

Rozdział I. Analiza pojęcia sztucznej inteligencji – wprowadzenie

§ 1. Sztuczna inteligencja w rozumieniu nauk informatycznych

W celu odniesienia się do wpływu samouczących się algorytmów na problematykę odpowiedzialności kontraktowej i deliktowej w prawie cywilnym w pierwszej kolejności należy dokonać analizy pojęcia sztucznej inteligencji w jej technicznym ujęciu. Dopiero zrozumienie, czym sztuczna inteligencja rzeczywiście jest lub czym powinna być, pozwoli na prowadzenie dalszych rozważań o charakterze już czysto prawnym. Nie oznacza to jednak, że celem niniejszego rozdziału jest wskazanie tylko jednej słusznej definicji. Wręcz przeciwnie, powołanie wielu różnych definicji służy temu, by zwrócić uwagę na te ich elementy, które się powtarzają i być może decydują o istocie pojęcia sztucznej inteligencji. Kolejną kwestię, która będzie rozważana w niniejszym rozdziale, stanowi próba wskazania punktu w czasie, w którym sztuczna inteligencja zaczyna działać samodzielnie, a w końcu – kiedy staje się niezależna od swojego twórcy. Wskazanie bowiem takiego przełomowego momentu będzie miało niebagatelny wpływ na przypisanie odpowiedzialności kontraktowej i deliktowej.

§ 2. Pojęcie inteligencji

Przed wszystkim należy wyjaśnić, czym ogólnie jest „inteligencja”, czy to naturalna, czy też sztuczna. Oczywiście nie będzie możliwe przytoczenie jednej definicji, bowiem taka nie została przez badaczy wypracowana (i nie wiadomo, czy w ogóle wypracowanie jej jest możliwe). Wskazać można na różne konkurujące ze sobą definicje. Zgodnie z encyklopedią PWN pojęcie inteligencji jest jednym z najbardziej niejednoznacznych pojęć występujących w psychologii

i odnosi się do sprawności w zakresie czynności poznawczych¹. Podkreślane jest więc, że mianem inteligencji należy określać zdolność logicznego myślenia, planowania czy rozumienia. Ponadto określane są tak również wiedza emocjonalna, kreatywność, zdolność rozwiązywania problemów, umiejętność uczenia się czy wreszcie samoświadomość². Z kolei *M. Tegmark*, fizyk i kosmolog, tworzy definicję, w której inteligencja ujmowana jest jako zdolność do osiągnięcia złożonych celów. Podkreśla on, że stworzona przez niego definicja jest wystarczająco szeroka, by objąć te wypracowane przez innych badaczy (takie jak stwierdzenie, że inteligencja sprowadza się do samoświadomości lub umiejętności zrozumienia problemów). Zdaniem badacza z istnienia wielu możliwych celów wynika istnienie wielu możliwych rodzajów inteligencji. Należy także dokonać rozróżnienia między inteligencją wąską i inteligencją szeroką. Jako przykład inteligencji wąskiej wskazać można jeden z popularniejszych przykładów pojawiających się w literaturze, jakim jest komputer *Deep Blue*, który w 1997 r. pokonał w grze w szachy ówczesnego mistrza świata *G. Kasparowa*. Oprócz umiejętności gry w szachy (przewyższającej umiejętności *G. Kasparowa*) nie posiadał on jednak żadnych innych zdolności. Z powyższego przykładu wynika, że ludzka inteligencja jest znacznie szersza niż ta komputera, którego umiejętność sprowadza się do jednej wąskiej dziedziny. Szczególnego podkreślenia wymaga więc pojęcie inteligencji ogólnej, która posiada maksymalnie szeroki zakres oraz jest zdolna do osiągnięcia praktycznie każdego celu (łącznie z nauką)³.

§ 3. Początki badań nad SI

Rozpoczynając rozważania na temat sztucznej inteligencji, warto przytoczyć pojęcie testu *Turinga*, przedstawionego przez *A. Turinga* w jego artykule „Computing Machinery and Intelligence”⁴, który został opublikowany w 1950 r. w czasopiśmie „Mind”. Mimo że praktyczne, naukowe i filozoficzne znaczenie tego testu nie jest przedmiotem konsensusu wśród naukowców, należy podkreślić, że otwiera on drogę do badań nad sztuczną inteligencją, skła-

¹ Encyklopedia PWN, dostępna pod adresem: <https://encyklopedia.pwn.pl/haslo/inteligencja;3915042.html>.

² *M. Tegmark*, *Życie 3.0. Człowiek w erze sztucznej inteligencji*, Warszawa 2019, s. 71–72.

³ *Ibidem*, s. 72–75.

⁴ *A. Turing*, *Computing Machinery and Intelligence*, *Mind*, październik 1950, s. 433–460.

nia do zastanawiania się nad jej istotą i zadawania pytań o jej znaczenie. Sama koncepcja testu *Turinga* stanowi nawiązanie do dawnej angielskiej gry. Odgadywano w niej pleć osoby ukrytej w zamkniętym pokoju wyłącznie na podstawie udzielanych przez nią pisemnych odpowiedzi. A. Turing wykorzystał zasady, na których gra bazowała, do odróżnienia komputera od człowieka. Pierwszym pytaniem było: „Czy maszyny mogą myśleć?”. Zdaniem A. Turinga było ono jednak zbyt wieloznaczne, postanowił więc zastąpić je innym. Osoba wykonująca test, mając do dyspozycji klawiaturę i ekran, miała za zadanie porozumiewać się za ich pomocą z ukrytym podmiotem będącym programem komputerowym albo człowiekiem. Tak jak we wspomnianej starej angielskiej grze testujący, który zadawał pytania, miał za zadanie, wyłącznie na podstawie pisemnych odpowiedzi, zdecydować, czy testowany podmiot jest człowiekiem, czy programem komputerowym⁵. A. Turing, komentując zaproponowany przez siebie test, stwierdził, że do roku 2000 zostaną stworzone takie programy komputerowe, że prawdopodobieństwo odgadnięcia przez przeciętną osobę, czy ma do czynienia z komputerem, czy człowiekiem, nie będą przewyższały 70%. Naukowcy są jednak co do zasady zgodni, że przewidywania te się nie sprawdziły. Należy przy tym podkreślić, że test *Turinga* traktowany jest jako operacyjna definicja inteligencji. Wykonując go, nie jesteśmy w stanie stwierdzić z pewnością, czy komputer faktycznie myśli, a jedynie możemy przekonać się o tym, czy komputer zachowuje się tak, jakby myślał. Z testu *Turinga* część naukowców wyciąga też wnioski, że istotą inteligencji jest wykonywanie pewnego algorytmu, czyli ciągu jasno zdefiniowanych czynności koniecznych do wykonania określonych zadań⁶.

Badania nad sztuczną inteligencją były już prowadzone na przełomie lat czterdziestych i pięćdziesiątych ubiegłego stulecia. Wtedy też naukowcy zaczęli zadawać sobie pytanie, czy możliwa jest imitacja przez komputer wybranych procesów intelektualnych występujących u człowieka, takich jak wnioskowanie, dowodzenie postawionych uprzednio hipotez czy nawet rozgrywanie gier. Tego rodzaju działania, w zakresie kształtujących się w tamtych latach nauk informatycznych, zaczęto określać jako inteligencję maszynową czy cybernetykę. W połowie lat pięćdziesiątych pojawiło się natomiast określenie w przeważającym stopniu używane obecnie, czyli sztuczna inteligencja⁷.

⁵ P. Łupkowski, Test Turinga. Perspektywa sędziego, Poznań 2010, s. 7–10.

⁶ A. Kisielewicz, Sztuczna inteligencja i logika, Warszawa 2014, s. 35–39.

⁷ M. Białko, Sztuczna inteligencja i elementy hybrydowych systemów ekspertowych, Koszalin 2005, s. 15.

Przedmiotem zainteresowań tak rozumianej sztucznej inteligencji jest niewątpliwie sposób funkcjonowania umysłu człowieka, uważanego za pewien system przetwarzania różnych docierających do niego informacji. Należy przy tym podkreślić, że umysł człowieka jest zdolny do przetwarzania dwóch rodzajów informacji. Pierwszym są informacje nabyte w procesie dziedziczenia od biologicznych rodziców, które człowiek posiada od urodzenia i które odtworzone są automatycznie, bez zastanowienia (przykładowo będą to wszystkie czynności związane fizjologicznie z ludzkim ciałem, jak poruszanie się). Drugim rodzajem są informacje nabyte przez człowieka w pewnych procesach wykonywanych, takich jak trenowanie, uczenie się poprzez oglądanie, czytanie, słuchanie, dotyk, mowę.

§ 4. Pojęcie SI – przegląd definicji

Podstawą sztucznej inteligencji jest założenie, że czynności wykonywane przez ludzki mózg i system nerwowy stanowią pewien specyficzny system obliczeń. Na poziomie elementarnym zarówno ludzki mózg, jak i komputer są utworzone z podstawowych elementów, którymi w przypadku mózgu są neurony oraz połączenia synaptyczne, a w przypadku komputera – bramki logiczne i ich odpowiednie połączenia. Uznanie, że struktura ludzkiego mózgu i komputera na poziomie elementarnym jest zbliżona (jeśli chodzi o występowanie podstawowych elementów, z których zbudowane są bardziej złożone struktury), pozwala na przyjęcie założenia, że możliwe jest odwzorowanie w programie komputerowym procesów logicznych zachodzących w ludzkim umyśle poprzez wykonywanie specyficznego rodzaju obliczeń⁸. Niektórzy naukowcy, jak Y.N. Harari, podkreślają wręcz, że organizmy żywe w istocie są algorytmami. Jego zdaniem ludzie nie są niepodzielnymi jednostkami, a zbiorem wielu różnych algorytmów, które z kolei kształtowane są przez geny i warunki środowiskowe, a decyzje podejmowane są przez nie w sposób deterministyczny albo przypadkowy⁹.

Celem badaczy zajmujących się konstruowaniem sztucznej inteligencji jest dążenie do zbudowania takich systemów komputerowych, w znacznym stopniu niezależnych od człowieka, które potrafią wykonać – zamiast człowieka – czynności wymagające od niego inteligencji. Według definicji M. Minsky'ego

⁸ *Ibidem*, s. 16.

⁹ Y.N. Harari, *Homo deus. Krótka historia jutra*, Wydawnictwo Literackie 2018, s. 417.

z 1960 r. sztuczna inteligencja jest nauką o maszynach realizujących zadania, które wymagają inteligencji, gdy są wykonywane przez człowieka¹⁰. Inna definicja podkreśla, że pod pojęciem sztucznej inteligencji kryje się wiele różnorodnych zastosowań w informatyce, technice, logice, psychologii, językoznawstwie czy filozofii, których celem jest wykreowanie autonomicznego oraz myślącego robota lub też które mogą mieć zastosowanie do tego¹¹.

Można zauważyć, że niemal każda publikacja odnosząca się do sztucznej inteligencji zawiera inną definicję tego pojęcia. Istnieje w konsekwencji wiele konkurujących ze sobą definicji¹². Znaczna ich część jest przy tym nieprecyzyjna, co może wynikać między innymi z tego, że tak naprawdę nie wiemy (brak jest co do tego powszechnej zgody), czym jest w rzeczywistości sama inteligencja. J. Chromiec i E. Strzemieczna wskazują na trzy konkurujące ze sobą definicje. Zgodnie z pierwszą sztuczna inteligencja to nauka mająca za zadanie nauczyć zachowania podobnego do ludzkiego. Wedle drugiej sztuczna inteligencja to nauka o tym, jak spowodować, aby komputery robiły rzeczy, które aktualnie ludzie robią lepiej. I wreszcie zgodnie z trzecią definicją sztuczna inteligencja to nauka o obliczeniach, które umożliwiają rozumienie, rozumowanie i działanie. Każda z tych definicji w inny sposób opisuje pojęcie sztucznej inteligencji, przy czym w ocenie wskazanych wyżej autorek z tego względu trudno wybrać najlepszą¹³.

Z kolei zdaniem A. Kościółek najbardziej użyteczną definicją, która jednocześnie nie jest skomplikowana technicznie, jest taka, że sztuczna inteligencja polega na opracowywaniu i tworzeniu systemów komputerowych zdolnych do wykonywania zadań, które normalnie wymagają ludzkiej inteligencji¹⁴.

Inny polski badacz, M. Rojszczak, zauważa z kolei, że sztuczna inteligencja nie powinna być zrównywana z tzw. uczeniem maszynowym (*machine learning*). Systemy sztucznej inteligencji powinny być postrzegane jako kolejny etap występujący po uczeniu maszynowym. Wskazał on ponadto, że o ile cechą charakterystyczną systemów ML jest zdolność do odkrywania wiedzy, o tyle

¹⁰ P. Wawrzyński, *Podstawy sztucznej inteligencji*, Warszawa 2014, s. 9.

¹¹ A. Kisielewicz, *Sztuczna inteligencja i logika*, Warszawa 2014, s. 33–34.

¹² K. Różanowski, *Sztuczna inteligencja: rozwój, szanse i zagrożenia*, Zeszyty Naukowe Warszawskiej Wyższej Szkoły Informatyki 2007, Nr 2/2, s. 109–135.

¹³ J. Chromiec, E. Strzemieczna, *Sztuczna inteligencja. Metody konstrukcji i analizy systemów eksperckich*, Warszawa 1994, s. 1–7.

¹⁴ A. Kościółek, *Wykorzystanie sztucznej inteligencji w sądowym postępowaniu cywilnym – zagadnienia wybrane* [w:] K. Flaga-Gieruszyńska, J. Gołaczyński, D. Szostek (red.), *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, Warszawa 2019.

systemy sztucznej inteligencji posiadają ponadto umiejętność podejmowania samodzielnych decyzji na podstawie zgromadzonej przez nie wiedzy. Wnioski formułowane przez systemy ML opierają się natomiast na zdobytych doświadczeniach, a podejmowane decyzje ograniczają się wyłącznie do przewidzianych w pierwotnym kodzie programu (nie występuje więc w tym wypadku swoboda ich powzięcia)¹⁵.

Zgodnie z inną definicją za sztuczną inteligencję należy uznać algorytm, który odbiera informacje z otoczenia i wykorzystuje je w celu podejmowania decyzji, by osiągnąć jakiś określony rezultat. Jako przykłady zastosowania sztucznej inteligencji w praktyce wskazuje się systemy rozpoznawania głosu, opisywanie fabuły filmu, prowadzenie pojazdów, odpowiadanie na pytania, sterowanie robotami przemysłowymi. Tutaj również sztuczna inteligencja odróżniana jest od uczenia maszynowego. Uczenie maszynowe jest bowiem podziedziną sztucznej inteligencji, tak samo jak robotyka. W konsekwencji uczenie maszynowe jest gałęzią informatyki skupiającą się na tworzeniu takiego oprogramowania, które służy do rozpoznawania wzorców bez potrzeby dokonywania jawnego kodowania przez programistę¹⁶. Utożsamianie sztucznej inteligencji i uczenia maszynowego byłoby więc błędem.

W konsekwencji należy stwierdzić, że w powołanych wyżej definicjach sztucznej inteligencji powtarza się między innymi to, iż SI ma imitować ludzkie zachowania czy odwzorowywać procesy logiczne zachodzące w ludzkim umyśle. Ponadto bardzo często wskazywane jest, że istotą SI jest zdolność do podejmowania samodzielnych decyzji (bez ingerencji człowieka) na podstawie zebranej wiedzy. W mojej ocenie, pomimo że być może niemożliwe jest dziś stworzenie jednej niebudzącej wątpliwości definicji SI, to właśnie ta ostatnia cecha sztucznej inteligencji (zdolność do podejmowania samodzielnych decyzji bez ingerencji człowieka) jest istotna z punktu widzenia późniejszych prawnych rozważań.

¹⁵ M. Rojszczak, Prawne aspekty systemów sztucznej inteligencji – zarys problemu [w:] K. Flaga-Gieruszyńska, J. Gołaczyński, D. Szostek (red.), *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, Warszawa 2019.

¹⁶ A. Bassens, G. Beylevend, J. Krohn, *Uczenie głębokie i sztuczna inteligencja. Interaktywny przewodnik ilustrowany*, Gliwice 2022, s. 49–53.

§ 5. Pojęcie SI w dyskursie prawniczym

Pojęcie sztucznej inteligencji pojawia się coraz częściej nie tylko w dyskusjach *stricto* dotyczących konstruowania nowych technologii, ale także w dyskursie prawniczym, zarówno na poziomie narodowym (w szczególności w publicystyce naukowej)¹⁷, jak i ponadnarodowym, jak struktury Unii Europejskiej. Można tu wskazać między innymi rezolucję Parlamentu Europejskiego zawierającą zalecenia dla Komisji w sprawie przepisów prawa cywilnego dotyczących robotyki¹⁸. W rezolucji PE zostały wskazane cechy charakterystyczne inteligentnych robotów, do których należy między innymi uzyskiwanie autonomii za pomocą czujników lub wymiany danych z otoczeniem (wzajemne połączenia) oraz wymiany i analizy tych danych, zdolność samokształcenia na podstawie zdobytego doświadczenia i interakcji z otoczeniem (kryterium fakultatywne), przynajmniej minimalna forma fizyczna, zdolność dostosowywania zachowania i działań do otoczenia, a w końcu brak funkcji życiowych w znaczeniu biologicznym. Również Komisja Europejska w komunikacie z 25.4.2018 r. „Sztuczna inteligencja dla Europy”¹⁹ zdefiniowała pojęcie sztucznej inteligencji. Zgodnie z definicją KE termin „sztuczna inteligencja” odnosi się do systemów, które wykazują inteligentne zachowania dzięki analizie otoczenia i podejmowaniu działań – do pewnego stopnia autonomicznie – w celu osiągnięcia konkretnych celów²⁰. KE podkreśla również, że wiele technologii SI potrzebuje dostępu do danych w celu poprawy swej wydajności. W literaturze

¹⁷ Zob. m.in. E. Kurowska-Tober, L. Czynieńnik, M. Komarska, Aspekty prawne sztucznej inteligencji – zarys problematyki, MoP 2019, Nr 21; A. Auleytner, M.J. Stępień, Dostęp do sztucznej inteligencji – równość i inne aspekty prawne dostępu do sztucznej inteligencji, dodatek do MoP 2019, Nr 21.

¹⁸ Rezolucja Parlamentu Europejskiego z 16.2.2017 r. zawierająca zalecenia dla Komisji w sprawie przepisów prawa cywilnego dotyczących robotyki (2015/2103(INL)) (Dz.Urz. UE C z 2018 r. Nr 252, s. 239).

¹⁹ Komunikat z 25.4.2018 r. Komisji do Parlamentu Europejskiego, Rady Europejskiej, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów „Sztuczna inteligencja dla Europy” (COM(2018) 237, dostępny pod adresem: <https://ec.europa.eu/transparency/reg-doc/rep/1/2018/PL/COM-2018-237-F1-PL-MAIN-PART-1.PDF>).

²⁰ Dalej KE wskazuje, że systemy SI mogą być oparte na oprogramowaniu, działając w świecie wirtualnym (np. asystenci głosowi, oprogramowanie do analizy obrazu, wyszukiwarki, systemy rozpoznawania mowy i twarzy), lub mogą być wbudowane w urządzenia (np. zaawansowane roboty, samochody autonomiczne, drony lub aplikacje Internetu rzeczy).

przedmiotu podkreśla się przy tym, że tworzone definicje powinny być otwarte i jasne, by umożliwiać ich nowe odczytywanie²¹.

§ 6. Prawa SI

W literaturze wskazywane są również trzy prawa sztucznej inteligencji. Zgodnie z pierwszym, nazywanym prawem *Ashby'ego*, każdy skuteczny system kontroli powinien być tak samo złożony jak kontrolowany przez niego system. Zgodnie z drugim prawem, wyrażonym przez *J. von Neumanna*, cechą charakterystyczną złożonego systemu jest to, że system taki sam w sobie jest najprostszym modelem swojego zachowania. Nie jest możliwe zredukowanie systemu do formalnej postaci, gdyż prowadzi to do skomplikowania modelu – nie zaś jego uproszczenia. **I wreszcie w myśl trzeciego prawa każdy system uznany za na tyle prosty, by móc go zrozumieć, nie będzie na tyle złożony, by wykazywać się inteligencją.** W konsekwencji system na tyle złożony, by można było nazwać go inteligentnym, będzie systemem o takim stopniu skomplikowania, że nie będzie można go zrozumieć. Odnosząc się do trzeciego z tych praw, podkreślić trzeba, że możliwe jest stworzenie systemu, którego się nie rozumie i którego działanie – nawet dla samego programisty – będzie stanowiło tajemnicę. Wskazane prawo nie stanowi więc sprzeczności²².

§ 7. Podział na słabą i silną SI

W ramach sztucznej inteligencji wyróżnia się **silną i słabą sztuczną inteligencję**. Pojęcia te zostały wprowadzone do piśmiennictwa w latach osiemdziesiątych ubiegłego stulecia przez *J. Searle'a*. Pierwotnie (w rozumieniu *J. Searle'a*) były one rozumiane inaczej niż współcześnie. U podstaw silnej sztucznej inteligencji stało przekonanie, że już samo wykonywanie algorytmu może wywoływać myślenie. Natomiast jeśli chodzi o słabą sztuczną inteligencję, podkreślano, że system wykonujący algorytm może najwyżej funkcjonować tak,

²¹ *M. Nowakowski*, O moralnej odpowiedzialności HAL-a 9000, czyli etyka sztucznej inteligencji w praktyce. Czy potrzebujemy definicji sztucznej inteligencji?, *Prawo Mediów Elektronicznych* 2022, Nr 1, s. 8.

²² *J. Brockman* (red.), *Człowiek na rozdrożu. Sztuczna inteligencja – 25 punktów widzenia*, Gliwice 2020, s. 55–61.

jakby myślał. Może on nawet przejść pozytywnie test *Turinga*, ale nie oznacza to, że zachodzą w nim jakiekolwiek procesy intelektualne²³. Obecnie **mianem silnej sztucznej inteligencji określa się sztuczne systemy myślące, których poziom intelektualny jest zbliżony do ludzkiego lub nawet go przewyższa**. Nie oznacza to oczywiście, że procesy myślowe człowieka i programu komputerowego powinny być z założenia takie same²⁴. Mogą być jedynie podobne w mniejszym lub większym stopniu, lecz mają prowadzić do tożsamego celu – a mianowicie do produktu w postaci myśli. Poszczególni badacze zajmujący się sztuczną inteligencją wskazują, że SI (w szczególności chodzi tutaj o SI „silną”) nie musi odwzorowywać umysłu ludzkiego. Prawdopodobne jest, że systemy sztucznej inteligencji będą funkcjonować na zupełnie odmiennych zasadach niż umysły organizmów ludzkich. Konsekwencją (na razie dosyć odległą) tych odmienności może być to, że systemy sztucznej inteligencji nie będą potrafiły odzwierciedlić typowo ludzkich uczuć (jak miłość, nienawiść, współczucie czy duma) czy też kierować się nimi²⁵.

Często w ramach rozważań nad silną sztuczną inteligencją przywoływana jest koncepcja samoświadomości, czyli świadomość własnej osobowości, własnych możliwości i odrębności od innych podmiotów. Kwestią sporną jest, czy za atrybut samoświadomości można uznać także tzw. instynkt samozachowawczy, występujący wśród wielu istot żywych. Nie jest też w obecnym stanie nauki możliwe twierdzenie, że instynkt samozachowawczy będzie cechował tzw. silną sztuczną inteligencję. Podstawowym pytaniem, na które aktualnie nie można udzielić jednoznacznej odpowiedzi, jest pytanie o możliwość określenia momentu, w którym dana istota staje się samoświadoma. Już na przykładzie człowieka należy stwierdzić, że nie jest możliwe wskazanie precyzyjnego momentu w jego rozwoju, kiedy uzyskuje on samoświadomość. Na tej podstawie można jednak z całą pewnością rzec, że uzyskiwanie samoświadomości jest procesem, a nie określonym punktem na osi czasu²⁶. Jest również pytaniem otwartym, na które nie można w obecnym stanie nauki udzielić jednoznacznej odpowiedzi, czy osiągalne jest stworzenie przez człowieka autonomicznej maszyny samoświadomej, która potrafiłaby się komunikować w języku natu-

²³ A. Kisielewicz, *Sztuczna inteligencja*, s. 39–40.

²⁴ Zresztą w obecnym stanie nauk biologicznych nie jesteśmy w stanie ze stuprocentową pewnością powiedzieć, jak faktycznie przebiegają procesy myślowe u ludzi, możemy zakładać tylko jakiś stopień prawdopodobieństwa. Ludzki mózg nadal jest przedmiotem intensywnych badań naukowców i wiele procesów w nim zachodzących pozostaje tajemnicą.

²⁵ N. Bostrom, *Superinteligencja. Scenariusze, strategie, zagrożenia*, Gliwice 2014, s. 55.

²⁶ A. Chłopecki, *Sztuczna inteligencja – szkice prawnicze i futurologiczne*, Warszawa 2018, s. 1.

ralnym, miałyby zdolność podejmowania procesów myślowych i na tej podstawie wyciągania logicznych wniosków, rozwiązywania postawionych przed nią problemów, podejmowania decyzji. Byłaby to przy tym maszyna zdolna do uczenia się, rozpoznawania otoczenia i poruszania się po nim. Na określenie takiego rodzaju sztucznej inteligencji często używa się angielskiego sformułowania *human-level artificial intelligence*.

Słaba sztuczna inteligencja dotyczy raczej rozwiązywania pewnych szczegółowych problemów, przy których określone są cele i kryteria ich osiągnięcia. Problemy koncentrują się na realizowaniu przez systemy informatyczne zadań, które są na tyle złożone, że jeśli są wykonywane przez człowieka, to wymagają przeprowadzenia przez niego konkretnych procesów myślowych, wymagają więc od jego wykonawcy inteligencji²⁷.

§ 8. SI a zdolność uczenia się

Główne działy sztucznej inteligencji to wnioskowanie, przeszukiwanie, planowanie i uczenie się. Wnioskowanie polega na mechanicznym wywiedzeniu logicznych konsekwencji przy użyciu reguł wnioskowania dostępnych w logice formalnej²⁸. Przeszukiwanie ma natomiast zastosowanie wszędzie tam, gdzie występują duże przestrzenie. Jego efektem końcowym powinno być odnalezienie najbardziej optymalnego rozwiązania przy jednoczesnym możliwie niewielkim zużyciu pamięci i czasu poświęconego na wykonanie obliczeń. Planowanie stanowi natomiast kombinację mechanizmów automatycznego wnioskowania i rozwiązywania problemów. Efektem planowania powinno być odnalezienie planu rozwiązania problemu, jednak w bardziej efektywny sposób niż poprzez przeszukiwanie, dzięki dostarczeniu systemowi wiedzy o efektach poszczególnych operatorów²⁹.

Przedstawiając problematykę sztucznej inteligencji, należy w sposób szczególny zaznaczyć rolę uczenia się. Nie bez przyczyny w niniejszej pracy podkreśliłam już na wstępie, że przedmiotem badań będą samouczące się algorytmy. Istotą inteligentnych systemów i ich koniecznym atrybutem, czy to naturalnych, czy też sztucznych, jest właśnie zdolność do uczenia się. Zdolność

²⁷ P. Cichosz, *Systemy uczące się*, Warszawa 2000, s. 49–50.

²⁸ Do wnioskowania wykorzystuje się różne systemy formalne, takie jak rachunek predykatów pierwszego rzędu, rachunek zdań i logiki nieklasyczne.

²⁹ P. Cichosz, *Systemy*, s. 50–51.

ta jest szczególnie powiązana ze zdolnością adaptacji, autonomiczności działania i podejmowania racjonalnych decyzji. Istotnym elementem uczenia się jest zmiana, bowiem system, który uczy się czegoś, nie może pozostawać niezmienny w czasie. Jest oczywiste, że zachodzą w nim różnego rodzaju zmiany. Nie może to być jednak zmiana dowolna, bowiem nie każda zmiana może zostać zaklasyfikowana jako uczenie się. Uczeniem się możemy nazwać tylko i wyłącznie taką zmianę, która niesie poprawę, powoduje jakiś progres, nigdy regres – przynajmniej w sensie poznawczym (bo nie zawsze emocjonalnym). Chodzi przede wszystkim o zmianę w ujęciu funkcjonalnym, poprzez poprawę wydajności, sprawności czy skuteczności systemu. Kolejną cechą uczenia się jest **autonomiczność**. **System powinien się ulepszać samodzielnie**, a nie poprzez działania (co nie dotyczy oczywiście dopływu nowych danych) pochodzące z zewnątrz. Nie można powiedzieć, że system uczy się, gdy poprawiany jest przykładowo przez swojego twórcę. Ponadto by autonomiczną progresywną zmianę móc określić jako uczenie się, należy wymagać, by wpływające na zmianę czynniki zewnętrzne można było traktować jako doświadczenie zdobywane przez dany system. Podsumowując, jest to proces polegający na autonomicznie dokonujących się zmianach w systemie na podstawie doświadczeń tego systemu. Proces ten ma na celu poprawę jakości jego funkcjonowania³⁰.

Podczas procesu uczenia się nabywane są informacje. Natomiast zasób zdobytych informacji należy nazwać wiedzą. „W SI podchodzi się do tego problemu pragmatycznie: reprezentacja wiedzy to kombinacja struktur danych i procedur interpretacyjnych tak dobranych, że właściwie użyte prowadzić będą do inteligentnego zachowania”³¹. Ponadto rozróżnić można różne rodzaje wiedzy. Jako pierwszy, podstawowy rodzaj wskazać można wiedzę o obiektach, czyli o przedmiotach (a także istotach żywych) znajdujących się w świecie rzeczywistym: będą to przykładowo kieliszki do wina, psy, ośmiornice czy plansze do gry w ruletkę. Kolejny podtyp stanowi wiedza o zdarzeniach będących działaniami podkreślającymi dynamikę występującą na świecie (opróżnienie kieliszka do wina, szczekanie psa, trzepotanie macek ośmiornicy, ruch na planszy do ruletki). Istnieje także rodzaj wiedzy praktycznej, dotyczący różnego rodzaju umiejętności. Kolejnym podtypem jest tzw. metawiedza, będąca w istocie wiedzą o samej wiedzy. Na koniec wyróżnić należy przekonania. W przeciwieństwie do wiedzy są one niepewne, zazwyczaj subiektywne, mogą być przy tym fałszywe (na przykład przekonanie, że przez przestrzeń kosmiczną pły-

³⁰ *Ibidem*, s. 63.

³¹ W. Duch, *Fascynujący świat komputerów*, Poznań 1997, podrozdział 1.2. Wiedza.

nie olbrzymi gwiazdny żółw niosący na swym grzbiecie cztery ogromne słonie³², które z kolei podtrzymują świat). Podział wiedzy na niniejsze podkategorie (choć niewykluczone jest, że można wyróżnić również inne) powoduje, że informacja może zostać przyporządkowana i zapisana w odpowiednim dla niej miejscu. Dzięki takiej kategoryzacji informacji możliwe do zastosowania są metody przeszukiwania baz danych. Im sprawniej bowiem będzie można znaleźć bądź dopasować daną informację, tym prędzej i efektywniej system będzie mógł na nią zareagować³³.

§ 9. Pojęcie algorytmu

Algorytmem jest „mechaniczny” sposób rozwiązywania danego zadania lub też pewnej klasy zadań, na który składają się pewne uporządkowane i uprzednio zdefiniowane czynności. Metoda ta określona została jako mechaniczna, bowiem nie wymaga od rozwiązującego jakiegokolwiek twórczej inwencji, wymaga natomiast postępowania w sposób ścisły, według wyznaczonej i opisanej szczegółowo instrukcji czynności. Algorytm nie stanowi więc konkretnego obliczenia, ale metodę, która jest stosowana podczas dokonywania obliczeń. Ważna jest przy tym sama kolejność poleceń wykonywanych zgodnie z instrukcją. Polecenia te należy wykonywać po kolei, chyba że na przejście do innego zezwala sama instrukcja. Algorytmy mogą być opisane różnymi metodami, takimi jak mniej lub bardziej precyzyjny opis słowny czy zapisy w różnych językach programowania. Podsumowując, należy podkreślić, że to właśnie na podstawie algorytmu można napisać program komputerowy. Algorytmy stanowią zatem niejako plany programów komputerowych, natomiast sam program komputerowy stanowi szczegółowy zapis algorytmu. Inną kwestią jest to, czy skonstruowany algorytm będzie poprawny i czy doprowadzi do oczekiwanego przez twórcę rezultatu³⁴.

Należy również odróżnić zwykły algorytm od algorytmu uczącego się. Nie zawsze będzie istniała potrzeba, by algorytm był algorytmem samouczącym się. Jednak w przypadku gdy nie będzie możliwe sformułowanie algorytmu w pełni zdefiniowanego, kompletnego i poprawnego, może się pojawić konieczność użycia algorytmu samouczącego się, który przykładowo będzie

³² Zob. T. Pratchett, cykl „Świat Dysku”.

³³ M.J. Kasperski, *Sztuczna inteligencja. Droga do myślących maszyn*, Gliwice 2003, s. 138–139.

³⁴ A. Kisielewicz, *Sztuczna inteligencja*, s. 70–71.

mógł samodzielnie eksperymentować, popełniać błędy i wyciągać z popełnionych błędów wnioski. Jako potrzebę zastosowania samouczącego się algorytmu wskazać tu można przykładowo przewidywaną zmienność otoczenia działania programu komputerowego lub inne okoliczności, w związku z którymi uzyskanie takiego algorytmu okaże się nieefektywne, nieekonomiczne czy nieracjonalne z innych względów. Konieczność zastosowania algorytmów samouczących się może wynikać także z potrzeby stworzenia systemu komputerowego niezależnego od ingerencji człowieka i zdolnego do działania bez jego interwencji. W celu stworzenia systemu komputerowego niezależnego od ingerencji człowieka konieczne jest zapewnienie jego adaptacyjności, czyli zdolności przystosowywania się do zmieniających się warunków środowiskowych. Kolejnym zastosowaniem algorytmów samouczących się jest potrzeba dokonywania operacji na zbiorach danych, które są na tyle obszerne i skomplikowane, że nie jest prawdopodobne uporządkowanie ich w nieautomatyzowany sposób³⁵. Współcześnie wskazywane są różne zastosowania samouczących się algorytmów w praktyce. Przykładowo jednym z kluczowych (i niewątpliwie przyszłościowym) zastosowań jest wykonywanie diagnostyki medycznej³⁶. Ponadto wskazać można zastosowane w motoryzacji, ze szczególnym uwzględnieniem autonomicznych samochodów, w grach planszowych, wykrywaniu nadużyć czy klasyfikacji dokumentacji³⁷. Proces uczenia się i jego definicja zostały już zaprezentowane wcześniej. Należy teraz odnieść te ustalenia do wykorzystywania uczenia się przez algorytmy.

W dyspozycji systemu uczącego się pozostają dwa rodzaje algorytmów. **Pierwszym z nich jest algorytm podstawowy, a drugim – algorytm adaptujący.** Zadaniem algorytmu adaptującego jest dokonywanie zmian w algorytmie będącym algorytmem podstawowym na podstawie wiadomych mu parametrów, czyli w istocie danych, które są w sposób automatyczny uzyskiwane przez system komputerowy z udostępnionego mu zasobu parametrów. Parametry określane są jako „hipoteza”, natomiast ich zasób zwany jest „przestrzenią hipotez”. Wyniki uzyskiwane przez adaptacyjny, samouczący się program komputerowy z definicji są więc niepewne, a sam system informatyczny nie

³⁵ P. Cichosz, *Systemy*, s. 39–40.

³⁶ Bardziej szczegółowo można tu wskazać między innymi wykorzystanie testów diagnostycznych, zaobserwowanych symptomów i ustaleń z wywiadów dotychczasowych pacjentów do diagnozowania kolejnych pacjentów i sugerowania potencjalnych terapii.

³⁷ P. Cichosz, *Systemy*, s. 3–39. Ponadto sztuczna inteligencja może być wykorzystywana przez twórców przy tworzeniu utworów, zob. P. Juściński, *Prawo autorskie w obliczu rozwoju sztucznej inteligencji*, Zeszyty Naukowe Uniwersytetu Jagiellońskiego 2019, Nr 1, s. 5–44.

może zostać uznany za niezawodny, ponieważ zawsze istnieje możliwość pełnienia przez niego błędu w procesie pozyskiwania wyników operacji. Wyniki uzyskane przez algorytm uczący się mogą wymagać wprowadzenia niezbędnych poprawek, nie zmienia to jednak głównej tezy, że uzyskane hipotezy są niezależne i odrębne wobec twórcy programu czy też osoby uruchamiającej program komputerowy³⁸.

Istotnym pojęciem dla funkcjonowania algorytmów jest tzw. teoria złożoności obliczeniowej. Jej główne zadanie stanowi określenie ilości zasobów niezbędnych do rozwiązania problemów obliczeniowych. Zasobami branymi pod uwagę są takie wielkości, jak pamięć, czas czy liczba procesorów.

Jako przykład algorytmu wskazać można wspomnianą już wcześniej maszynę *Turinga*. Język programowania maszyny *Turinga* składa się z jednej dość prymitywnej struktury danych, tj. z ciągu symboli. Do dyspozycji maszyny pozostają operacje, które umożliwiają między innymi przemieszczanie kursora po ciągu – w lewo i prawo – pisanie na bieżącej pozycji ciągu, a także decydowanie o późniejszym kierunku obliczeń na tle wartości bieżącego symbolu. Język programowania maszyny *Turinga* można więc określić jako ubogi i nieskomplikowany. Jednak możliwe jest w nim wyrażenie jakiegokolwiek algorytmu z nieistotną stratą efektywności i symulowanie dowolnego języka programowania. Ponadto maszyny *Turinga* są algorytmami rozwiązującymi problemy związane ze słowami³⁹.

Przedmiotem zainteresowania w niniejszej pracy będą między innymi systemy wspomagania decyzji. Główne dążenie takich systemów stanowi przygotowanie decyzji, która będzie odpowiadać danej sytuacji. Należy zwrócić uwagę na dwa możliwe warianty. W pierwszym przygotowana decyzja będzie wymagała zaaprobowania przez człowieka obeznanego z danym problemem; dopiero wówczas będzie mogła zostać zastosowana. W drugim wariantcie natomiast wypracowana przez system decyzja będzie natychmiast (bez przedstawiania jej do zaaprobowania człowiekowi) wykonywana. Obecnie systemy wspomagania decyzji mają szerokie zastosowanie do programów rozgrywających gry. Należy jednak zauważyć, że po odpowiednim zmodyfikowaniu systemy takie będą mogły mieć zastosowanie przykładowo w wojskowości czy być wykorzystywane do celów związanych z ekonomią. Wykorzystywanie kompu-

³⁸ C.H. Papadimitriou, *Złożoność obliczeniowa*, Warszawa 2002, s. 37–49.

³⁹ *Ibidem*.

terów w celu rozwiązywania czy wspomagania problemów decyzyjnych wymagać będzie wypracowania odpowiednich algorytmów postępowania⁴⁰.

Problem podejmowania samodzielnych działań przez systemy sztucznej inteligencji i uzyskania niezależności od twórcy można zobrazować na przykładzie tzw. systemów ekspertowych. Są to, ogólnie rzecz biorąc, inteligentne programy komputerowe, które są w stanie zastąpić eksperta z danej dziedziny wiedzy. Korzystają one z wiedzy bazowej, wprowadzonej do systemu, i na jej podstawie potrafią wyciągać wnioski. Zaletą systemu ekspertowego jest to, że osoba nieznająca się na danej dziedzinie jest w stanie otrzymać profesjonalną poradę w sposób szybszy, mniej kosztowny i wygodniejszy, niż gdyby korzystała z porad eksperta⁴¹. Jeżeli system ekspertowy korzysta wyłącznie z wiedzy, która została mu przekazana przez twórcę, i na tej podstawie wyciąga wnioski, nie można powiedzieć, że działa on w sposób niezależny od programisty.

Inna sytuacja występuje natomiast, kiedy mamy do czynienia z systemem ekspertowym będącym programem samouczącym się. Twórca w takim przypadku ma wpływ na wiedzę wyjściową, dostarczoną programowi. Jednak taki system dzięki wykorzystywaniu procesu uczenia może zdobywać także nową wiedzę, niedostępną przykładowo dla jego twórcy (twórca może mieć mniejszą wiedzę, a przy tym bardziej ograniczone zdolności poznawcze). Dzięki pozyskiwaniu nowych informacji system ekspertowy będzie miał możliwość jeszcze lepszego rozwiązywania danej kategorii problemów, w sposób niezależny od zasobów wiedzy posiadanych przez programistę. Oczywiście to rolę programisty jest przekazanie systemowi metod rozwiązywania (heurystyk) problemów. Nie zawsze jednak będzie do przewidzenia uzyskane rozwiązanie, w sytuacji gdy komputer będzie dysponował wiedzą, która nie została mu przekazana w postaci wiedzy bazowej przez twórcę. Problem niniejszy jest punktem wyjścia do rozważań, w jakim stopniu twórca sztucznej inteligencji będzie mógł przewidzieć działania SI w sytuacji, gdy będzie miał do czynienia z systemem samouczącym się. Czy będzie mógł przewidzieć (i w jakim stopniu), jakie rozwiązanie zostanie wypracowane przez komputer, który potrafi samodzielnie się doskonalić? Czy będzie znana granica samouczenia się programu komputerowego, do jakiej rozwiązania wdrożone przez program będą przewidywalne? Powyższe pytania i rozważania są w sposób szczególnie istotne dla kwestii od-

⁴⁰ P. Lula, G. Paliwoda-Pękosz, R. Tadeusiewicz, *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Kraków 2007, s. 26–28.

⁴¹ *Ibidem*, s. 36–40.

powiedzialności za działania sztucznej inteligencji i będą wymagały rozwinęcia na gruncie nauk prawnych.

§ 10. Sztuczna inteligencja a automatyzacja

Na zakończenie tych rozważań należy w sposób wyraźny wskazać na **różnicę pomiędzy samouczącymi się algorytmami a postępującą w dzisiejszym świecie automatyzacją**⁴². Zgodnie z definicją z Encyklopedii PWN za automatyzację należy uznać „stosowanie określonych urządzeń i maszyn do zbierania oraz przetwarzania informacji, przejmujących pewne działania poznawcze, intelektualne i decyzyjne człowieka, wykonywane dotąd przez niego w trakcie użytkowania obiektu (np. obrabiarki, samolotu, banku) lub w trakcie prac twórczych (np. projektowania, konstruowania, uczenia)”⁴³. Podkreślić przy tym trzeba, że w przypadku automatyzacji katalog działań, które mogą zostać podjęte przez urządzenie czy maszynę, jest zamknięty, bowiem sytuacje te zostały uprzednio zaprogramowane w sposób skończony. Nie ma więc ryzyka, że działanie podjęte przez maszynę będzie nieznane twórcy i niemożliwe do uprzedniego przewidzenia.

Samouczący się algorytm działa natomiast nie tylko w sposób zautomatyzowany, ale również autonomiczny, a jego cechą wyróżniającą jest to, że działa niezależnie od twórcy i charakteryzuje się między innymi adaptacyjnością, czyli zdolnością do przystosowywania się do zmieniających się warunków środowiskowych. W przypadku samouczącego się algorytmu nie jest zatem możliwe przewidzenie końcowego efektu jego działań. Brak jest zamkniętego katalogu decyzji, które mogą zostać podjęte przez samouczący się algorytm. W konsekwencji odróżnić należy działanie, które jest wyłącznie zautomatyzowane, od działania autonomicznego. Pojęcia te nie mogą być ze sobą zrównywane i stosowane zamiennie. Rozróżnienie to jest szczególnie widoczne w odniesieniu do samochodów autonomicznych. W powstałych klasyfikacjach pojazdów odróżnia się bowiem samochody zautomatyzowane (w części lub w pełni) od tzw. samochodów autonomicznych⁴⁴. Pojazd autonomiczny za-

⁴² T. Kirchberger, European Union policy-making on robotics and artificial intelligence: Selected issues, Croatian Yearbook of European Law and Policy 2017, vol. 13, s. 197–198.

⁴³ Automatyzacja, Encyklopedia PWN: źródło wiarygodnej i rzetelnej wiedzy.

⁴⁴ W. Robaczyński, Czy „ruch pojazdu” powinien być przesłanką odpowiedzialności za pojazd autonomiczny?, PS 2022, Nr 4, s. 7–19; A. Kulickowska, Problematyka uregulowania statusu