

Zbigniew ORBIK

W POSZUKIWANIU FILOZOFICZNYCH ŹRÓDEŁ PROBLEMATYKI SZTUCZNEJ INTELIGENCJI

Rozwój sztucznej inteligencji implikuje problemy natury filozoficznej. Wielu filozofów od czasów starożytnych stawiało pytania istotne z punktu widzenia metodologii sztucznej inteligencji. Dotyczyły one między innymi natury procesów poznawczych, możliwości konstrukcji języków formalnych opisujących w ścisły sposób rzeczywistość czy idei automatyzacji rozumowań. Refleksja filozoficzna wpisana jest niejako w naturę zagadnień pojawiających się w trakcie rozwoju badań nad sztuczną inteligencją.

Pojawienie się sztucznej inteligencji stanowi jeden z najważniejszych momentów w dotychczasowym rozwoju naukowo-technologicznym. Sam termin „sztuczna inteligencja” – mimo że używany nader często w rozmaitych kontekstach – nie doczekał się dotychczas powszechnie akceptowanej definicji. Wynika to głównie z powodu trudności ze sprecyzowaniem pojęcia samej inteligencji. W definicjach podkreśla się najczęściej, że jest to pewna zdolność do przetwarzania informacji na poziomie abstrakcyjnym, adaptacji do nowych warunków otoczenia, uczenia się czy rozwiązywania skomplikowanych zagadnień natury teoretycznej lub praktycznej. Pojęcie inteligencji implikuje zdolność myślenia. We współczesnych ujęciach określenie to rozszerzono poza sferę intelektualną również na sferę emocjonalną. Tak rozumianą inteligencję przypisać można nie tylko człowiekowi, ale także niektórym wyżej rozwiniętym zwierzętom.

Termin „sztuczna inteligencja” (ang. artificial intelligence – AI) po raz pierwszy pojawił się w roku 1955 w projekcie badawczym profesora Dartmouth College Johna McCarthy’ego¹, poświęconym możliwości stworzenia maszyny symulującej ludzką inteligencję. To pojęcie odnosi się do dziedziny informatyki, której celem jest stworzenie inteligentnych, zatem także myślących maszyn. W tradycji filozoficznej inteligencja (intelekt) jest zazwyczaj przeciwstawiana intuicji. Pojęcie sztucznej inteligencji jest w pełni uprawnione o tyle, o ile obdarzonej nią maszynie (komputerowi) przypisać można funkcje myślenia charakteryzujące umysł ludzki – funkcje mentalne. W ten sposób pytanie o istotę sztucznej inteligencji prowadzi do pytania o istotę świadomości.

¹ Por. J. M c C a r t h y, i in., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955*, „AI Magazine” 27(2006) nr 4, s. 12.

W prezentowanym artykule podejmuje się zagadnienie sztucznej inteligencji w aspekcie filozoficznym. Pomimo że jej powstanie datuje się na drugą połowę dwudziestego stulecia, pewne intuicje, które można określić jako „informatyczną preświadomość”², istniały już wcześniej. Przejawiały się one w postaci idei filozoficznych dotyczących możliwości ujęcia rozumowań w postaci symbolicznej, natury uzyskiwanej wiedzy, jej zapisu i przechowywania. Były to, innymi słowy, idee, których realizacja praktyczna stała się możliwa dopiero po osiągnięciu pewnego poziomu rozwoju technologicznego. Warto jednak pamiętać, że postęp ten poprzedzony był wielowiekową historią zmagania się umysłu ludzkiego z najbardziej fundamentalnymi problemami dotyczącymi natury świata i jego poznania oraz możliwościami usprawnienia własnych władz poznawczych. Artykuł ma na celu nakreślenie tej drogi oraz wskazanie na wybrane problemy natury filozoficznej, jakie niesie ze sobą rozwój sztucznej inteligencji.

ZARYS HISTORII SZTUCZNEJ INTELIGENCJI

Pojęcie sztucznej inteligencji pojawiło się po raz pierwszy w połowie dwudziestego wieku. W tamtym okresie ukuto używany do dzisiaj termin, lecz sama idea AI jest o wiele starsza. Teoretycznych początków długiego procesu kształtowania się tej idei, który doprowadził do jej urzeczywistnienia, można poszukiwać już w dociekaniach greckich filozofów nad istotą rozumowania, co zwieńczyły osiągnięcia w dziedzinie logiki i matematyki (między innymi pitagorejczyków, Arystotelesa czy Euklidesa). Grecy myśliciele zapoczątkowali przede wszystkim racjonalne myślenie o świecie, czego wynikiem w późniejszych stuleciach było powstanie nowożytnej nauki i techniki, której jednym z rezultatów są obecnie prowadzone badania nad sztuczną inteligencją. Ich rozważania nad strukturą formalną systemów abstrakcyjnych czynią z nich także twórców idei inicjujących proces, który doprowadził po wielu wiekach rozwoju nauki do badań nad sztuczną inteligencją. Jako pierwsi byli bowiem wyrazicielami poglądu, że każde rozumowanie może zostać ujęte w ramach pewnego rachunku (algorytmu).

Badania nad automatyzacją rozumowań i jej skutkami na polu filozofii prowadził w średniowieczu, żyjący na przełomie trzynastego i czternastego wieku, kataloński filozof, teolog i poeta Rajmundus Lullus³. Trudno tutaj

² Por. P. Staciewicz, *Światopogląd informatyczny. Naukowe podstawy i filozoficzne perspektywy*, w: *Informatyka a filozofia. Od informatyki i jej zastosowań do światopoglądu informatycznego*, red. P. Staciewicz, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2015, s. 13.

³ Por. K. Trzęsicki, *Idea sztucznej inteligencji*, „Filozofia i Nauka” 8 (2020) cz. 1, s. 69-96.

oczywiście mówić o badaniach w znaczeniu ściśle technicznym. Lullus był raczej typem wizjonera-marzyciela, dla którego skonstruowana przez niego maszyna logiczna, zwana „volvelle” (łac. *volvere*, pol. obracać), miała być narzędziem rozstrzygania dysput teologicznych. Volvelle może być traktowana jako fizyczna egzemplifikacja idei logicznych Arystotelesa. Konstrukcja tego urządzenia była stosunkowo prosta. Było ono wykonane z papieru lub pergaminu. Rozstrzygając jakiś problem, odwoływano się do zestawu liczb i odpowiadających im symboli. Przymioty Boga reprezentowane były przez dziewięć odpowiadających im liter. Ich kombinacja stanowiła odpowiedź na postawiony problem teologiczny. Były one umieszczone na jednym lub dwóch ruchomych kołach.

Sama idea maszyny Lullusa stanowiła próbę mechanizacji procesu rozumowania. Warto w tym momencie rozróżnić pojęcia mechanizacji i automatyzacji rozumowań. Pierwsze z nich ma swoje źródło w zapoczątkowanej w średniowieczu idei formalizacji rozumowań, która znalazła pełny wyraz we współczesnych systemach logicznych Gottloba Fregego, Giuseppe Peana, Bertranda Russella, Davida Hilberta, Jana Łukasiewicza, Kurta Gödla, Alfreda Tarskiego i innych. W ramach tego nurtu doszło do algebraizacji logiki w postaci rozwiniętej przez George’a Boole’a algebry binarnej, co przyczyniło się do powstania komputerów, które umożliwiły mechanizację rozumowań dedukcyjnych⁴. Mechanizacja rozumowania, ujmując rzecz najprościej, to po prostu techniczna realizacja dowodzenia sformalizowanego, w której używa się nie długopisu czy ołówek, ale impulsów elektrycznych, a adnotacji dokonuje nie ręka człowieka, lecz określony proces fizyczny. Automatyzację rozumowań należy rozumieć jako zastępowanie pracy człowieka w procesie rozumowania przez działanie maszyny bez udziału człowieka; w tym sensie automatyzacja jest kolejnym etapem następującym po mechanizacji. Lullus uczynił zatem ważny krok zmierzający do skonstruowania maszyny, która w założeniu byłaby zdolna do dokonywania pewnych operacji intelektualnych, w szerokim oczywiście tego słowa znaczeniu. Uprawniony wydaje się więc wniosek, że „Lullus jest tym, od którego można rozpocząć historię idei myślących maszyn, a więc historię idei sztucznej inteligencji”⁵. Znaczenie Lullusa polega także na tym, że oddziałał on na wielu późniejszych myślicieli. Zaliczyć do nich można takich uczonych, jak na przykład Giovanni de la Fontana (ok. 1390-1455/56), Mikołaj z Kuzy (1401-1464), Giordano Bruno (1548-1600), Thomas Hobbes (1588-1679) czy Athanasius Kircher (1602-1680). Jak łatwo dostrzec z powyższego zestawienia, były to postaci wielkiej miary, które wywarły znaczący wpływ na europejską filozofię i naukę.

⁴ Por. W. Marciszewski, R. Murawski, *Mechanization of Reasoning in a Historical Perspective*, Editions Rodopi, Amsterdam–Atlanta 1995, s. 12.

⁵ Trzęsicki, dz. cyt., s. 74.

Każdy z wymienionych intelektualistów wniósł wkład w rozwój idei mechanizacji rozumowań i/ ostatecznie powstanie sztucznej inteligencji. Fontana interesował się nie tylko sztuką, ale i inżynierią. Jednym z jego osiągnięć jest opis konstrukcji maszyny ułatwiającej zapamiętywanie oraz wywody dotyczące sztucznej pamięci⁶. Mikołaj z Kuzy, który był autorem licznych dzieł z zakresu matematyki, wzorując się na pomysłach Lullusa, opracował za pomocą ruchomych diagramów i symboli sztukę wnioskowania w warunkach niepewności (łac. *ars coniecturalis*). Interesujące są jego obserwacje dotyczące umysłu, prowadzone w kontekście rozważań na temat różnic zachodzących między umysłem Boga a umysłem człowieka. Mikołaj studiował również możliwość eksperymentalnego zdobywania wiedzy. Można go uznać za prekursora nowoczesnej nauki⁷. Także Bruno nawiązuje do idei *volvelle* Lullusa. W swoich badaniach dążył do stworzenia sztucznej pamięci, którą później wykorzystywał w retoryce. Hobbes formułuje koncepcję operacji syntaktycznych rozumianych jako pewien rodzaj rachunku. Dochodzi on do wniosku, że rozumowanie jest rachunkiem. Gramatyka języka naturalnego to, w jego ujęciu, rachunek dokonywany na słowach. Hobbes pisze: „Gdy człowiek rozumuje, to nie czyni nic innego, tylko ujmuje pojęciowo (*conceive*) całkowitą sumę, dodając składniki, albo też przedstawia sobie pojęciowo (*conceive*) resztę, odejmując jedną sumę od drugiej. Jeżeli jest to wykonywane na słowach, to jest to przedstawieniem sobie pojęciowym (*conceive*) szeregu nazw wszystkich składników i łączeniem ich w nazwę całości; albo też przejściem od nazw całości i jednego składnika do nazwy innego składnika”⁸.

Mało znany współczesnym naukowcom Kircher był uczonym, jezuitą, a jego prace miały istotny wpływ na rozwój idei, które doprowadziły do powstania sztucznej inteligencji. Początki zainteresowań jego osobą i twórczością datuje się na lata osiemdziesiąte dwudziestego wieku. Określany jest jako „ostatni człowiek, który wiedział wszystko”⁹; był osobowością prawdziwie renesansową. Poszukiwał uniwersalnego języka, rozwijał mnemotechnikę oraz kombinatorykę. Był przekonany, że słowa są rzeczywistością fundamentalną, świat bowiem został stworzony Słowem Bożym. Przedmioty są zatem upo-

⁶ Zob. A.C. Sparravigna, *Giovanni de la Fontana, Engineer and Magician*, <https://arxiv.org/abs/1304.4588>.

⁷ Por. M. Grządziel, *Filozofia, teologia i matematyka w pismach Mikołaja z Kuzy – szkic poglądów*, 2008, s. 6, https://www.chfnp.pl/new05/new05_grzadzziel.pdf. Wyraźnie widoczny jest na przykład wpływ Kuzaińczyka na poglądy Johanna Keplera. Zob. C. Boyer, *Historia rachunku różniczkowego i całkowitego i rozwój jego pojęć*, tłum. S. Dobrzycki, Państwowe Wydawnictwo Naukowe, Warszawa 1964.

⁸ T. Hobbes, *Lewiatan, czyli materia, forma i władza państwa kościelnego i świeckiego*, tłum. C. Znamierowski, Aletheia, Warszawa 2009, s. 118.

⁹ P. Finken, *Athanasius Kircher: The Last Man Who Knew Everything*, Routledge, New York–London 2004.

rządkowane według reguł gramatyki, która określa sposób użycia wyrazów¹⁰. Ten ontologiczny pogląd Kirchera powodował istotne implikacje: jeżeli bowiem między rzeczami a językiem istnieje ściśle przyporządkowanie, wówczas uprawniony staje się pogląd, że dokonując operacji na słowach (pojęciach), jesteśmy w stanie poznawać rzeczywistość. Uczony, stosując opracowaną przez Lullusa metodę ruchomych kół, stosował taką kombinatorykę do pojęć. Jego zamierzenia w tym zakresie uznać należy za maksymalistyczne: „Próbował wytwarzać możliwe kombinacje wszystkich skończonych alfabetów (nie tylko graficznych, lecz również matematycznych)”¹¹. Wymagało to rozwijania umiejętności szyfrowania. W tym celu Kircher projektował automaty i maszyny kodujące i dekodujące, a nawet roboty, które były w stanie poprzez specjalne tuby wydawać dźwięki¹².

Kolejnym myślicielem, który istotnie przyczynił się do powstania sztucznej inteligencji w dwudziestym wieku, był Gottfried Wilhelm Leibniz (1646-1716)¹³. Był on jednym z najwybitniejszych reprezentantów siedemnastowiecznego racjonalizmu oraz jednym z najbardziej wszechstronnych umysłów w dziejach ludzkości. Skonstruował między innymi maszynę liczącą, na której cześć zaprojektował medal z napisem: „*temu, co przewyższa człowieka*”¹⁴.

Leibniz realizował ideę języka uniwersalnego, która była przedmiotem rozważań wielu filozofów siedemnastego wieku. Jak zauważa Adam Nowaczyk: „Formalizacja języków i teorii jest produktem marzeń o zastąpieniu rozumowań rachunkiem”¹⁵. Leibniz w następujący sposób opisuje swoją koncepcję języka uniwersalnego (łac. *characteristica universalis*): „Widać stąd, że gdyby udało się wynaleźć znaki graficzne lub umowne nadające się do wyrażenia naszych myśli równie jasno i ściśle, jak arytmetyka wyraża liczby lub analiza geometryczna linie, to można byłoby w każdej dziedzinie – o ile tylko podlega rozumowaniu – zrobić wszystko, co można zrobić w arytmetyce i geometrii. Albowiem wszystkich dociekań zależnych od rozumowania dokonywano by przez przedstawienie znaków i rodzaj rachunku. [...] Nie trzeba by było tak so-

¹⁰ Por. O. B r e i d b a c h, M.T. G h i s e l i n, *Athanasius Kircher (1602-1680) on Noah's Ark: Baroque „Intelligent Design”*, „Theory: Proceedings of the California Academy of Sciences” 57(2006), nr 36, s. 999.

¹¹ T r z ę s i c k i, dz. cyt., s. 79.

¹² Por. K. V e r m e i r, *Athanasius Kircher's Magical Instruments: An Essay on „Science”, „Religion” and Applied Metaphysics*, „Studies in History and Philosophy of Science Part A” 38(2007) nr. 2, s. 363-400.

¹³ Zob. O. S t e f f e n s, *Leibniz, Lull and the Logic of Truth: Precursors of Artificial Intelligence*, Ostbayerische Technische Hochschule Regensburg, Regensburg 2020.

¹⁴ K. T r z ę s i c k i, *Leibnizjańskie inspiracje informatyki*, „Filozofia Nauki” 14(2006) nr 3(55), s. 43.

¹⁵ A. N o w a c z y k, *Wprowadzenie do filozofii matematyki*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2019, s. 63.

bie łamać głowy, jak jesteśmy do tego zmuszeni dzisiaj, a miałyby się mimo to pewność, że w naszej mocy leży dokonanie wszystkiego, co byłoby wykonalne *ex datis*¹⁶. Chociaż projekt Leibniza nie doczekał się pełnej realizacji, do jego prac odwoływali się kolejni uczeni podejmujący wyzwanie skonstruowania języka uniwersalnego. Jednym z nich był Frege (1848-1925).

Wpływ Fregego na ewolucję idei sztucznej inteligencji widoczny jest przede wszystkim w jego pracach z zakresu logiki. Nie sposób oczywiście omówić tutaj, choćby w zarysie, koncepcji niemieckiego logika. Rewolucyjne znaczenie dla rozwoju logiki miała przede wszystkim jego praca z 1879 roku zatytułowana *Begriffsschrift*¹⁷. Jest ona, z jednej strony, zerwaniem z sylogistyką Arystotelesa, jako jedyną uznawaną do tej pory teorią logiczną, z drugiej – stanowi fundament pod budowę nowoczesnej logiki matematycznej. Frege odrzuca tradycję sylogistyczną, według której zdanie składa się z nazw oraz stosunków zakresowych wyznaczonych przez określone spójniki międzyzdaniami. W swoich badaniach sięga do średniowiecznej koncepcji Tomasza z Akwinu, zgodnie z którą zdanie zbudowane jest z funktorów i ich argumentów¹⁸. Program Fregego (logicyzm) był osłabioną wersją idei Leibniza. Obu filozofom towarzyszyło przekonanie o izomorficzności struktur świata, myśli oraz języka. Znaczenie badań logicznych w kontekście problematyki sztucznej inteligencji polega na tym, że logika jest pewnego rodzaju łącznikiem między językiem naturalnym a językami sformalizowanymi teorii matematycznych¹⁹.

Wkład Leibniza w powstanie sztucznej inteligencji przejawia się także w jego teorii kodowania binarnego. Idea kodu binarnego wiąże się w jego ujęciu z symboliką liczb. Według religii monoteistycznych stworzenie świata polegało na powołaniu go do istnienia przez Boga z nicości. Bóg i nicność są to dwie podstawowe zasady ontologiczne. Wychodząc od tej przesłanki, Leibniz proponuje, aby cyfra jeden stanowiła symbol odwiecznej światłości Boga, cyfra zero natomiast – nicność i ciemność. W swojej koncepcji odwołuje się także do tradycji chińskiej, którą żywo się interesował²⁰. System rachunku

¹⁶ G. W. L e i b n i z, *Przedmowa do nauki ogólnej*, w: tenże, *Wyznanie wiary filozofa*, tłum. S. Cichowicz, Wydawnictwo Naukowe PWN, Warszawa 1969, s. 71n.

¹⁷ Zob. G. F r e g e, *Ideografia. Język formalny czystego myślenia wzorowany na języku arytmetyki*, w: *Próby gramatyki filozoficznej. Antologia*, red. i tłum. K. Rotter, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 1997.

¹⁸ Por. H. J a s t r z ę b s k i, S. C h l e b o w s k i, *Frege i Łukasiewicz o wartościach logicznych*, „Hybris” 27(2014), s. 122.

¹⁹ Por. J. M a c i a s z e k, *Od matematyki do analizy języka naturalnego. Kilka uwag o filozofii języka Gottloba Fregego*, „Językoznawstwo. Współczesne badania, problemy i analizy językoznawcze” 7(2013), s. 58.

²⁰ Por. K. T r z ę s i c k i, *Leibnizjańskie inspiracje informatyki*, „Filozofia Nauki” 14(2006) nr 3(55), s. 31.

binarnego stworzony przez Leibniza ma fundamentalne znaczenie dla informatyki. Nie był on wprawdzie pierwszym filozofem, który zainteresował się tym systemem, lecz jako pierwszy doprowadził go do postaci umożliwiającej wykorzystanie praktyczne²¹. Kazimierz Trzęsicki podkreśla znaczenie Leibniza jako właściwego twórcy systemu binarnego: „System dwójkowy został wynaleziony przez Leibniza około 1679 r., kiedy to napisał *De progressionem dyadica* [...]. W pracy tej Leibniz analizuje możliwości tego systemu oraz pokazuje, jak wykonywać podstawowe operacje arytmetyczne: dodawanie i odejmowanie, mnożenie i dzielenie. W artykule tym Leibniz pisze o maszynie działającej w oparciu o system binarny”²². Mimo że idea systemu binarnego nie znalazła uznania wśród siedemnastowiecznych uczonych, projekty i koncepcje Leibniza przetrwały i okazały się niezwykle płodne. Są one realizowane we współczesnej informatyce wykorzystującej koncepcję maszynowego „myślenia” zapisanego językiem binarnym²³.

Prace skutkujące późniejszym powstaniem sztucznej inteligencji nabrały tempa w dziewiętnastym wieku. Poza wspomnianymi już dokonaniem logikami Fregego kluczowe okazały się dokonania angielskiego matematyka i filozofa George’a Boole’a (1815-1864). Był on pierwszym uczonym, który – w roku 1847 – stworzył język formalny dla rozumowań logicznych. Tak zwana algebra Boole’a jest strukturą algebraiczną, która znalazła zastosowanie w informatyce oraz elektronice cyfrowej²⁴. Istota przekształceń logicznych zaproponowanych przez angielskiego matematyka nawiązuje do koncepcji Leibniza (a wcześniej Arystotelesa), że każde zdanie w sensie logicznym może być wyłącznie prawdziwe bądź fałszywe²⁵. Stosowana obecnie elektronika cyfrowa opiera się na układach logicznych, których operacje wykonywane są na prawach algebry Boole’a²⁶.

Kolejnym ważnym wydarzeniem było zbudowanie przez Alana Turinga w 1936 roku słynnej maszyny nazwanej jego imieniem. Maszyna Turinga to

²¹ Można wskazać na przykład na pitagorejczyka Filolaosa z Tarentu (ok. 480-405 p.n.e.), dla którego podstawową zasadą świata była opozycja skończone–nieskończone, czy indyjskiego poety i matematyka Pingalę (III-II w. p.n.e.). W czasach późniejszych idee systemu binarnego rozważał Thomas Hariot (1560-1621). Por. J.W. Shirley, *Binary Numeration before Leibniz*, „American Journal of Physics” 19(1951) nr 8, s. 452-454.

²² Trzęsicki, *Leibnizjańskie inspiracje informatyki*, s. 27.

²³ Por. tamże, s. 45.

²⁴ Zob. L.M. Laita, L. de Ledesma, E. Roanes-Lozano, A. Brunori, *George Boole, a Forerunner of Symbolic Computation*, w: *Artificial Intelligence and Symbolic Computation: International Conference AISC 2000 Madrid, Spain, July 2000; Revised Papers*, red. J.A. Campbell, E. Roanes-Lozano, Springer, Berlin–Heidelberg 2001, s. 1-19.

²⁵ Por. A. Piecuch, *Zaniedbana algebra a nauczanie informatyki*, „Edukacja – Technika – Informatyka” 8(2017) nr 3, s. 289.

²⁶ Por. tamże, s. 290.

teoretyczny projekt maszyny matematycznej – automatu wykonującego samodzielnie operacje na symbolach, których celem jest znalezienie rozwiązania określonego zadania matematycznego. Założeniem, na którym opierał się pomysł Turinga, było przekonanie, że każde obliczalne (algorytmizowalne, rozstrzygalne) zadanie matematyczne może zostać rozstrzygnięte przez maszynę. Idea maszyny Turinga sformułowana została w jego artykule *On Computable Numbers with an Application to the Entscheidungsproblem*²⁷. Rewolucyjne w pomysle Turinga było wyrażone explicite przekonanie, że możliwa jest mechanizacja procesu obliczeń oraz że może tego dokonać maszyna²⁸. Maszyna Turinga była początkowo tworem wyłącznie teoretycznym. Dopiero w trakcie drugiej wojny światowej oraz po jej zakończeniu, pod wpływem doświadczeń w związku z udziałem w pracach nad łamaniem szyfrów maszyn kodujących, Turing dostrzegł w swoim metamatematycznym projekcie problemy natury praktycznej, dotyczące możliwości skonstruowania maszyny dokonującej operacji rachunkowych. Przyjmuje się, że komputery stanowią fizyczną realizację matematycznej maszyny Turinga. Sam Turing nie był oczywiście filozofem, lecz jego wkład w powstanie sztucznej inteligencji jest znaczący. Kierował się on w swoich dociekaniach filozoficzno-psychologicznymi przesłankami zaczerpniętymi z koncepcji behawioryzmu funkcjonalnego Gilberta Ryle’a (1900-1976).

Niewiele później, w roku 1943 Warren McCulloch i Walter Pitts stworzyli model sztucznych neuronów. A już w roku następnym John von Neumann i Oskar Morgenstern opracowali teorię decyzji (teorię gier)²⁹. Sformalizowali oni klasyczną teorię decyzji, co znalazło zastosowanie nie tylko w informatyce, lecz także w wielu innych naukach, między innymi w ekonomii.

Wszystkie omówione powyżej dokonania stanowiły etapy przygotowawcze. Projekt sztucznej inteligencji ze sfery rozważań teoretycznych przeszedł do etapu urzeczywistnienia w połowie dwudziestego wieku. Trafnie ujął to Bruce G. Buchanan: „Jednak dopiero w ostatnim półwieczu byliśmy w stanie, jako społeczność AI, zbudować maszyny zdolne do testowania hipotez dotyczących mechanizmów myślenia i inteligentnych zachowań, a tym samym zademonstrować mechanizmy istniejące wcześniej jedynie jako teoretyczne możliwości”³⁰. Za rok narodzin sztucznej inteligencji uznaje się 1956, kiedy

²⁷ Por. A. M. T u r i n g, *On Computable Numbers, with an Application to the Entscheidungsproblem*, „Journal of Mathematics” 58(1936) no. 5, ss. 345-363.

²⁸ Por. M. H e t m a n s k i, *Maszyna Turinga a umysł ludzki*, „Annales Universitatis Mariae Curie-Skłodowska. Sectio I. Philosophia-Sociologia” 23(1988), s. 156.

²⁹ Zob. J. v o n N e u m a n n, O. M o r g e n s t e r n, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton 1944.

³⁰ B. G. B u c h a n a n, *A (Very) Brief History of Artificial Intelligence*, „AI Magazine” 26(2005) nr 4, s. 53 (tłum. fragm. – Z.O.).

to McCarthy po raz pierwszy użył tego terminu. Odtąd obserwuje się znacznie żywsze zainteresowanie naukowców reprezentujących różne dziedziny zagadnieniem sztucznej inteligencji³¹.

SZTUCZNA INTELIGENCJA JAKO PROBLEM FILOZOFICZNY

Rozwój sztucznej inteligencji implikuje problemy natury filozoficznej. Są to głównie problemy ontologiczne, epistemologiczne oraz etyczne. Jak już wspomniano, wielu filozofów od czasów starożytnych stawiało pytania istotne z punktu widzenia metodologii sztucznej inteligencji³². Dotyczyły one między innymi natury procesów poznawczych, możliwości konstrukcji języków formalnych opisujących w ścisły sposób rzeczywistość czy idei automatyzacji rozumowań. Do wskazanych już wcześniej filozofów podejmujących tę problematykę dodać można na przykład Williama Ockhama, św. Tomasza z Akwinu czy Kartezjusza.

Wydaje się, że związek filozofii z zagadnieniami sztucznej inteligencji nie jest jedynie akcydentalny. W interesujący sposób kwestię tę ujmuje Michał Heller: „Jednym z głównych powodów nawrotu do «naturalizmu» (można wręcz mówić o «nowym naturalizmie») są ogromne postępy w naukach neurokognitywnych i w dziedzinie tak zwanej sztucznej inteligencji. Działa tu następujący mechanizm napędzający. Z jednej strony postęp w wyżej wspomnianych dziedzinach jest istotnie imponujący; z drugiej jednak strony o ludzkim umyśle wiemy ciągle bardzo mało, a to skłania do «uzupełniania» naukowych wyników filozoficznymi domysłami. Wielu neurofizjologów i specjalistów od sztucznej inteligencji, nie będąc filozofami z wykształcenia, podejmuje próby filozofowania na własną rękę. Utwierdzają ich w tym sporadyczne kontakty (bezpośrednie lub za pośrednictwem lektury) z zawodowymi filozofami, którzy bardzo często okazują się naukowo nieprzygotowani (a więc nieużyteczni) do prowadzenia takiego dialogu. Mówi się już dziś o tak zwanej neurofilozofii; powstaje ona w obszarze mało zagospodarowanym przez tradycyjną filozofię i z natury rzeczy ma ścisły związek z naukami neurokognitywnymi (niekiedy nauki te w sposób prawie niezauważalny przechodzącą w rozważania filozoficzne)”³³. Wydaje się zatem, że refleksja filozoficzna wpisana jest niejako w naturę zagadnień pojawiających się w trakcie rozwoju badań nad sztuczną inteligencją. Można nawet postawić tezę, że źródłem idei sztucznej inteligencji

³¹ Por. V. K a u l, S. E n s l i n, S. A. G r o s s, *History of Artificial Intelligence in Medicine*, „Gastrointestinal Endoscopy” 92(2020) nr 4, s. 807-812.

³² Por. M. F l a s i Ń s k i, *History of Artificial Intelligence*, w: tenże, *Introduction to Artificial Intelligence*, Springer, Cham 2016, s. 3-13.

³³ M. H e l l e r, *Chrześcijański naturalizm*, „Roczniki Filozoficzne” 51(2003) nr 3, s. 44.

jest, obok logiki matematycznej, właśnie filozofia. Pogląd ten wydaje się tym bardziej uzasadniony, jeżeli weźmie się pod uwagę, że przez wieki logika traktowana była jako część filozofii, a dokładniej narzędzie jej uprawiania.

Jednym z problemów natury ontologicznej pojawiających się przy rozważaniu zagadnienia sztucznej inteligencji jest kwestia umysłu człowieka, która, rozpatrywana w tym aspekcie, sprowadzałaby się przede wszystkim do odpowiedzi na pytanie, czy istnieje zasadnicza różnica między umysłem człowieka a inteligentną maszyną. Odpowiedź nie jest, jak mogłoby się na pierwszy rzut oka wydawać, prosta – aby odpowiedzieć na postawione pytanie, należy wcześniej ustalić, jaka jest natura umysłu, a ujmując rzecz ściślej, czy umysł jest zdeterminowany (zaprogramowany) w swoich działaniach, w jakiś sposób analogiczny do sztucznej inteligencji, która działa na podstawie stworzonego przez człowieka programu, czy też operacje umysłu cechuje jakiś rodzaj wolności. Jak dotychczas nie udało się uzyskać jednoznacznej odpowiedzi.

Jedna z przyjmowanych w naukach kognitywnych definicji umysłu stwierdza, że jest on „systemem kontrolnym określającym zachowanie się systemu przy oddziaływaniach ze złożonym, zmiennym w czasie środowiskiem”³⁴. Nie jest to jednak definicja precyzyjna, lecz jedynie relacyjne ujęcie funkcji umysłu, które niewiele, tak naprawdę, mówi o jego naturze. Urszula Żegleń zwraca uwagę na wielowymiarowy charakter koncepcji umysłu. Wielowymiarowość ta implikuje tezę o konieczności zastosowania w badaniach nad istotą umysłu podejścia interdyscyplinarnego. Żegleń utożsamia umysł z ludzkim systemem poznawczym, ale nie z mózgiem. „Przez system poznawczy rozumiem taki system, w którym zachodzą procesy poznawcze, co znaczy, że jest to system zdolny do przyjmowania informacji z zewnątrz i przekształcania ich w celu uzyskiwania (przez system) wiedzy o świecie”³⁵. Zaproponowana perspektywa przeciwstawia się zdecydowanie naturalistycznemu ujęciu umysłu. Naturalizm ten występuje w dwóch zasadniczych postaciach: (1) naturalizmu ontologicznego, głoszącego tezę o istnieniu umysłu w świecie, oraz (2) naturalizmu metodologicznego, zgodnie z którym właściwą metodą badań umysłu jest metoda naukowa. Wobec licznych trudności towarzyszących dotychczasowym próbom zdefiniowania pojęcia umysłu i naukowego opisu zjawisk mentalnych przyjąć można rozumienie umysłu jako „funkcji mózgu”³⁶. W przyjętej w niniejszym artykule perspektywie filozoficznej jego akceptacja nie rozstrzyga sporu naturalizm–antynaturalizm. Stwierdza jedynie związek umysłu z procesami neuronalnymi zachodzącymi w mózgu. Interesującą tezę wysuwa Roger Penrose,

³⁴ W. Duch, *Umysł, świadomość i działanie twórcze*, „Kognitywistyka i Media w Edukacji” 2008, nr 1-2, s. 10.

³⁵ U. Żegleń, *Wielowymiarowość umysłu*, „Filozofia Nauki” 12(2004) nr 3-4, s. 154.

³⁶ Duch, dz. cyt., s. 7.

według którego umysł (świadomość) nie daje się wyjaśnić na gruncie współczesnych neuronauk³⁷. Zdaniem brytyjskiego fizyka procesy mentalne mogą zostać wyjaśnione dopiero na poziomie kwantowym.

Wacław Branicki, referując toczącą się w powyższej sprawie dyskusję, zauważa, że założenie, że różnica między inteligentną maszyną a człowiekiem jest jedynie różnicą techniczną, związaną ze stanem rozwoju sztucznej inteligencji, zdaje się wątpliwe. „Przedstawione tutaj rozumowanie – pisze Branicki – wydaje się spójne, jednak warto pamiętać, że opiera się ono na dyskusyjnym założeniu definiującym ludzką podmiotowość jako zbiór schematów uwarunkowanych biologicznie, społecznie i językowo. Z perspektywy dezintegracji pozytywnej rysuje się inna hipoteza. Zakłada ona, że taki redukcjonistyczny obraz człowieka nie znajduje potwierdzenia w badaniach empirycznych i jest przykładem błędu logicznego typu *pars pro toto*”³⁸.

Dyskusyjne są także inne kwestie, jak zdolność zdobywania wiedzy czy wręcz jej tworzenia, charakteryzująca umysł ludzki – ale czy również sztuczną inteligencję? Pozytywnej odpowiedzi w tym względzie udziela między innymi Ryszard Tadeusiewicz, który uważa, że komputer jest w stanie generować nową wiedzę³⁹. Rodzi się pytanie o naturę tej wiedzy. Czy sztuczna inteligencja potrafi tworzyć wiedzę nową jakościowo? Czy wiedza ta ogranicza się wyłącznie do nauk formalnych, czy też maszyny zdolne są już do odkrywania wiedzy o świecie transcendentnym? Wydaje się, że obecnie maszyny tworzą jedynie wiedzę funkcjonalną. Nauki kognitywne wkroczyły dopiero na drogę swojego rozwoju. Być może ich dalszy postęp udzieli nam odpowiedzi na pytania dotyczące natury umysłu oraz pozwoli lepiej zrozumieć różnego typu procesy poznawcze.

Działanie sztucznej inteligencji rodzi także cały szereg pytań natury etycznej. Czy maszyna może stanowić moralny podmiot działania? Pojawia się zagadnienie ewentualnej odpowiedzialności moralnej sztucznej inteligencji, a także odpowiedzialności uczonych zaangażowanych w prace nad nią. Sztuczna inteligencja – jako dziedzina nauki – implikuje zarówno problemy etyczne należące do innych obszarów badań naukowych, ale też takie, które są charakterystyczne wyłącznie dla niej. Paweł Łupkowski wyróżnia trzy grupy zarysowujących się tutaj problemów: (1) problemy dotyczące odpowiedzialności uczonych pracujących na polu sztucznej inteligencji; (2) problemy odnoszące

³⁷ Zob. R. Penrose, *Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*, tłum. P. Amsterdamski, T. Lanczewski, Zysk i S-ka Wydawnictwo, Poznań 2021.

³⁸ W. Branicki, *Autentyczność osobowa w pespektywie teorii dezintegracji pozytywnej*, „Studia Humanistyczne AGH” 11(2012) nr 3, s. 128.

³⁹ Por. R. Tadeusiewicz, *Awangarda sztucznej inteligencji – maszyny, które potrafią same tworzyć nowe pojęcia*, w: *Pojęcia. Jak reprezentujemy i kategoryzujemy świat*, red. J. Bremer, A. Chuderski, Universitas, Kraków 2011, s. 414.

się do oceny etycznej postępu naukowo-technologicznego; (3) zagadnienie przyszłego statusu inteligentnych maszyn obdarzonych świadomością⁴⁰. Nie podejmując tego niezwykle interesującego, ale także trudnego wątku, zauważyć należy, że dalszy rozwój sztucznej inteligencji wymagał będzie wzmoczonych działań ze strony etyków. Być może niezbędne stanie się nawet utworzenie odrębnej subdyscypliny w ramach etyki.

PROBLEMATYKA ŚWIADOMOŚCI SZTUCZNEJ INTELIGENCJI

W roku 1950 Turing zaproponował sposób rozstrzygnięcia, czy dane oprogramowanie (maszyna) jest inteligentne. Jego zdaniem za inteligentne należy uznać takie urządzenie, w odniesieniu do którego rozmawiający z nim człowiek nie jest w stanie rozpoznać, czy prowadzi dialog z maszyną, czy z innym człowiekiem. Pomijając długą historię dyskusji nad testem Turinga, zgodzić się należy z negatywną oceną uzyskanych za jego pomocą wyników. Kluczowy wydaje się argument natury filozoficznej, który w sformułowaniu Pawła Łupkowskiego brzmi następująco: „Idea tego typu argumentacji sprowadza się do wykazania, że nawet jeśli maszyna zdałaby test Turinga, to i tak nie moglibyśmy powiedzieć o niej, że jest inteligenta, ponieważ TT dostarcza z gruntu błędnej definicji posiadania inteligencji przez maszyny”⁴¹. Wspomniany autor odwołuje się do koncepcji Neda Blocke’a⁴², który zwraca uwagę, że test Turinga nie odróżnia kompetencji umysłu od jego zachowania. Umysł natomiast charakteryzuje jego kompetencje, a nie obserwowalne zachowania. Powyższy argument jest wersją „argumentu z drzewa konwersacji”⁴³. Zakłada on maksymalny czas trwania testu Turinga na jedną godzinę. Ma to umożliwić utworzenie wspomnianego drzewa konwersacji, zawierającego kompletną listę możliwych pytań i odpowiedzi, dających się sformułować w ciągu godziny. Inną jego odmianę stanowi argument „chińskiego pokoju” zaproponowany przez Johna R. Searle’a⁴⁴, również prowadzący do negatywnego wniosku odnośnie do możliwości przypisania inteligencji maszynie. Wydaje się zatem, że test Turinga nie jest w stanie dostarczyć przekonującej odpowiedzi na pytanie,

⁴⁰ Por. P. Ł u p k o w s k i, *Rola etyki i antropologii w rozważaniach o sztucznej inteligencji*, „Ethos” 18(2005) nr 1-2, s. 240.

⁴¹ T e n z e, *Test Turinga. Perspektywa sędziego*, Wydawnictwo Naukowe UAM, Poznań 2010, s. 32.

⁴² Por. N. B l o c k, *The Mind as the Software of the Brain*, w: *An Invitation to Cognitive Science*, t. 3, *Thinking*, red. E. Smith, D. Osherson, The MIT Press, Londyn, 1995, s. 377-425.

⁴³ Ł u p k o w s k i, *Test Turinga*, s. 32.

⁴⁴ Por. J.R. S e a r l e, *Minds, Brains, and Programs*, „Behavioral and Brain Sciences” 3(1980) nr 3, 417-457.

czy inteligentne maszyny są zdolne myśleć bądź czy cechuje je jakiś rodzaj świadomości. Sam Turing zresztą traktował sformułowany przez siebie test nie tyle jako sposób odpowiedzi na pytanie, czy maszyna jest w stanie myśleć, ile jako narzędzie do rozwiązywania skromniejszego problemu, a mianowicie, czy maszyna, biorąc udział w określonej grze, poradzi sobie równie dobrze jak człowiek⁴⁵. Pytanie o świadomość inteligentnych maszyn pozostaje zatem otwarte. Jego rozstrzygnięcie wymaga odpowiedzi na inne fundamentalne i niezwykle trudne pytanie: Czym jest świadomość?

Problematyka świadomości stanowi jedno z kluczowych zagadnień psychologii, nauk kognitywnych oraz filozofii. Świadomość to zjawisko trudne do zdefiniowania. Pytanie o jej istotę nurtowało wielu myślicieli na przestrzeni dziejów. Przedstawienie ich rozważań nie jest tutaj oczywiście możliwe. Warto jednak przynajmniej zwrócić uwagę na pewne aspekty filozoficznych analiz świadomości, które ujawniły się w trakcie trwającego wiele wieków dyskursu filozoficznego.

Znamienne jest, w jak różny sposób poszczególne szkoły filozoficzne traktowały fenomen świadomości i jak różne znaczenia przypisywały tej problematyce. Filozofowie starożytni i średniowieczni nie pisali o świadomości, lecz o duszy⁴⁶. Pojęcie świadomości odgrywało fundamentalną rolę w myśli Kartezjusza oraz wyrosłej na jej gruncie tradycji filozoficznej. Według Kartezjusza funkcją świadomości jest myślenie oraz wytwarzanie idei (łac. *cogitationes*). Myślenie, w ujęciu francuskiego filozofa, obejmuje zarówno akty poznawcze, jak i wolicjonalne oraz namiętności i emocje. Jest to najszersze możliwe ujęcie świadomości. Wymienione zjawiska psychiczne różnią się, według Kartezjusza, jedynie stopniem wysiłku jakiego wymaga ich pojawienie się w umyśle, traktowanym jako byt wytwarzający świadomość, której najwyższą formą jest samoświadomość. Świadomość zaś to jedna z dwóch podstawowych kategorii ontologicznych. Jej treścią są idee. Pojęcie „ja” zostaje przez Kartezjusza zredukowane do pojęcia świadomości, ducha, umysłu (łac. *res cogitans*)⁴⁷. Przeciwnieństwo świadomości – rzecz rozciąglą, materia (łac. *res extensa*) jest maszyną, mechanizmem.

W wywodzącej się od Franza Brentana (1838-1917) tradycji podkreśla się intencjonalny charakter przeżyć świadomych, czyli fakt, że są one zawsze

⁴⁵ Por. A.M. Turing, *Computing Machinery and Intelligence*, „Mind” 59(1950) nr 236, s. 434.

⁴⁶ Por. A. Revonsuo, S. Sajama, M. Kammipinen, *General Introduction: The Riddle of Consciousness*, w: *Consciousness in Philosophy and Cognitive Neuroscience*, red. A. Revonsuo, M. Kammipinen, Psychology Press, New York 2014, s. 5.

⁴⁷ Por. R. Descartes, *Medytacje o pierwszej filozofii. Wraz z zarzutami uczonych mężów i odpowiedziami autora*, tłum. M. Ajdukiewicz, K. Ajdukiewicz, Państwowe Wydawnictwo Naukowe, Warszawa 1958, s. 22.

skierowane na jakiś przedmiot. Rozważania nad istotą świadomości stanowią centralny punkt filozofii ucznia Brentana – Edmunda Husserla (1859-1938). Twórca fenomenologii akceptuje w początkowej fazie swojej twórczości Brentanowską tezę o intencjonalnym charakterze przeżyć świadomych⁴⁸. W późniejszym okresie koncepcja świadomości ulega w myśli Husserla licznym modyfikacjom. W *Ideach czystej fenomenologii i fenomenologicznej filozofii* pojawia się koncepcja tak zwanej czystej transcendentalnej świadomości⁴⁹. Tak rozumiana świadomość jest bytem absolutnym stanowiącym podstawę ontologiczną świata, w którym konstytuują się sensy wszystkich przedmiotów. Prawdopodobnie pojmowana tak świadomość nie może, ex definitione, przysługiwać inteligentnym maszynom.

Rozwijana przez tradycję brentanowską koncepcja intencjonalności świadomości znalazła kontynuację we współczesnej filozofii umysłu. Zdaje się do niej nawiązywać między innymi Searle, który wskazuje, że umysłu ludzkiego nie da się sprowadzić wyłącznie do czysto poznawczych – obliczeniowych funkcji. Odrzuca on mocną wersję koncepcji sztucznej inteligencji, zgodnie z którą komputerom można przypisać świadomość. Searle uważa, że komputery nie myślą, gdyż nie posiadają nastawienia intencjonalnego: „Ich myśli nie są myślami o czymś ani też [komputery – Z.O.] nie posiadają emocji i nie przeżywają uczuć”⁵⁰. Pojęcie intencjonalności jest także kluczowe w zaproponowanym przez Searle’a słynnym argumencie „chińskiego pokoju”.

Wskazane tu wybrane filozoficzne koncepcje świadomości nie są oczywiście jedynymi, jakie można odnaleźć w europejskiej myśli filozoficznej. Świadomość jest w nich traktowana jako byt sui generis, nieredukowalny do zjawisk fizycznych. Można także odnaleźć inne perspektywy (redukcjonistyczne), na przykład ujęcie Daniela Clementa Dennetta⁵¹. Traktuje się w nich świadomość jako tak czy inaczej rozumiany wytwór procesów natury fizjologicznej zachodzących w mózgu. Istotne staje się tutaj, w kontekście rozważań nad sztuczną inteligencją, rozróżnienie tego, co „naturalne”, od tego, co „sztuczne”. Anna Latawiec zwraca uwagę, że kluczowe w tej dyskusji jest opowiedzenie się za wyróżnionymi przez Searle’a mocną bądź słabą wersją sztucznej inteligencji⁵². Uznanie pierwszej z nich prowadzi do antropomorfizowania inteligentnych

⁴⁸ Zob. E. Husserl, *Badania dotyczące fenomenologii i teorii poznania*, cz. 2, tłum. J. Sidorek, Wydawnictwo Naukowe PWN, Warszawa 2000.

⁴⁹ Zob. tenże, *Idee czystej fenomenologii i fenomenologicznej filozofii*, t. 1, tłum. D. Gierulanka, Państwowe Wydawnictwo Naukowe, Warszawa 1975.

⁵⁰ Z. Piątek, *Spory o naturę świadomości a resuscytacja*, „Humanistyka i Przyrodznawstwo” 10(2004), s. 36.

⁵¹ Zob. D.C. Dennett, *Natura umysłów. Jak zrozumieć świadomość*, tłum. W. Turopolski, Copernicus Center Press, Kraków 2021.

⁵² Por. A. Latawiec, *Naturalne i sztuczne w świetle rozważań o sztucznej inteligencji*, „Studia Philosophiae Christianae” 32(1996) nr 1, s. 192.

maszyn, akceptacja drugiej może implikować negację możliwości istnienia sztucznej inteligencji w ogóle. Nie rozstrzygając, czy opisywany spór daje się sprowadzić wyłącznie do werbalnej jego postaci, uznać należy, że ustalenia terminologiczne są ważną jego częścią.

Filozoficzne ujęcia świadomości cechuje wysoki stopień ogólności oraz brak odniesień do danych empirycznych. Podstawowym problemem dotyczącym zagadnienia sztucznej inteligencji jest odpowiedź na pytanie, czy i jak świadomość może wyłaniać się z materii⁵³. I tutaj rozważania filozoficzne mogą okazać się pomocne. Negatywna odpowiedź na powyższe pytanie zdaje się wykluczać możliwość mocnej wersji sztucznej inteligencji. Próbę systematyki występujących w filozofii koncepcji świadomości przeprowadził Jakub Jonkisz. Została ona dokonana w oparciu o cztery kryteria: rodzaj epistemiczny, rząd semantyczny, typ pragmatyczny oraz stan fizjologiczny. W jej wyniku powstała tabela zawierająca dwanaście wyróżnionych typów świadomości⁵⁴. Pomimo mogących pojawiać się wątpliwości, dotyczących chociażby takiego, a nie innego wyboru kryteriów podziału, można mieć wrażenie, że próby uporządkowania problematyki świadomości podejmowane zarówno w filozofii, jak i naukach szczegółowych, należy uznać za wartościowe. Nie jest to jedyna próba systematyzacji filozoficznych koncepcji świadomości. Joseph Levine za pierwotny w różnych ujęciach świadomości uznaje podział na koncepcje materialistyczne i antymaterialistyczne⁵⁵. Pomimo dominującego w ostatnim okresie interdyscyplinarnego podejścia do problematyki świadomości, podział ten najlepiej odzwierciedla ogólne ramy prowadzonych od stuleci filozoficznych rozważań nad świadomością.

W przypadku świadomości mamy do czynienia z interesującą sytuacją poznawczą. Otóż, jak dotychczas ani filozofom, ani też przedstawicielom nauk szczegółowych podejmującym tę problematykę nie udało się zgłębić natury świadomości. Dopóty to nie nastąpi, wszelkie dyskusje dotyczące świadomości inteligentnych maszyn wydają się bezprzedmiotowe, a w każdym razie dalekie od jednoznacznych rozstrzygnięć. Uzasadnieniem powyższej tezy może być fakt istnienia wielu rozmaitych, często bardzo od siebie różnych, filozoficznych koncepcji świadomości. Nie udało się także, jak dotąd, sformułować powszechnie akceptowanej definicji świadomości.

⁵³ Por. E. H i l d t, *Artificial Intelligence: Does Consciousness Matter?*, „Frontiers in Psychology” 10(2019), s. 1535. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01535/full>.

⁵⁴ Por. J. J o n k i s z, *Pojęcie świadomości w kognitywistyce i filozofii umysłu – próba systematyzacji*, „Filozofia Nauki” 20(2012) nr 2, s. 55.

⁵⁵ Por. J. L e v i n e, *Recent Work on Consciousness*, „American Philosophical Quarterly” 34(1997) nr 4, 379-404.

*

Sztuczna inteligencja jest jednym z najodważniejszych zamierzeń umysłu ludzkiego. Pomimo faktu, że sam termin pojawił się stosunkowo późno, bo dopiero w połowie dwudziestego wieku, wyrosła ona z wielowiekowej refleksji nad naturą poznania, a w szczególności stanowi realizację pragnienia przejawianego przez wielu myślicieli na przestrzeni wieków: ujęcia ludzkiego rozumowania w ramy logicznego formalizmu, jego algorytmizację. Idea ta traktowana początkowo w kategoriach bardziej marzenia niż naukowego projektu, w miarę doskonalenia narzędzi logicznych, wyłaniała się z mroków spekulacji, aby w drugiej połowie dwudziestego stulecia znaleźć swoją egzemplifikację w postaci inteligentnych maszyn.

Wraz z rozwojem sztucznej inteligencji rodzą się pytania natury filozoficznej. Dotyczą one przede wszystkim natury umysłu, świadomości oraz informacji. Pojawia się także pytanie o ograniczenia poznawcze uzyskiwane sztucznie za pomocą cyfrowych technik przetwarzania informacji. Mogą one być związane na przykład z odkryciem tak zwanych liczb nieobliczalnych. Są to „liczby niewymierne, których przedstawienia dziesiętnego nie może wyznaczyć, z dowolną zadaną dokładnością, żaden układ do obliczeń mechanicznych, zwany dziś maszyną Turinga”⁵⁶. Istnieje możliwość, że pewne rozumowania intuicyjne nie mogą zostać poddane mechanizacji, ponieważ nie mają one reprezentacji w liczbach obliczalnych. Dyskusja w tym temacie trwa i trudno tutaj formułować jakieś ostateczne wnioski. Warto jednak zauważyć, że wykazanie istnienia w świecie fizycznym wielkości reprezentowanych przez liczby nieobliczalne nakładałoby istotne ograniczenia natury epistemologicznej na możliwości sztucznej inteligencji. Być może pewnym wyjściem z impasu stałoby się wynalezienie komputerów kwantowych, lecz jest to wciąż kwestia przyszłości.

Jeżeli przyjmiemy, że informacja jest podstawowym elementem konstytuującym zarówno świat biologiczny, jak i technologiczny, wówczas uznać możemy ją za kategorię filozoficzną. Mieczysław Lubański zauważa, że pojęcie informacji przeszło znamiennej ewolucję. W jej wyniku ukonstytuowało się ogólnonaukowe oraz odpowiednio filozoficzne pojęcie informacji⁵⁷. Istotą wspomnianej ewolucji jest poszerzanie zakresu pojęcia informacji, z jednoczesnym zubożaniem jego treści. Staje się to bardziej zrozumiałe, gdy uświadomimy sobie, że powstanie i rozwój sztucznej inteligencji jest zjawiskiem

⁵⁶ P. S t a c e w i c z, *Liczby nieobliczalne a granice kodowania w informatyce*, „Studia Semiotyczne” 32(2018) nr 2, s. 137.

⁵⁷ Por. M. L u b a ń s k i, *Informacja i jej nośniki*, „Studia Philosophiae Christianae” 16(1980) nr 2, s. 56.

interdyscyplinarnym. Zadanie filozofów polegałoby przede wszystkim na dostarczeniu ram pojęciowych niezbędnych chociażby do zrozumienia podstawowych pojęć, takich jak: inteligencja, umysł, świadomość czy informacja.

Projekt sztucznej inteligencji jest zamysłem o charakterze technologicznym, ale wyrasta on z ducha refleksji filozoficznej nad podstawowymi zagadnieniami umysłu ludzkiego, mechanizmami jego funkcjonowania i możliwościami symulowania jego działania przez maszyny. Refleksja filozoficzna nie jest tylko źródłem procesu, który doprowadził do powstania sztucznej inteligencji. Jej rozwój niesie ze sobą nowe problemy natury filozoficznej, lecz także rzuca pewne światło na te, które rozważane są niemal od początku europejskiej filozofii. Zakres problematyki filozoficznej, jaki zarysowuje się w kontekście badań nad sztuczną inteligencją, jest bardzo szeroki. Obejmuje on między innymi filozofię umysłu, neurofilozofię, filozofię języka, etykę i epistemologię procesów poznawczych.

Jeżeli zaakceptujemy tezę, że problematyka filozoficzna wpisana jest niejako w naturę ludzką, wówczas także sztuczna inteligencja, jako wytwór człowieka, nie może być od tej problematyki wolna. Rozwój sztucznej inteligencji stawia zatem nowe wyzwania także przed filozofami. Uzasadnienie tej tezy może być dwojakie. Po pierwsze historyczne: pamiętać należy, że powstanie nauki poprzedzone było pojawieniem się refleksji filozoficznej, z której nauka wyrosła. Akceptacja poglądu, uznającego poznanie naukowe za jeden z wyznaczników ludzkiego sposobu bycia w świecie, implikuje zgodę na uznanie myśli, głoszącej, że filozofia stanowi jeden z przejawów natury ludzkiej. Drugi argument wynika z faktu, że pomimo rozwoju wiedzy naukowej i jej specjalizacji, problemy filozoficzne nie tylko nie są eliminowane z pola zainteresowań badaczy, lecz w wielu przypadkach rozwój nauki przyczynia się do wyłaniania się nowych. Istnieje także w człowieku potrzeba poddawania filozoficznym analizom szczegółowych wyników uzyskiwanych na gruncie rozmaitych dziedzin poznania naukowego. Jest tak również, jak się zdaje, w przypadku badań nad sztuczną inteligencją.

BIBLIOGRAFIA / BIBLIOGRAPHY

- Block, Ned. "The Mind as the Software of the Brain." In *An Invitation to Cognitive Science*. Vol. 3. *Thinking*. Edited by Edward E. Smith and Daniel N. Osherson. London: The MIT Press, 1995.
- Boyer, Carl. *Historia rachunku różniczkowego i całkowego i rozwój jego pojęć*. Translated by S. Dobrzycki. Warszawa: Państwowe Wydawnictwo Naukowe, 1964.
- Branicki, Wacław. "Autentyczność osobowa w pespektywie teorii dezintegracji pozytywnej." *Studia Humanistyczne AGH* 11, no. 3 (2012): 127–40.

- Breidbach, Olaf, and Michael T. Ghiselin. "Athanasius, Kircher (1602-1680) on Noah's Ark: Baroque 'Intelligent Design' Theory." *Proceedings of the California Academy of Sciences* 57, no. 36 (2006): 991–1002.
- Buchanan, Bruce G. "A (Very) Brief History of Artificial Intelligence." *AI Magazine* 26, no. 4 (2005).
- Descartes, René. *Medytacje o pierwszej filozofii: Wraz z zarzutami uczonych mężów i odpowiedziami autora*. Translated by Maria Ajdukiewicz and Kazimierz Ajdukiewicz. Warszawa: Państwowe Wydawnictwo Naukowe, 1958.
- Dennett, Daniel C. *Natura umysłów: Jak zrozumieć świadomość*. Translated by Witold Turopolski. Kraków: Copernicus Center Press, 2021.
- Duch, Włodzisław. "Umysł, świadomość i działanie twórcze." *Kognitywistyka i Media w Edukacji*, nos. 1–2 (2008): 9–37.
- Findlen, Paula. *Athanasius Kircher: The Last Man Who Knew Everything*. New York and London: Routledge, 2004.
- Flasiński, Mariusz. "History of Artificial Intelligence." In Flasiński, *Introduction to Artificial Intelligence*. Cham: Springer, 2016.
- Gottlob, Frege. "Ideografia: Język formalny czystego myślenia wzorowany na języku arytmetyki." In Franz Brentano, Gottlob Frege, and Christian Thiel, *Próby gramatyki filozoficznej: Antologia*. Edited and translated by Krzysztof Rotter. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 1997.
- Grządziel, Mariusz. *Filozofia, teologia i matematyka w pismach Mikołaja z Kuzy – szkic poglądów*, 2008. https://www.chfnp.pl/new05/new05_grzadzziel.pdf.
- Heller, Michał. "Chrześcijański naturalizm." *Roczniki Filozoficzne* 51, no. 3 (2003): 41–58.
- Hetmański, Marek. "Maszyna Turinga a umysł ludzki." *Annales Universitatis Mariae Curie-Skłodowska. Sectio I. Philosophia-Sociologia*, no. 23 (1988): 155–77.
- Hildt, Elisabeth. "Artificial Intelligence: Does Consciousness Matter?" *Frontiers in Psychology*, no. 10 (2019): 1535. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01535/full>.
- Hobbes, Thomas. *Lewiatan, czyli materia, forma i władza państwa kościelnego i świeckiego*. Translated by Czesław Znamierowski. Warszawa: Fundacja Aletheia, 2009.
- Husserl, Edmund. *Badania dotyczące fenomenologii i teorii poznania*. Part 2. Translated by Janusz Sidorek. Warszawa: Wydawnictwo Naukowe PWN, 2000.
- . *Idee czystej fenomenologii i fenomenologicznej filozofii*. Vol. 1. Translated by Danuta Gierulanka. Warszawa: Państwowe Wydawnictwo Naukowe, 1975.
- Jastrzębski, Hubert, and Szymon Chlebowski. "Frege i Łukasiewicz o wartościach logicznych." *Hybris*, no. 27 (2014): 121–35.
- Jonkisz, Jakub. "Pojęcie świadomości w kognitywistyce i filozofii umysłu – próba systematyzacji." *Filozofia Nauki* 20, no. 2 (2012): 29–55.
- Kaul, Vivek, Sarah Enslin, and Seth A. Gross. "History of Artificial Intelligence in Medicine." *Gastrointestinal Endoscopy* 92, no. 4 (2020): 807–12.

- Latawiec, Anna. "Naturalne i sztuczne w świetle rozważań o sztucznej inteligencji." *Studia Philosophiae Christianae* 32, no. 1 (1996): 191–202.
- Laita, Luis M., et al. "George Boole, a Forerunner of Symbolic Computation." In *Artificial Intelligence and Symbolic Computation: International Conference AISC 2000 Madrid, Spain, July 2000 Revised Papers*. Edited by John A. Campbell and Eugenio Roanes-Lozano. Berlin and Heidelberg: Springer, 2001.
- Leibniz, Gottfried Wilhelm. "Przedmowa do nauki ogólnej." In Leibniz, *Wyznanie wiary filozofa*. Translated by Stanisław Cichowicz. Warszawa: Wydawnictwo Naukowe PWN, 1969.
- Levine, Joseph. "Recent Work on Consciousness." *American Philosophical Quarterly* 34, no. 4 (1997): 379–404.
- Lubański, Mieczysław. "Informacja i jej nośniki." *Studia Philosophiae Christianae* 16, no. 2 (1980): 55–66.
- Łupkowski, Paweł. "Rola etyki i antropologii w rozważaniach o sztucznej inteligencji." *Ethos* 18, nos. 1-2 (2005): 239–41.
- Łupkowski, Paweł. *Test Turinga. Perspektywa Sędziego*. Poznań: Wydawnictwo Naukowe UAM, 2010.
- Maciaszek, Janusz. "Od matematyki do analizy języka naturalnego: Kilka uwag o filozofii języka Gottloba Fregego." *Językoznawstwo: Współczesne badania, problemy i analizy językoznawcze* 7 (2013): 57–75.
- Marciszewski, Witold, and Roman Murawski. *Mechanization of Reasoning in a Historical Perspective*. Amsterdam and Atlanta: Editions Rodopi, 1995.
- McCarthy, John, et al. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955." *AI Magazine* 27, no. 4 (2006): 12–14.
- Nowaczyk, Adam. *Wprowadzenie do filozofii matematyki*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 2019.
- von Neumann, John, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.
- Penrose, Roger. *Nowy umysł cesarza: O komputerach, umyśle i prawach fizyki*. Translated by Piotr Amsterdamski and Tomasz Lanczewski. Poznań: Wydawnictwo Zysk i S-ka, 2021.
- Piątek, Zdzisława. "Spory o naturę świadomości a resuscytacja." *Humanistyka i Przyrodoznawstwo* 10 (2004): 29–38.
- Piecuch, Aleksander. "Zaniedbana algebra a nauczanie informatyki." *Edukacja—Technika—Informatyka* 8, no. 3 (2017): 288–94.
- Revonsuo, Antti, Seppo Sajama, and Matti Kamppinen. „General Introduction: The Riddle of Consciousness.” In *Consciousness in Philosophy and Cognitive Neuroscience*. Edited by Antti Revonsuo and Matti Kamppinen. New York: Psychology Press, 2014.
- Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–57.
- Shirley, John W. "Binary Numeration before Leibniz." *American Journal of Physics* 19, no. 8 (1951): 452–54.

- Sparavigna, Amelia Carolina. *Giovanni, de la Fontana, Engineer and Magician*. <https://arxiv.org/abs/1304.4588>.
- Stacewicz, Paweł. "Liczby nieobliczalne a granice kodowania w informatyce." *Studia Semiotyczne* 32, no. 2 (2018): 131–52.
- Stacewicz, Paweł. "Światopogląd informatyczny. Naukowe podstawy i filozoficzne perspektywy." In *Informatyka a filozofia: Od informatyki i jej zastosowań do światopoglądu informatycznego*. Edited by Paweł Stacewicz. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2015.
- Steffens, Oliver. *Leibniz, Llull and the Logic of Truth: Precursors of Artificial Intelligence*, Regensburg: Ostbayerische Technische Hochschule Regensburg, 2020.
- Tadeusiewicz, Ryszard. "Awangarda sztucznej inteligencji – maszyny, które potrafią same tworzyć nowe pojęcia." In *Pojęcia: Jak reprezentujemy i kategoryzujemy świat*. Edited by Józef Bremer i Adam Chuderski. Kraków: Universitas, 2011.
- Trzęsicki, Kazimierz. "Idea sztucznej inteligencji." *Filozofia i Nauka* 8, no. 1 (2020): 69–96.
- . "Leibnizjańskie inspiracje informatyki." *Filozofia Nauki* 14, no. 3(55) (2006): 21–48.
- Turing, Alan Mathison. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433–60.
- . "On Computable Numbers, with an Application to the Entscheidungsproblem." *Journal of Mathematics* 58, no. 5 (1936): 345–63.
- Vermeir, Koen. "Athanasius Kircher's Magical Instruments: An Essay on 'Science', 'Religion' and Applied Metaphysics." *Studies in History and Philosophy of Science Part A* 38(2007) no. 2: 363–400.
- Żegleń, Urszula. "Wielowymiarowość umysłu." *Filozofia Nauki* 12, no. 3-4 (2004): 153–70.

ABSTRAKT /ABSTRACT

Zbigniew ORBIK – W poszukiwaniu filozoficznych źródeł problematyki sztucznej inteligencji

DOI 10.12887/36-2023-3-143-05

Celem artykułu jest poszukiwanie filozoficznych korzeni idei sztucznej inteligencji. Dokonano w nim przeglądu długiego procesu rozwoju pojawiających się w myśli filozoficznej już od czasów starożytnych idei, które uznać można za antecedensy współczesnej koncepcji sztucznej inteligencji. Wprawdzie sztuczna inteligencja rozumiana jest obecnie jako rodzaj inteligencji, którą posiadają maszyny będące w stanie symulować procesy myślenia charakterystyczne dla człowieka i jest dziedziną zaliczaną do informatyki i nauk kognitywnych, okazuje się jednak, że ma ona swoje źródło w refleksji filozoficznej. Koncepcja ta rodzi także wiele problemów natury filozoficznej, które w niniejszym tekście zostały, z oczywistych względów, jedynie zasygnalizowane. Są to zagadnienia

obejmujące swoim zakresem niemal cały przekrój problematyki tradycyjnie zaliczanej do filozofii. Dotyczą one między innymi struktury rozumowań, możliwości ich automatyzacji, natury procesów poznawczych, możliwości konstrukcji języków formalnych czy ontologicznej natury umysłu. Jednym z najważniejszych jest zagadnienie świadomości. W artykule zastosowano analizę wybranej, dostępnej literatury dotyczącej historii sztucznej inteligencji oraz jej filozoficznych aspektów. Oprócz przeglądu literatury autor stosuje metodę analizy i konstrukcji logicznej. Rozwój sztucznej inteligencji implikuje także konieczność dalszych badań filozoficznych nad naturą umysłu, istotą świadomości, procesów poznawczych oraz licznych kwestii etycznych. Filozoficzna refleksja nad sztuczną inteligencją wpisuje się w starożytny postulat samopoznania człowieka.

Słowa kluczowe: nauki kognitywne, sztuczna inteligencja, umysł, świadomość

Kontakt: Katedra Stosowanych Nauk Społecznych, Wydział Organizacji i Zarządzania, Politechnika Śląska, ul. Roosevelta 26-28, 41-800 Zabrze
E-mail: zbigniew.orbik@polsl.pl
Tel. 721 433 549
ORCID: 0000-0001-7092-1567

Zbigniew ORBIK, In Search of Philosophical Sources of the Problem of Artificial Intelligence

DOI 10.12887/36-2023-3-143-05

The aim of the article is to inquire into the philosophical origins of the idea of artificial intelligence. The author provides a review of the long process of developing ideas which have emerged in philosophical thought since antiquity and can be regarded as antecedents to the modern concept of artificial intelligence. Currently understood as a type of intelligence possessed by machines capable of stimulating human-specific thinking processes and, as a research area, considered as part of computer science and cognitive sciences, artificial intelligence has its source in philosophical reflection. Also, the concept of artificial intelligence poses numerous philosophical problems, which, for obvious reasons, were only mentioned in the article. They are issues that cover almost the entire cross-section of problems traditionally considered as philosophical and concern, among others, the structures of reasoning, the possibilities of their automation, the nature of cognitive processes, the possibilities of constructing formal languages or the ontological nature of mind. One of the most important question is that of consciousness. The article analyzes selected available literature on the history of artificial intelligence and its philosophical aspects. Apart from the literature review, the author uses the method of analysis and logical construction. The development of artificial intelligence also implies the need for further philosophical research on the nature of mind, the essence of consciousness, cognitive processes, and a vast range of emerging ethical issues.

Philosophical reflection on artificial intelligence meets the ancient postulate of human self-knowledge.

Keywords: cognitive science, artificial intelligence, mind, consciousness

Contact: Department of Applied Social Sciences, Faculty of Organization and Management, Silesian University of Technology, ul. Roosevelta 26-28, 41-800 Zabrze, Poland

E-mail: zbigniew.orbik@polsl.pl

Phone: +48 721 433 549

ORCID: 0000-0001-7092-1567