej inteligencji[5].

Trzecią potrzebą sztucznej inteligencji są ulepszone **modele i algorytmy**. Systemy oparte na danych nadal dokonują szybkich postępów w kolejnych dziedzinach, które kiedyś uważano za obszar wyłącznie ludzkiego poznania. Dobrym przykładem rewolucyjnego zastosowania sztucznej inteligencji jest system opieki zdrowotnej. Technologie oparte na danych zapewniają potencjalnie ważne możliwości poprawy opieki indywidualnej i wspierania innowacji w badaniach medycznych. Kluczową cechą jest ich potencjał do analizowania dużych i złożonych struktur danych w celu tworzenia modeli prognostycznych, które personalizują i poprawiają diagnozę, rokowanie, monitorowanie i administrowanie leczeniem w celu poprawy indywidualnych wyników zdrowotnych[6].

Boom w wykorzystaniu sztucznej inteligencji wynika z połączenia przełomowych technologii, rosnącej dostępności danych i zwiększonej świadomości ich potencjału. Jednakże termin „sztuczna inteligencja" po raz pierwszy został publicznie zaprezentowany przez Johna McCarthy'ego w 1956 r. podczas konferencji organizowanej przez Dartmouth College. Mimo iż od tego momentu upłynął już dłuższy czas, do chwili obecnej nie ma przyjętej powszechnie definicji sztucznej inteligencji. Zadaniem przekraczającym ramy niniejszej publikacji byłoby zebranie wszystkich dostępnych definicji sztucznej inteligencji w jednym miejscu. W kontekście europejskim należy jednak wskazać, że w dniu 13.03.2024 r. Parlament Europejski przyjął tzw. AI Act – rozporządzenie Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze Unii[7]. W przyjętym brzmieniu tego aktu prawnego brakuje definicji sztucznej

---

[4] A. de Vries, *The growing energy footprint of artificial intelligence*, „Joule" 2023/7, iss. 10, s. 2191–2194.

[5] J. Jiang Yu, *The Future of AI Is Nuclear*, https://builtin.com/articles/nuclear-power-artificial-intelligence (dostęp: 27.06.2024).

[6] G.S. Collins, K.G.M. Moons, *Reporting of artificial intelligence prediction models*, „Lancet", April 2019, s. 1577–1579.

[7] COM(2021)206 final.

inteligencji, natomiast w art. 3 pkt 1 rozporządzenia wskazano, że system sztucznej inteligencji („system AI") „oznacza system maszynowy, zaprojektowany do działania z różnym poziomem autonomii, który może po wdrożeniu wykazywać zdolność adaptacji i który – do wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie danych wejściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne"[8]. W pierwotnej wersji rozporządzenia zaproponowano natomiast definicję systemu sztucznej inteligencji jako oprogramowania opracowanego przy użyciu co najmniej jednej spośród technik i podejść wymienionych w załączniku I do rozporządzenia, które może – dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję. Z kolei na stronie internetowej Parlamentu Europejskiego znajduje się definicja sztucznej inteligencji, według której jest to: „zdolność maszyn do wykazywania ludzkich umiejętności, takich jak rozumowanie, uczenie się, planowanie i kreatywność". Jako doprecyzowanie takiej definicji wskazano tam również, że „sztuczna inteligencja umożliwia systemom technicznym postrzeganie ich otoczenia, radzenie sobie z tym, co postrzegają, i rozwiązywanie problemów, działając w kierunku osiągnięcia określonego celu. Komputer odbiera dane (już przygotowane lub zebrane za pomocą jego czujników, np. kamery), przetwarza je i reaguje"[9]. Definicja ta wywodzi się pośrednio z dokumentu *A definition of Artificial Intelligence: main capabilities and scientific disciplines*, w którym grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji opracowała definicję na potrzeby trwających prac kodyfikacyjnych Unii Europejskiej. Grupa ta zaproponowała następującą definicję: „Systemy sztucznej inteligencji to systemy oprogramowania (a być może także sprzętowe) zaprojektowane przez ludzi, które gdy zadane im zostanie złożone zadanie, działające w fizycznym lub cyfrowym wymiarze, postrzegają swoje otoczenie poprzez gromadzenie danych, interpretowanie strukturyzowanych lub nieustrukturyzowanych danych, rozumując w oparciu o wiedzę lub przetwarzanie informacji wywiedzionych z tych danych oraz podejmujące decyzję co do

[8] https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_PL.docx (dostęp: 27.06.2024).

[9] *Sztuczna inteligencja: co to jest i jakie ma zastosowania?*, https://www.europarl.europa.eu/topics/pl/article/20200827STO85804/sztuczna-inteligencja-co-to-jest-i-jakie-ma-zastosowania (dostęp: 27.06.2024).

najlepszego działania w celu osiągnięcia powierzonego jej zadania. Systemy sztucznej inteligencji mogą używać reguł symbolicznych lub uczyć się modeli numerycznych, mogą również dostosowywać swoje działanie na podstawie analizy tego, jaki wpływ na otoczenie miały ich poprzednie zachowania. Jako dyscyplina naukowa, sztuczna inteligencja składa się z szeregu podejść i technik, takich jak uczenie maszynowe (którego przykładem jest uczenie głębokie oraz uczenie wspomagane), wnioskowanie maszynowe (które zawiera planowanie, tworzenie harmonogramów, przedstawienie wiedzy i rozumowanie, wyszukiwanie i optymalizacja) oraz robotyka (na którą składają się kontrola, percepcja, czujniki i siłowniki, jak również integracja wszystkich innych technik w cyber-fizyczne systemy)"[10]. W chwili przygotowania niniejszego opracowania nie udostępniono wspomnianej definicji w języku polskim. Definicja ta jest jednak bardzo rozbudowania i kazuistyczna, dlatego jej wykorzystanie może generować pewne problemy. Z tego powodu istotne jest zaprezentowanie innych, prostszych definicji, które dla osoby niemającej doświadczenia informatycznego będą znacznie bardziej obrazowe i przystępne. Jako jeden z takich przykładów należy wskazać propozycję definicji zaprezentowaną w Tymczasowym Raporcie (*Interim Report*) z listopada 2019 r. przygotowanym przez National Security Commission on Artificial Intelligence (NSCAI). Wskazano tam, że „sztuczną inteligencję należy rozumieć nie jako samodzielne oprogramowanie czy rozwiązanie sprzętowe, a raczej konstelację technologii, która daje systemowi komputerowemu umiejętność do rozwiązywania problemów i wykonywania zadań, które w przeciwnym wypadku wymagałyby ludzkiej inteligencji"[11]. Kluczowe wydaje się odwołanie w tej definicji do elementu ludzkiej inteligencji jako wzorca, który omawiane systemy mają imitować, a który dla laika jest jasny i zrozumiały. Inne definicje sztucznej inteligencji wskazują także na takie elementy, jak „zdolność systemu do prawidłowego interpretowania danych pochodzących z zewnętrznych źródeł, nauki na ich podstawie oraz wykorzystywania tej wiedzy, aby wykonywać określone zadania i osiągać cele poprzez elastyczne dostosowanie"[12]. Najprostszą i przez to jednocześnie

---

[10] *A definition of Artificial Intelligence: main capabilities and scientific disciplines,* https://digital-strategy.ec.europa.eu/pl/node/2226 (dostęp: 27.06.2024).

[11] https://www.nscai.gov/wp-content/uploads/2021/01/NSCAI-Interim-Report-for-Congress_201911.pdf (dostęp: 27.06.2024).

[12] *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*, https://www.sciencedirect.com/science/article/abs/pii/S0007681318301393 (dostęp: 27.06.2024).

nienarażoną na szybką dezaktualizację jest definicja zbudowana na podstawie założeń testu Turinga. Alan Turing, brytyjski matematyk i informatyk, uważany za ojca sztucznej inteligencji, w roku 1950 zaproponował tzw. test Turinga (zob. pkt 1 w niniejszym rozdziale). Testowi miałyby być poddawane maszyny w celu sprawdzenia, czy są zdolne do myślenia w sposób podobny do ludzkiego[13]. Definicja zbudowana na podstawie założeń tego testu stanowi, że „sztuczna inteligencja to zdolność maszyny do naśladowania lub imitowania ludzkiej inteligencji"[14].

Należy mieć na uwadze, że pod pojęciem sztucznej inteligencji można rozumieć zarówno imitowanie ludzkiej inteligencji w procesie technicznym (maszynowym), jak i ogólnie – technologię, która umożliwia wykorzystanie sztucznej inteligencji, oraz – również bardziej ogólnie – dziedzinę naukową, część informatyki.

1. Historia AI – *rollercoaster* sukcesów i niepowodzeń

Historia sztucznej inteligencji obejmuje wiele kluczowych wydarzeń, które przyczyniły się do rozwoju tej dziedziny.

W 1943 r. neurofizjolog Warren McCulloch i matematyk Walter Pitts wydali publikację na temat funkcjonowania ludzkich neuronów, jednocześnie przedstawiając pierwszy matematyczny model sztucznego neuronu, czyli systemu przetwarzającego wartości sygnałów wprowadzanych na jego wejścia w pojedynczą wartość wyjściową, wysyłaną na jego jedynym wyjściu. Sztuczne neurony stanowią podstawę budowy sztucznych sieci neuronowych[15].

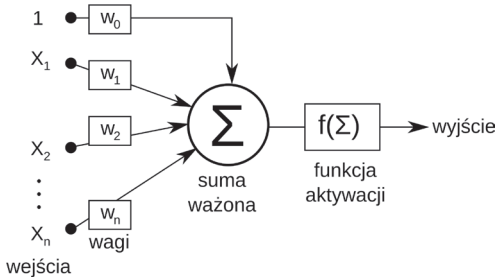
---

[13] P. Łupkowski, *Test Turinga – perspektywa sędziego*, Poznań 2010, s. 41, https://repozytorium.amu.edu.pl/server/api/core/bitstreams/a401990a-4a04-4c89-b102-81ca7c041e55/content (dostęp: 27.06.2024).

[14] T. Zalewski [w:] *Prawo sztucznej inteligencji*, red. L. Lai, M. Świerczyński, Warszawa 2020, nb 1.

[15] Szerzej W.S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity,* „Bulletin of Mathematical Biophysics" 1943/5, s. 115–133.

Rysunek 1.1. Schemat neuronu McCullocha-Pittsa (CC BY-SA 2.5)



Źródło: wersję rastrową rysunku wykonał użytkownik polskiego projektu Wikipedii, Andrut; zwektoryzował Krzysztof Zajączkowski.

W 1950 r. **Alan Turing**, uznawany za ojca sztucznej inteligencji w swojej słynnej pracy *Computing Machinery and Intelligence*, opublikowanej w czasopiśmie filozoficznym „Mind", jako pierwszy zadał pytanie: „czy komputery mogą myśleć?". Opracował również test, uznawany powszechnie za operacyjną definicję inteligencji. Idea testu nawiązywała do starej angielskiej zabawy polegającej na odgadywaniu płci osoby ukrytej w innym pokoju wyłącznie na podstawie pisemnych pytań i odpowiedzi. Turing koncept ten wykorzystał do odróżnienia komputera od człowieka. Osoba testująca dysponowała terminalem składającym się z klawiatury i ekranu, za pomocą którego miała porozumieć się z ukrytym podmiotem: człowiekiem albo komputerem. Zadając pytania, testujący miał rozstrzygnąć, czy testowany podmiot jest człowiekiem, czy programem komputerowym[16]. Pierwsze próby prostej konwersacji pokazywały, że poprzeczka umieszczona jest wysoko. Maszyny podejmowały dialog, ale szybko okazywało się, że nie są w stanie sprostać wyzwaniu. Przyjmuje się, iż test pomyślnie przeszedł dopiero algorytm nazwany Eugene Gootsman w 2014 r., udający nastolatka. Przekonał do siebie 1/3 testujących[17]. Ówczesne realia nie pozwoliły jednak Turingowi kontynuować badań. Komputerom brakowało wówczas kluczowego warunku inteligencji: nie mogły przechowywać poleceń, a jedynie je wykonywać, tzn. komputerom można było powiedzieć, co mają robić, ale nie pamiętały, co zrobiły. Dodatkowo były one w tym

[16] A. Kisielewicz, *Sztuczna inteligencja i logika*, Warszawa 2011, s. 35–36.
[17] A.M. Turing, *Computing Machinery and Intelligence*, „Mind" 1950/49, s. 433–460.

czasie niezwykle drogie[18]. Na początku lat pięćdziesiątych XX wieku koszt leasingu komputera sięgał 200 tys. dolarów miesięcznie. Tylko prestiżowe uniwersytety i duże firmy technologiczne mogły sobie pozwolić na szaleństwo na tych niezbadanych wodach. Potrzeba było zatem czasu, aby sztuczna inteligencja mogła się w pełni rozwinąć.

Latem 1956 r. na kampusie uczelni Dartmouth College zorganizowane zostały warsztaty dla naukowców, na których zebrani stwierdzili, że na zbudowanie maszyn, które będą potrafiły rozumować równie sprawnie jak ludzie, będzie potrzeba więcej niż jednego pokolenia. Założenia te okazały się zbyt optymistyczne. Dopiero aktualnie funkcjonują maszyny potrafiące rozumować matematycznie i logicznie z taką samą skutecznością jak człowiek. Zorganizowaniu konferencji towarzyszyło stwierdzenie, że „badania będą opierać się na założeniu, że każdy aspekt uczenia się albo jakąkolwiek inną cechę inteligencji można w zasadzie opisać tak precyzyjnie, żeby sprawić, by maszyna ją symulowała". W Dartmouth John McCarthy, Marvin Minsky, Allen Newell i Herbert Simon użyli terminu „sztuczna inteligencja" i rozpoczęli badania w tej dziedzinie[19]. Termin ten został stworzony jako zwrot marketingowy w celu uzyskania funduszy na badania[20]. W tym okresie powstały pierwsze programy komputerowe, takie jak Logic Theorist i General Problem Solver, które miały zdolność rozwiązywania problemów na poziomie ludzkim[21].

W latach 1957–1974 sztuczna inteligencja rozkwitła. Komputery mogły przechowywać więcej informacji, stały się szybsze, tańsze i bardziej dostępne. W latach 60. XX wieku rozwinięto systemy eksperckie, które

---

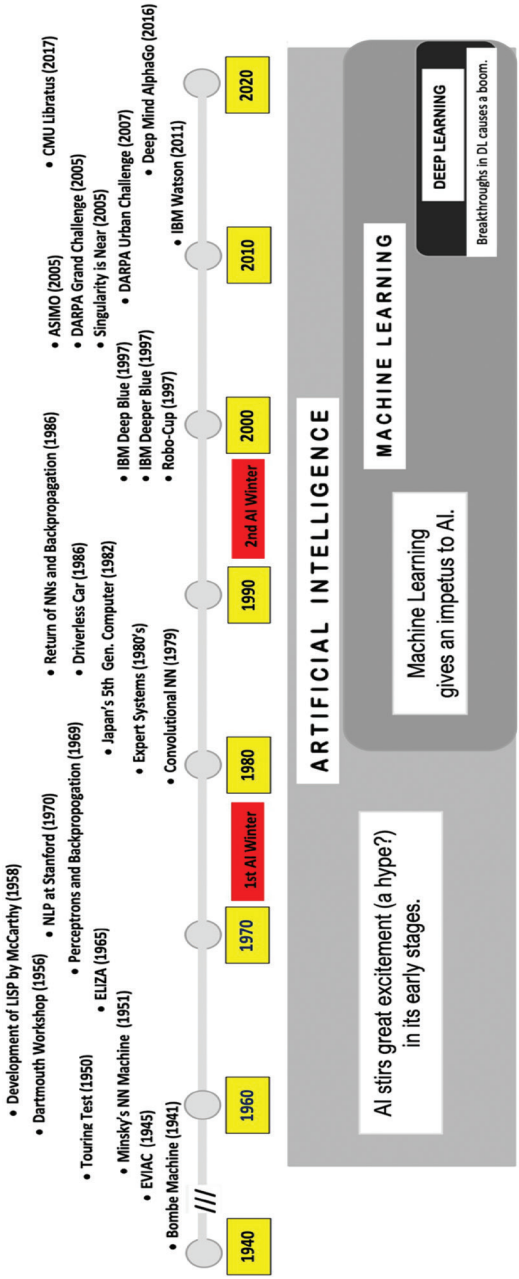
[18] Wczesne komputery: https://ethw.org/Early_Popular_Computers,_1950_-_1970 (dostęp: 27.06.2024).

[19] Wybór terminu „sztuczna inteligencja" dla konferencji w Dartmouth miał na celu odróżnienie od ugruntowanej dziedziny, tj. cybernetyki. W wywiadzie z 2000 roku McCarthy wskazał: „Jeśli chodzi o mnie, to jednym z powodów wymyślenia terminu «sztuczna inteligencja» była chęć uniknięcia skojarzeń z «cybernetyką». Jej skupienie na analogowym sprzężeniu zwrotnym wydawało się mylne" – za: J. Kaplan, *Sztuczna Inteligencja. Co każdy wiedzieć powinien*, Warszawa 2019, s. 30.

[20] M. Tłuczek, *Jak sztuczna inteligencja zmieni twoje życie*, Gliwice 2023, s. 9.

[21] J. Moor, *The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years*, „AI Magazine" 2006/27, no. 4.

Rysunek 1.2. Oś czasu rozwoju AI (CC BY 4.0)



Źródło: Y. Jiang, X. Li, H. Luo, S. Yin, O. Kaynak, *Quo vadis artificial intelligence?*, „Discover Artificial Intelligence" 2022/2, no. 4.

opierały się na wiedzy ekspertów w konkretnych dziedzinach. Były one próbą ograniczenia wymagań obliczeniowych stawianych przez sztuczną inteligencję poprzez wykorzystanie wiedzy ekspertów. Pojawienie się systemów regułowych (*rule-based*), ramowych (*frame-based*), logicznych (*logic-based*) stanowiło pierwsze przydatne i udane wdrożenie sztucznej inteligencji. Z systemami eksperckimi mamy do czynienia do dziś. Przykładem takiego systemu jest np. moduł sprawdzania gramatyki i pisowni w edytorach tekstu. Systemy te osiągnęły tak szerokie zastosowanie, że zostały wchłonięte przez aplikacje, które miały wspomagać[22]. Zaczęto inwestować także środki publiczne do finansowania badań nad sztuczną inteligencją w kluczowych instytucjach obronnych[23].

W 1965 r. ukraiński matematyk Olieksij Grigorowicz Iwachnienko, nazywany ojcem głębokiego uczenia maszynowego, opracował metody grupowego przetwarzania danych (*Group Method of Data Handling* – GMDH) oraz indukcyjnego uczenia statystycznego. Rok później niemiecki informatyk Joseph Weizenbaum, zainspirowany testem Turinga, napisał program – nazwany ELIZA[24] – czyli stworzył chatbota (pojęcie to zostało ukute w 1994 r.) zdolnego przekonać rozmówcę, że jest człowiekiem udającym psychoterapeutę[25]. Program rozpoznawał kluczowe słowa i zwroty w odpowiedziach ludzi i konstruował własne zdania, wykorzystując wcześniej zaprogramowane odpowiedzi. W ten sposób wywoływał wrażenie, że człowiek rozmawia z innym człowiekiem. Program ELIZA był wykorzystywany z powodzeniem do terapii osób z zaburzeniami psychicznymi, a po nim powstawały kolejne chatboty: w 1972 r. – Parry, w 1983 – Racter, a w 1997 – Jabberwacky[26].

---

[22] J.P. Mueller, L. Massaron, *Sztuczna inteligencja*, Gliwice 2021, s. 33–34.

[23] Rząd był szczególnie zainteresowany maszyną, która mogłaby transkrybować i tłumaczyć język mówiony, a także przetwarzać dane o dużej przepustowości: https://livinginternet.com/i/ii_darpa.htm (dostęp: 27.06.2024).

[24] Obecnie program można wciąż testować: https://www.masswerk.at/eliza/ (dostęp: 27.06.2024).

[25] J. Weizenbaum, *ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine*, „Communications of the ACM" 1966/9, iss. 1, s. 36–45.

[26] https://www.sztucznainteligencja.org.pl/portfolio/1966-eliza-bot-udaje-psychoterapeute (dostęp: 27.06.2024).

Rysunek 1.3. Interfejs użytkownika programu ELIZA

```
==============================================================================
          EEEEEEEE  L         IIIIIII  ZZZZZZZ      AAA
          E         L            I         Z      A   A
          E         L            I        Z      A     A
          EEEEE     L            I       Z       A     A
          E         L            I      Z        AAAAAAA
          E         L            I     Z         A     A
          EEEEEEEE  LLLLLLLL  IIIIIII  ZZZZZZ     A     A

==============================================================================
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE... !
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...

==============================================================================
```

Źrodło: https://medium.com/nlp-chatbot-survey/computational-lingustics-754c16fc7355 (dostęp: 27.06.2024).

Przeniknięcie przez początkową mgłę AI odsłoniło masyw przeszkód. Największą był brak mocy obliczeniowej, gdyż komputery nie były w stanie przechowywać wystarczającej ilości informacji ani przetwarzać ich wystarczająco szybko. Przykładowo, aby się porozumieć, trzeba znać znaczenie wielu słów i rozumieć je w wielu kombinacjach. Komputery tego nie potrafiły. W miarę jak malała cierpliwość inwestorów, zmniejszały się fundusze, a badania toczyły się latami. Pojawienie się tzw. zimy AI nastąpiło około lat 70. XX w., kiedy to nie udało się osiągnąć oczekiwanych rezultatów, co prowadziło do spadku zainteresowania i inwestycji w sztuczną inteligencję. Trudno określić, kiedy dokładnie rozpoczęła się i zakończyła zima AI, ponieważ był to proces stopniowy, a różne obszary AI mogły być poddane różnym okresom stagnacji i ożywienia[27].

W latach 80. zainteresowanie AI zaczęło ponownie rosnąć dzięki postępom w dziedzinie uczenia maszynowego i rozwoju algorytmów genetycznych. Edward Feigenbaum wprowadził systemy ekspertowe, które naśladowały proces podejmowania decyzji przez ludzkiego eksperta. Program pytał eksperta w danej dziedzinie, jak zareagować w określonej sytuacji, a gdy

[27] C. Smith, B. McGuire, G. Yang, *The History of Artificial Intelligence*, Washington 2006.

już dowiedział się tego w praktycznie każdej sytuacji, osoby niebędące ekspertami mogły uzyskać porady w ramach tego programu[28]. Systemy ekspertowe znalazły szerokie zastosowanie w przemyśle. W 1986 r. pojawił się system ekspertowy o nazwie DENDRAL, który mógł identyfikować związki chemiczne na podstawie analizy spektroskopowej. Jest on uważany za pierwszy system ekspertowy, ponieważ zautomatyzował proces podejmowania decyzji i rozwiązywania problemów chemii organicznej[29].

W latach 80. i 90. XX wieku powstały również systemy AI do przetwarzania języka naturalnego, takie jak system MYCIN do diagnozowania chorób. Wraz z rosnącą mocą obliczeniową i dostępem do dużych zbiorów danych wzrosło zainteresowanie uczeniem maszynowym i głębokimi sieciami neuronowymi. Co ciekawe, przy braku funduszy rządowych i zainteresowania społecznego sztuczna inteligencja rozkwitła, czego dowodem jest Kismet, opracowany przez Cynthię Breazeal. Kismet to pierwszy robot zaprojektowany, aby angażować ludzi w naturalną i ekspresyjną interakcję twarzą w twarz. Jest on powszechnie uznawany za pionierskie przedsięwzięcie w nowej dziedzinie robotyki społecznej[30].

W latach 90. i na początku XXI wieku osiągnięto wiele przełomowych celów sztucznej inteligencji. W 1997 r. arcymistrz szachowy Gary Kasparow został pokonany przez szachowy program komputerowy Deep Blue firmy IBM. Zwycięstwo w pojedynku składającym się z sześciu gier oznaczało punkt zwrotny w informatyce i zwiastowało przyszłość, w której superkomputery i sztuczna inteligencja będą mogły symulować ludzkie myślenie. Deep Blue czerpał swoje szachowe umiejętności z mocy obliczeniowej 32 procesorów do równoległego wykonywania zestawu skoordynowanych, szybkich obliczeń. Deep Blue był w stanie ocenić 200 milionów pozycji

[28] „E.A. Feigenbaum Papers", *Work in artificial intelligence and computer science at Stanford University*, https://exhibits.stanford.edu/feigenbaum (dostęp: 27.06.2024).

[29] *25th Anniversary of the DENDRAL Project,* Buchanan, Barr, Friedland, Clancy, Bennett, Dietterich, Kunz, Englemore, Session #3, Stanford University, Libraries, Department of Special Collections and University Archives, https://purl.stanford.edu/wr759qp9369 (dostęp: 27.06.2024).

[30] C. Breazeal, *Sociable Machines: Expressive Social Exchange Between Humans and Robots*, Doctoral Dissertation, Department of Electrical Engineering and Computer Science, MIT, https://groups.csail.mit.edu/lbr/hrg/2000/phd.pdf (dostęp: 27.06.2024).

szachowych na sekundę, osiągając prędkość przetwarzania 11,38 miliarda operacji zmiennoprzecinkowych na sekundę (flopów). Dla porównania: pierwszy superkomputer IBM, Stretch, wprowadzony na rynek w 1961 r., osiągał prędkość przetwarzania mniejszą niż 500 flopów[31].

W lutym 2011 r. superkomputer IBM, Watson, rywalizował w teleturnieju *Jeopardy!* (którego polskim odpowiednikiem jest *Va banque*) przeciwko dwóm największym mistrzom wszechczasów teleturnieju, co ugruntowało pozycję AI w świadomości publicznej[32]. Watson zajął pierwsze miejsce z wynikiem 77 147 dolarów. Ken Jennings uzyskał wynik 24 000 dolarów, a Brad Rutter – 21 600 dolarów[33]. Cały odcinek można obejrzeć w serwisie YouTube.

---

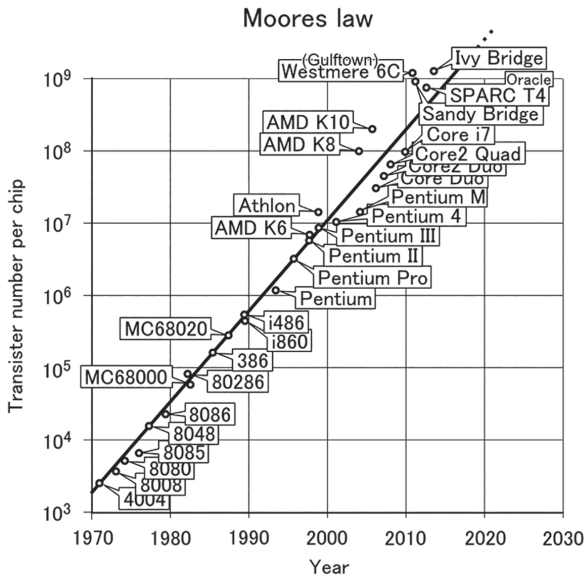
[31] https://www.ibm.com/history/deep-blue (dostęp: 27.06.2024).

[32] Podczas gdy wielkim wyzwaniem związanym z projektem było zwycięstwo w teleturnieju *Jeopardy!*, szerszym celem Watsona było stworzenie technologii nowej generacji, która może skuteczniej znajdować odpowiedzi w nieustrukturyzowanych danych niż standardowa technologia wyszukiwania. Komputery nigdy nie były dobre w znajdowaniu odpowiedzi. Pytania w tym teleturnieju pełne były subtelności, kalamburów i gier słownych – rzeczy, które zachwycają ludzi, ale dławią komputery. „Czym jest czarna śmierć komiwojażera?" jest poprawną odpowiedzią na pytanie w *Jeopardy!*: „Kolorowa czternastowieczna zaraza, która stała się hitem sztuki Arthura Millera". Jedynym sposobem na uzyskanie tej odpowiedzi jest zebranie fragmentów informacji z różnych źródeł, ponieważ dokładnej odpowiedzi raczej nie da się nigdzie zapisać. Watson działał na klastrze komputerów Power 750™ – dziesięciu szafach mieszczących 90 serwerów, łącznie 2880 rdzeni procesorów, na których działa oprogramowanie i pamięć masowa DeepQA. Może pomieścić równowartość około miliona książek. Przez lata Watson był karmiony górami informacji, w tym tekstem ze źródeł komercyjnych, takich jak World Book Encyclopedia, oraz źródeł, które umożliwiają otwarte kopiowanie ich treści, takich jak Wikipedia i książki z Projektu Gutenberg. Kiedy Watsonowi zadawane jest pytanie, ponad 100 algorytmów analizuje je na różne sposoby i znajduje wiele różnych prawdopodobnych odpowiedzi – wszystko to w tym samym czasie. Jeszcze inny zestaw algorytmów klasyfikuje odpowiedzi i przyznaje im punkty. Dla każdej możliwej odpowiedzi Watson znajduje dowody, które mogą potwierdzić lub obalić tę odpowiedź. Tak więc dla każdej z setek możliwych odpowiedzi znajduje setki bitów dowodów, a następnie za pomocą setek algorytmów ocenia stopień, w jakim dowody wspierają odpowiedź. Odpowiedź z najlepszą oceną dowodów zdobędzie największe zaufanie. Najwyżej oceniona odpowiedź staje się odpowiedzią. Jednak podczas gry *Jeopardy!*, jeśli najwyższa możliwa odpowiedź nie została oceniona wystarczająco wysoko, aby dać Watsonowi wystarczającą pewność siebie, Watson postanowił nie wtrącać się i nie ryzykować utraty pieniędzy, jeśli odpowiedź byłaby błędna. Watson robi to wszystko w około trzy sekundy.

[33] https://www.ibm.com/history/watson-jeopardy (dostęp: 27.06.2024).

W ostatnich latach osiągnięto wiele postępów w dziedzinie AI, takich jak rozpoznawanie obrazów, przetwarzanie języka naturalnego, autonomiczne pojazdy i wiele innych zastosowań. Okazuje się, że ograniczenie pamięci komputera, które powstrzymywało postęp przez lata, nie stanowi już problemu. Opisane wcześniej prawo Moore'a (nie jest to prawo naukowe, a pierwotnie było to jedynie stwierdzenie), które szacuje, że pamięć i szybkość komputerów podwaja się co roku, sprawdza się i przeszło do kanonu informatyki. Deep Blue pokonał Kasparowa, Google Alpha Go pokonało chińskiego mistrza gry w Go, Ke Jie. W skrócie można rzec, że nasycamy AI do dostępnego poziomu mocy obliczeniowej, a następnie czekamy, by prawo Moore'a nadrobiło zaległości.

Rysunek 1.4. Wykres logarytmiczny pokazujący oś czasu, na której liczba tranzystorów w mikroczipach niemal podwaja się co dwa lata od 1970 do 2020 r.; prawo Moore'a, Max Roser, Hannah Ritchie (CC BY 4.0)



Źródło: https://commons.wikimedia.org/wiki/File:Moores_law_%281970-2011%29.PNG

2. Czy maszyny mogą myśleć?

Sztuczna inteligencja jest wszędzie i przekształca nasz świat. Żyjemy w czasach *big data* – epoce, w której potrafimy gromadzić ilości danych, które dla człowieka są niemożliwe do przetworzenia. Zastosowanie AI w branżach, takich jak opieka zdrowotna, bankowość, marketing i rozrywka, przynosi znaczne korzyści. Gdyby nie narzędzia oparte na tej technologii, firma Pfizer nie opracowałaby w rekordowym tempie szczepionki na COVID-19[34]. Nawet jeśli postęp w dziedzinie algorytmów nie przyspiesza, to i tak duże zbiory danych i możliwości masowego ich przetwarzania pozwalają sztucznej inteligencji uczyć się za pomocą „metody siłowej" (*brute force*) – metody rozwiązywania problemów polegającej na próbowaniu każdej możliwej opcji aż do znalezienia tej poprawnej[35].

Chociaż niektórzy twierdzą, że prawo Moore'a zwalnia (zob. podrozdział 1 w tym rozdziale), to ilość wytworzonych przez nas danych coraz szybciej się zwiększa. Tak jak energia parowa i mechaniczne silniki zmieniły świat w XVIII wieku, tak technologia sztucznej inteligencji zaczyna zmieniać społeczeństwo i gospodarkę, a także tradycyjne sposoby pracy. Jesteśmy uczestnikami kolejnej rewolucji cywilizacyjnej. Jednakże o tym, jak buduje się systemy AI, decyduje niewielka grupa innowatorów. Ponieważ technologia ta w coraz większym stopniu będzie przekształcać nasze życie, w naszym interesie powinno być jej zrozumienie, choćby na podstawowym poziomie. Warto bowiem zastanowić się, jak będzie wyglądał świat, jeśli jako społeczeństwo nie będziemy poinformowani i wystarczająco zaangażowani, pozostawiając kontrolę nad tym garstce przedsiębiorców i inżynierów, którzy zdecydują, w jaki sposób technologia będzie zmieniać świat.

---

[34] Zob. *mRNA and Artificial Intelligence for Advanced Vaccine Innovation,* https://www.pfizer.com/news/articles/how_a_novel_incubation_sandbox_helped_speed_up_data_analysis_in_pfizer_s_covid_19_vaccine_trial (dostęp: 27.06.2024).

[35] U. Taşova, *The Dictionary of Artificial Intelligence*, Istambuł 2023, s. 56.